Université de Liège

Faculté de Philosophie et Lettres

Département de Langues et Littératures françaises et romanes

UR Traverses – Centre de Sémiotique et Rhétorique

Chiffrer et déchiffrer les larges ensembles d'images

Contribution à l'étude sémio-rhétorique des stratégies de communication en ligne et des applications du traitement informatique d'images

Thèse présentée par M. Cédric HONBA HONBA en vue de l'obtention du titre de Docteur en Langues, Lettres et Traductologie sous la direction de Mme Maria Giulia DONDERO

Membres du jury : MM. Björn-Olav DOZO (ULiège), Stéphane POLIS (FNRS/ULiège), Everardo REYES-GARCIA, (Université Paris 8) et Peter STOCKINGER (Institut national des langues et civilisations orientales)

Année académique 2023 - 2024

Les recherches qui ont permis de rédiger cette thèse de doctorat ont été financées par le Fonds de la recherche scientifique (FRS-FNRS) dans le cadre du PDR FNRS de Maria Giulia DONDERO intitulé : *Les Big Visual Data. Approche sémiotique de l'indexation et de la visualisation des archives d'images* (*Big Visual Data. A Semiotic Approach to Indexing and Visualizing Image Archives*).

Remerciements

D’emblée, ma gratitude va à ma promotrice, Mme Maria Giulia Dondero, qui m’a accordé sa confiance depuis mon travail de fin d’études jusqu’au dépôt de ma thèse. En plus du profond respect que j’éprouve au regard de la rigueur et de la qualité de ses recherches, je suis surtout reconnaissant d’avoir pu rencontrer cette personnalité chaleureuse, toujours en mouvement, qui a encadré mon parcours. Encore une fois, mille mercis.

Je tiens à remercier les membres du jury : MM. Björn-Olav Dozo, Stéphane Polis, Everardo Reyes-Garcia et Peter Stockinger, ainsi que M. Nicolas Mazziotta. Leurs suggestions ont largement contribué à la fois au contenu et à la forme du présent travail. De plus, la disponibilité et la sympathie de ce comité, mais aussi la franchise de chacun des membres, ont favorisé un contexte où je ne pouvais que m’améliorer.

Bien entendu, mes remerciements vont aussi à mes collègues du Centre de Sémiotique et Rhétorique : Enzo d’Armenio, Sémir Badir, Alexandre Lansmans, Ingrid Mayeur, François Provenzano et Élise Schürgers. Discuter et travailler avec eux est un réel plaisir.

Je souhaite en passant exprimer ma gratitude au département Médias, Cultures et communication et au département de Langues et Littératures françaises et romanes de la Faculté de Philosophie et Lettres (ULiège). Le premier pour la formation, riche et variée, qui est enseignée par les professeurs, les chercheurs et les assistants de ce département. Le second pour m’avoir permis d’y réaliser ma recherche doctorale.

Ma reconnaissance s’exprime également pour mes camarades de l’université, en particulier Fabio Recchia et Pablo Luca, deux excellents philosophes avec lesquels j’ai passé de très bons moments, qu’il s’agisse de réflexions scientifiques ou d’authentiques rigolades. De la même façon, je tiens à saluer et à remercier mes amis proches ; ces derniers se reconnaîtront sans que j’évoque leurs noms.

Mes derniers remerciements s’adressent à ma famille et ma belle-famille, qui ont été un soutien moral constant pendant ces dernières années, et surtout à Nora, *die Liebe meines Lebens*.

Sommaire

INTRODUCTION **13**

PREMIÈRE PARTIE. SÉMIOTIQUE DE LA COMMUNICATION EN LIGNE : ÉNONCIATION ET STRATÉGIES MULTIMÉDIAS **25**

CHAPITRE 1 POUR UNE APPROCHE SÉMIO-RHÉTORIQUE DES ARCHIVES ET DES COLLECTIONS D'IMAGES NUMÉRIQUES **27**

1.1. Accroissement des archives et des collections d'images **28**

1.2. Traitement documentaire des images : conserver, décrire, classer **32**

1.3. Point de vue sémiotique sur l'archive **36**

1.4. Approche théorique du support numérique **57**

1.5. Rhétorique et pragmatique de la communication en ligne **67**

1.6. Conclusions **73**

CHAPITRE 2 STRATÉGIES DE COMMUNICATION EN LIGNE D'IMAGES ET DÉMATÉRIALISATION DE LA CULTURE VISUELLE **77**

2.1. Humanités numériques : une réaction transdisciplinaire **78**

2.2. Archives d'images numériques et musée virtuel **81**

2.3. Patrimoine numérique : tensions entre technique et interprétation **93**

2.4. Conclusions **109**

DEUXIÈME PARTIE. ANALYSES D'IMAGES ET MÉTHODES ASSISTÉES PAR UN ORDINATEUR **113**

CHAPITRE 3 PROGRÈS DES MÉTHODES D'ANALYSE D'IMAGES EN HISTOIRE DE L'ART ET EN SCIENCES DU LANGAGE **115**

3.1. De l'histoire de l'art aux sciences des images **116**

3.2. Sciences du langage et interprétation d'images **137**

3.3. Indexation d'images et limites du principe de métalangage de description **155**

3.4. Revenir à l'analyse des formes visuelles **160**

3.5. Conclusions **167**

CHAPITRE 4 IMAGES NUMÉRIQUES, NOUVEAUX MÉDIAS ET ANALYSE INFORMATIQUE D'IMAGES 171

4.1. De l'analyse d'images à l'analyse d'images numériques : quelle différence ? 172

4.2. *Computer vision*, *data visualization* et *deep learning* 193

4.3. Conclusions 213

TROISIÈME PARTIE. ANALYSE DU CORPUS (1) : BASES D'IMAGES NUMÉRIQUES ET EXPOSITIONS VIRTUELLES 217

CHAPITRE 5 PREMIÈRE TENDANCE STRATÉGIQUE : LES BASES D'IMAGES NUMÉRIQUES DES INSTITUTIONS PATRIMONIALES 219

5.1. Des archives et des collections d'images aux bases d'images numériques 220

5.2. L'interface Web en tant qu'espace graphique et multimédia 226

5.3. Fonctionnalités des moteurs de recherche et présentation des résultats 238

5.4. Conclusions 252

CHAPITRE 6 RENOUVELLEMENT DES BASES D'IMAGES NUMÉRIQUES (1) : PARTICIPATION ET COOPÉRATION 255

6.1. Médiathèques en ligne sur le modèle de l'encyclopédie 256

6.2. Agglomérer et standardiser les bases d'images numériques 281

6.3. Conclusions 302

CHAPITRE 7 RENOUVELLEMENT DES BASES D'IMAGES NUMÉRIQUES (2) : STRATÉGIES DE L'INTERACTIVITÉ 305

7.1. Des bases d'images numériques aux expositions virtuelles 306

7.2. Renouvellement d'un service Web à travers les applications interactives 320

7.3. Autres exemples de communication en ligne dans une perspective interactive 334

7.4. Conclusions 342

QUATRIÈME PARTIE. ANALYSE DU CORPUS (2) : CULTURE VISUELLE ET *BIG VISUAL DATA* 345

CHAPITRE 8 TROISIÈME TENDANCE STRATÉGIQUE : COMMUNICATION EN LIGNE ET VISUALISATION D'IMAGES 347

8.1. La *data visualization* au service des recherches en sciences humaines 349

8.2. De l'observation à l'expérimentation : *computer vision* et *media visualization* 364

8.3. Conclusions 375

CHAPITRE 9 DE LA COMMUNICATION EN LIGNE D'IMAGES À L'ANALYSE INFORMATIQUE DES FORMES VISUELLES 377

9.1. *Media visualization* et analyse des formes visuelles 379

9.2. Autres projets qui se basent sur la *media visualization* et sur les algorithmes 405

9.3. Conclusions 415

CONCLUSION 419

RÉFÉRENCES BIBLIOGRAPHIQUES ET WEBOGRAPHIE 443

TABLES DES ILLUSTRATIONS 469

INDEX 475

TABLE DES MATIÈRES 477

Liste des abréviations et des sigles

Chap. = Chapitre(s)

Sect. = Section(s)

Fig. = Figure(s)

API = Application Programming Interface

ConvNet = Convolutional Neural Networks

CSS = Cascading Style Sheets

HTML = Hypertext Markup Language

IIIF = International Image Interoperability Framework

ISO = International Organization for Standardization

JPEG = Joint Photographics Expert Groupe

OSI = Open Systems Interconnection

SVS = Semantic Vector Space

SGBD = Système de gestion de base de données

t-SNE = t-distributed stochastic neighbor embedding

URL = Uniform Resource Locator

BALaT = Belgian Art Links and Tools

EAT = Experiment in Art and Technology

GSG = Getty Search Gateway

IRPA = Institut royal du patrimoine artistique

MAP = Médiathèque de l'architecture et du patrimoine

POP = Plateforme ouverte du patrimoine

WC = WikimediaCommons

WF = Wikimedia Foundation

« [Les images] considèrent le monde selon deux directions opposées. Elles oscillent dans une zone située au-dessus de l'univers des choses "pratiques" et au-dessous des forces désincarnées qui animent ces choses. Elles servent en quelque sorte de charnière entre les deux. »

> Rudolf ARNHEIM, *La pensée visuelle*, tr. fr. de Claude NOËL et Marc LE CANNU, Paris, Flammarion, coll. « Champs », 1976 [1969], p. 142.

« [Représenter] signifie se présenter représentant quelque chose et toute représentation, tout signe ou procès représentationnel comprend une double dimension – dimension réflexive, se présenter ; dimension transitive, représenter quelque chose – et un double effet : l'effet de sujet et l'effet d'objet. Sémanticiens et pragmaticiens contemporaines les nommeront, et ce n'est point un hasard, "opacité" et "transparence". Inutile de dire que toute la fantasmatique de la description et de la mimésis s'est édifiée sur la dimension transitive de la représentation (représenter quelque chose) par oubli de son opacité réflexive et de ses modalités (se présenter) : une mimétique de l'image de peinture ouvre ainsi la voie des mots aux images parce que les images des choses (en peinture) sont déjà les noms des choses (en langage), parce qu'une description nominale des choses est déjà inscrite dans l'image. »

> Louis MARIN, *De la représentation*, Paris, Seuil/Gallimard, coll. « Hautes Études », 1994, p. 255.

INTRODUCTION

La communication en ligne des archives d'images et des collections d'œuvres d'art visuel est actuellement en plein essor. En effet, plusieurs institutions patrimoniales ont entrepris la *numérisation* et la *diffusion en ligne* des images qu'elles conservent. Le but est d'intégrer des copies de ces images dans un environnement informatique qui permet aux internautes de découvrir une grande quantité d'images depuis leurs ordinateurs personnels. Pourtant, les conséquences de cette numérisation des archives et des collections d'images commencent seulement à apparaître. L'une des conséquences les plus évidentes est qu'un système informatique traite uniquement des valeurs numériques et non pas les images telles que nous les regardons. Par conséquent, il faut que les images changent de support, c'est-à-dire qu'elles sont préalablement converties en données numériques (*digital data*). Cette conversion des images en nombres est la condition *sine qua non* pour appliquer différents types d'opérations de traitement informatiques sur des images.

D'un autre côté, des progrès dans plusieurs domaines du traitement informatique d'images aboutissent en outre sur des *méthodes d'analyse d'images* qui intéressent les institutions patrimoniales et les spécialistes des images et du visuel. Ces méthodes sont théorisées et développées par les recherches en *image processing*, en *data visualization* et en *computer vision*, ou encore dans le cadre des *software studies* et de la *digital art history*. Dans certains cas, la collaboration des institutions patrimoniales avec des spécialistes dans ces domaines aboutit sur des expérimentations intéressantes dont le but

est de proposer des services Web pour visualiser des archives et des collections d'images. En plus des bases d'images numériques et des expositions virtuelles, d'autres services sont conçus sur le modèle de la *media visualization*. Contrairement à d'autres types de visualisation (par exemple, les dispositifs qui représentent les données à l'aide de symboles qui sont organisés dans des diagrammes ou des réseaux), la *media visualization* vise quant à elle à construire des espaces de représentation qui intègrent des reproductions numériques des médias analysés.

Notre recherche vise à rassembler plusieurs travaux et projets qui cherchent à éclairer deux questions. Quel est l'impact de la numérisation et de la diffusion en ligne des images sur les « études visuelles » (*visual studies*), c'est-à-dire sur les disciplines qui étudient des objets liés à une « culture visuelle » ? Quels sont les usages possibles des méthodes informatiques d'analyse et de visualisation d'images ? En effet, l'utilisation de ces instruments informatiques dans les recherches sur les images et le visuel suscite encore des débats, car le fait que les opérations d'analyse sont appliquées sur des nombres, et non pas sur les supports originaux des images, peut paraître déconcertant pour les spécialistes des images et du visuel.

Afin de contribuer aux réflexions sur la numérisation et l'analyse informatique des images, l'objectif du présent travail est double.

1. Étant donné que l'accès aux archives et aux collections d'images s'adapte à des pratiques numériques où différents services Web mettent à disposition une grande quantité de copies numériques de ces images, le premier objectif est de *montrer que ces pratiques s'appuient sur une nouvelle forme de médiation* : la *communication en ligne d'images*. Nous proposons d'étudier cette forme de médiation à l'aide d'une approche « sémio-rhétorique » qui allie des travaux sur la sémiotique des pratiques et de l'énonciation avec des travaux sur les stratégies de diffusion du patrimoine numérique. En outre, plusieurs travaux en sémiotique mettent en évidence que la communication en ligne d'images met en scène un processus d'interaction qui construit une relation énonciative entre l'utilisateur et le service. Cette interaction dépend à la fois des caractéristiques spécifiques du support numérique et des choix stratégiques des producteurs des dispositifs de communication en ligne qui encadrent l'expérience des utilisateurs de ces services. Nous défendons l'hypothèse que l'observation de ces choix fait ressortir *trois tendances stratégiques* et *quatre types de construction d'un utilisateur*.

2. Un système informatique n'analyse pas des images de la même manière qu'un spécialiste des images. Un système informatique traite seulement une transformation mathématique d'une image en un ensemble de *pixels* (*picture elements*). Chaque pixel constituant la surface de l'image est associé à des valeurs numériques qui sont traitées par l'ordinateur. Pour éclairer ce changement, le second objectif est de confronter les *méthodes d'analyse d'images* (l'humain interprète un corpus d'images, une culture visuelle, etc.) aux *méthodes assistées par un ordinateur* (un système informatique analyse des données numériques qui représentent des images). Le but de cette confrontation est de mettre en évidence les limites et les opportunités que l'analyse informatique apporte à deux approches qui étudient habituellement des images (l'iconologie et la sémiotique). Ainsi, notre deuxième hypothèse est de *montrer que la communication en ligne est progressivement devenue à la fois un moyen de valorisation des archives et des collections d'images et un terrain d'expérimentation des performances de l'analyse informatique d'images*. Cette situation impose, selon nous, un dialogue transdisciplinaire que nous voulons défendre et soutenir.

Ce double objectif constitue les deux axes principaux du présent travail. En effet, nous choisissons de traiter la problématique de notre thèse en l'abordant par deux angles différents, mais qui se rejoignent dans nos analyses. Pour ce faire, ce travail s'articule sur quatre parties.

La **première partie** et la **deuxième partie** ont pour but d'éclaircir les termes et les hypothèses de ces deux axes en passant en revue plusieurs travaux théoriques sur les archives d'images, sur les pratiques d'énonciation, sur le support numérique, sur les stratégies de diffusion du patrimoine numérique et sur les méthodes d'analyses d'images. Tandis que dans la **troisième partie** et la **quatrième partie**, nous analysons des exemples concrets de services Web dont le but est de mettre en place des dispositifs de communication en ligne pour organiser et pour diffuser une grande quantité d'images.

Cette division en quatre parties, nous la justifions pour une raison simple : si chaque partie concerne en fait un sujet qui est construit comme un tout indépendant, la pertinence d'une partie n'apparaît que relativement aux autres. C'est pourquoi nous renvoyons à plusieurs reprises le lecteur d'un chapitre à un autre, voire d'une section d'un chapitre à une autre.

La suite de l'**introduction** présente en résumé les points importants du parcours que nous proposons au lecteur. Chaque **chapitre** est conclu par un bilan. Dans la **conclusion**, nous résumons les arguments de notre travail et, finalement, nous proposons quelques pistes de réflexion pour les recherches à venir.

Sémiotique de la communication en ligne : énonciation et stratégies multimédias

La **première partie** précisera le cadre théorique et méthodologique de notre recherche.

Dans le **chapitre 1**, nous présenterons rapidement quelques travaux sur l'accroissement des supports d'images et sur le rôle de la photographie dans la reproduction des œuvres d'art et du patrimoine. Cette multiplication des images contribue à augmenter considérablement les archives d'images. Cela nous permettra notamment d'indiquer les objectifs et les opérations qui guident la classification des images dans les archives. Nous remarquerons que, dans les pratiques d'archivage, les outils de classification d'images accordent une place centrale à la description verbale, sans doute parce que le traitement des images s'appuyait déjà avant la conversion numérique sur des méthodes standardisées de description et d'organisation des archives.

Pour éclairer cette relation entre les archives et la description verbale, nous présenterons un point de vue sémiotique sur l'archive à l'aide de plusieurs travaux en sciences du langage. Nous commencerons par remarquer que Michel Foucault a rapproché l'archive de l'énonciation à peu près au même moment que des travaux sur l'énonciation proposaient des remaniements importants dans la théorie du langage initiée par Ferdinand de Saussure et Louis Hjemslev. Après une brève présentation de cette théorie du langage et de quelques développements postérieurs de cette théorie, en particulier les propositions d'Algirdas J. Greimas, d'Émile Benveniste ou, plus récemment, de François Rastier, nous considèrerons les travaux de Jacques Fontanille, en particulier ses travaux sur rôle du support dans les pratiques énonciatives et ceux sur les stratégies qui aménagent ces pratiques. Les propositions de ce dernier auteur serviront à souligner un point important au sujet du support des textes informatiques. En effet, nous verrons que toutes les composantes et les opérations relatives à la médiation informatique ne sont pas accessibles, car cette médiation comprend à la fois un support imperceptible et un support perceptible. L'utilisateur interagit uniquement avec le second.

Ensuite, nous présenterons l'approche théorique du support de Bruno Bachimont. À l'aide de cette théorie, nous soulignerons certaines caractéristiques du support numérique à travers la notion de dispositif. De plus, cette théorie permettra de comprendre les similitudes et les différences entre le numérique et d'autres types de supports d'inscription. Cette approche s'appuie également sur la sémiotique pour caractériser ces supports, sauf qu'elle ne se confond pas exactement avec les travaux de Fontanille. Néanmoins, nous verrons que certains rapprochements sont envisageables. En effet, si l'une des spécificités du support numérique est qu'il se rapporte principalement à des opérations sur des nombres, nous remarquerons que les archives numériques impliquent toujours des usages pratiques et une dimension rhétorico-herméneutique, puisqu'il s'agit d'interpréter et discuter des connaissances ; cette question est notamment traitée par Bachimont.

Enfin, nous proposerons d'aborder la dimension énonciative et les stratégies de la communication en ligne des archives d'images en rapprochant la sémiotique des pratiques de Fontanille de la perspective proposée par l'approche sémio-pragmatique théorisée par Roger Odin. Nous verrons que ces auteurs mettent en avant la dimension pragmatique de la communication en soulignant que l'expérience d'un lecteur correspond à la production d'un texte dans un certain contexte qui implique une compétence communicationnelle et qui est nécessairement influencé par différentes contraintes spécifiques de tel ou tel contexte de communication. Ce sont les contraintes de la communication en ligne d'images que nous considèrerons dans le présent travail.

Pour ce faire, dans le **chapitre 2**, nous présenterons rapidement quelques considérations générales sur les « humanités numériques » (*digital humanities*), appellation qui désigne un domaine de recherche dont le but est de comprendre comment l'informatique influence les approches en sciences humaines pour analyser des connaissances. Nous discuterons ensuite différents travaux sur les archives d'images, en particulier les travaux qui étudient l'impact de cette communication en ligne sur les pratiques d'archivage ou, plus généralement, l'impact de la conversion numérique sur l'expérience de la culture visuelle. Ces travaux s'accordent sur un point : la reproduction numérique des archives et des collections d'images se base surtout sur la dématérialisation des ressources documentaires.

En effet, nous verrons que la dématérialisation des archives est à la base du principe d'accessibilité à distance qui remplace la manipulation directe des supports par

la manipulation de substituts. Nous constaterons quelques désavantages de cette dématérialisation, en indiquant que plusieurs travaux fournissent déjà un cadre de réflexion pour comprendre les changements qui sont en cours dans les pratiques patrimoniales.

Enfin, nous remarquerons que les stratégies de reproduction numérique des archives et des collections d'images se fondent généralement sur des systèmes de gestion des bases de données et sur la conception d'une interface-utilisateur. Sur ce sujet, nous rapprocherons les discussions sur l'énonciation (**chapitre 1**) des travaux en humanités numériques de Johanna Drucker. Cette auteure a notamment mis en évidence que les interfaces impliquent une énonciation qui n'est pas toujours prise en compte. Pour finir, nous proposerons un point de rencontre entre l'utilisateur et le système informatique : le dispositif de lecture des ressources numériques. Étant donné que la communication en ligne s'appuie sur un dispositif de médiation, nous proposerons d'étudier la communication en ligne comme une médiation ouverte en reprenant à notre compte la définition de l'« œuvre ouverte » proposée par Umberto Eco, c'est-à-dire une œuvre où la participation active de l'utilisateur est indispensable pour produire du sens. Dans le cas de la communication en ligne, la question est de savoir comment cette participation est construite et encadrée par une stratégie qui prépare un point de vue pour l'utilisateur.

Analyses d'images et méthodes assistées par un ordinateur

Le but de la **deuxième partie** est de contraster les approches en histoire de l'art et en sciences du langage qui étudient les images et les méthodes d'analyse informatique des images. Ces explications sont essentielles, car les dispositifs récents de communication en ligne d'archives et de collections d'images proposent de plus en plus des instruments d'analyse pour aider les chercheurs qui étudient des images. Il faut donc préciser les disciplines qui sont concernées.

Dans le **chapitre 3**, nous présenterons d'abord les travaux d'Erwin Panofsky sur l'iconologie et l'interprétation des œuvres d'art visuel. Pour faire ressortir les acquis et les limites de l'approche iconologique, nous mettrons en évidence deux moments importants. Premièrement, le renouvellement de l'iconologie à travers des travaux qui se concentrent, plus généralement, sur l'étude de la culture visuelle (par exemple, les travaux de Richard Howells, de Laurent Gervereau ou d'Emmanuel Alloa), ou encore à travers des travaux sur la comparaison et la résurgence des formes visuelles qui sont mobilisées

dans les productions visuelles en fonction des époques (Aby Warburg ; Maud Hagelstein). Deuxièmement, le rapprochement entre l'iconologie et d'autres disciplines, notamment l'anthropologie culturelle (Hans Belting). Ce contexte favorise le dialogue entre différentes disciplines et aboutit même sur l'idée d'une approche pluridisciplinaire pour étudier les images et le visuel (par exemple, la *Bildwissenschaft*).

Ensuite, nous présenterons rapidement des travaux en sémiologie sur l'analyse d'images. Partant du point de vue de Roland Barthes, nous verrons que ces travaux partagent certaines problématiques avec l'iconologie (avant son renouvellement), notamment l'idée que le processus d'analyse et d'interprétation des images se fonde avant tout sur une description verbale. Néanmoins, nous remarquerons que Louis Marin a mis en évidence l'importance du support et du cadre de la représentation dans l'analyse des peintures. Cela amène l'auteur à distinguer deux dimensions de la représentation : une dimension transitive (ou transparente) et une dimension réflexive (ou opaque) concernant la relation d'énonciation qui est en jeu dans la représentation.

D'un autre côté, en nous appuyant sur les propositions théoriques de Greimas, nous remarquerons que la sémiotique visuelle a proposé une approche analytique qui fait ressortir deux niveaux de lecture des images (figuratif et plastique). Nous verrons que cette approche sémiotique des images a par la suite favorisé des travaux sur la rhétorique visuelle (Groupe μ) et sur l'analyse énonciative des images (Jacques Fontanille ; Jean-François Bordron ; Maria Giulia Dondero). Ces derniers travaux ont montré que le positionnement de l'observateur (son « point de vue ») est construit par rapport à la production d'une image comme un lieu de savoir, c'est-à-dire un lieu d'où les choses sont vues et à partir duquel l'observateur construit sa propre vision.

Lorsque nous présenterons ces travaux en iconologie et en sémiotique visuelle, nous verrons en outre que des historiens de l'art et des sémioticiens envisagent actuellement d'utiliser des instruments informatiques pour analyser des images, notamment parce que ces instruments classent rapidement des images en fonction de mots-clés ou de caractéristiques plastiques.

Pour finir, nous considèrerons deux problématiques au sujet de l'analyse et de la classification d'images. D'une part, nous mettrons en évidence certains problèmes concernant l'indexation et le métalangage de description des images. D'autre part, nous présenterons plusieurs travaux en histoire, en sociologie et en psychologie de l'art qui

insistent sur l'importance de la comparaison de nombreuses figures et formes visuelles lorsqu'il s'agit d'étudier des images. Selon nous, cet intérêt pour l'étude de la dimension plastique de l'image et la comparaison des formes visuelles présentes dans un corpus d'images est justement un point de rencontre possible entre ces travaux et les méthodes informatiques d'analyse.

Dans cette perspective, le **chapitre 4** présentera les usages plus récents du traitement informatique dans l'analyse d'images et l'étude du visuel. Nous voulons montrer que d'autres approches ont émergé dans ce domaine d'analyse. Pour ce faire, nous présenterons successivement différents domaines de l'informatique dont le but est de développer des instruments pour analyser des images (*image processing*, *computer vision*, *data visualization, deep learning*). Ces instruments se basent sur la transformation en données numériques des caractéristiques plastiques des images grâce à une quantification complète de la surface des images. Il s'agit dans la plupart des cas de programmer des ordinateurs pour traiter ces nombres à l'aide d'algorithmes et en fonction de différents paramètres d'analyse.

Néanmoins, comme nous l'avons suggéré, ces nouvelles approches n'analysent pas les images dans le même cadre de réflexion que l'histoire de l'art ou les sciences du langage. Du point de vue informatique, les images numériques sont composées de pixels correspondant à des valeurs numériques ; ce sont donc uniquement des nombres qui sont analysés par un ordinateur. Malgré ce décalage, nous soulignerons un point en commun entre les méthodes d'analyse issues de l'iconologie et des sciences du langage et ces « méthodes assistées par un ordinateur » (*computer-assisted methods*) : les unes comme les autres concentrent de plus en plus leurs approches sur la reconnaissance, sur la description et sur la compréhension des structures et des formes visuelles qui constituent les images.

Analyse du corpus : tendances stratégiques et types d'utilisateurs

Les **deux premières parties** constitueront l'état de l'art sur lequel se basent les analyses de notre corpus. Dans la **troisième partie** (*Bases d'images numériques et expositions virtuelles*), nous étudierons l'évolution des services informatiques qui classent uniquement des images en fonction de mots-clés vers des expositions virtuelles d'images qui impliquent de plus en plus la participation des internautes à travers des applications interactives. Ensuite, dans la **quatrième partie** (*Chiffrer et déchiffrer la culture visuelle*

à travers les big visual data ?), nous montrerons comment cette situation a encore une fois évolué à la suite des progrès en informatique qui seront présentés dans la **deuxième partie**.

Dans la plupart des cas, les exemples du corpus proviennent de projets institutionnels qui proposent des services Web dont le but est de diffuser des archives ou des collections d'images. Nous distinguerons cependant *trois tendances stratégiques* en ce qui concerne la communication en ligne d'images. Ces tendances visent à construire différents *points de vue* pour l'utilisateur du service (utilisateur-explorateur ; utilisateur-lecteur ; utilisateur-observateur ; utilisateur-expérimentateur).

Une *première tendance* fera ressortir une stratégie visant à mettre en place des bases d'images numériques qui sont rattachées à des institutions patrimoniales. Le **chapitre 5** se focalisera sur des projets mis en place par des institutions patrimoniales qui gèrent des archives d'images en relation avec le patrimoine culturel d'un pays. Ces archives sont principalement des archives photographiques qui ont été converties en ressources numériques. Le *premier corpus* regroupe quatre projets (*Belgian Art Links and Tools* ; *Plateforme ouverte du patrimoine* ; *Memobase* ; *Getty Search Gateway*). Dans tous les cas, il s'agit de représenter les images par le moyen d'un service Web à peu près similaire. À l'aide de ces exemples, nous verrons que les possibilités d'exploration sont encadrées par la conception d'une interface Web, et surtout par la mise en avant des activités de l'institution patrimoniale à travers une énonciation verbale. Les services Web des institutions fonctionnent surtout comme des inventaires débouchant sur des listes de résultats. Dans ce cas, nous verrons que la stratégie favorise uniquement un double point de vue : celui de l'*utilisateur-explorateur* et celui de l'*utilisateur-lecteur*. Selon nous, cette première tendance promeut un dispositif de communication en ligne qui, selon nous, est limité par l'importance de la description verbale dans les premières bases d'images numériques.

Ensuite, nous observerons un renouvellement des services Web qui diffusent des archives d'images. Dans les **chapitres 6** et **7**, nous présenterons une *deuxième tendance* qui renouvelle : (1) l'origine et le type d'images qui sont agglomérées dans des bases d'images numériques (stratégie de coopération) ; (2) le rôle et la participation des utilisateurs dans la constitution des ressources de la base d'image (stratégie de participation) ; (3) les fonctionnalités des interfaces en favorisant une utilisation interactive (stratégie de l'interactivité).

Dans le **chapitre 6**, nous considèrerons d'abord la stratégie liée à *Wikimedia Foundation* qui propose, d'une part, la plus importante médiathèque en libre partage sur Internet (*WikimediaCommons*), et d'autre part, une encyclopédie des arts visuels (*WikiArt*). Dans le premier cas, le but de la stratégie est la *participation de l'utilisateur*. Comme les bases d'images, cette médiathèque en ligne nécessite des données structurées qui sont généralement des métadonnées descriptives réduisant les images à des descriptions verbales. Dans ce cas, le traitement informatique ne consiste donc pas à classer ou à analyser des images, mais plutôt à classer des descriptions simplifiées de ces images ou d'autres informations sous une forme verbale. Ensuite, nous considèrerons un *projet de coopération* qui vise à rassembler les ressources documentaires de plusieurs archives d'images de différents pays en Europe dans le but de promouvoir un sentiment d'identité européenne à travers le patrimoine culturel (*Europeana*). Par rapport aux services Web qui sont regroupés dans la première tendance, nous remarquerons que ce projet européen intègre des discours de vulgarisation à propos des ressources qui sont agglomérées par ce projet. Dans ce cas, la stratégie se fonde encore sur la construction d'un *utilisateur-explorateur* et d'un *utilisateur-lecteur*, mais en favorisant beaucoup plus des modalités de communication en rapport avec le second type d'utilisateur.

Dans le **chapitre 7**, nous remarquerons qu'un autre renouvellement des services Web se généralise à travers d'autres projets institutionnels qui proposent des expositions virtuelles. Néanmoins, nous remarquerons que les possibilités de manipulation permises par ces services sont limitées parce que certaines caractéristiques de ces expositions sont similaires à celles des bases d'images numériques des institutions, notamment la manière de classer les images. De plus, les expositions virtuelles utilisent surtout la communication en ligne pour transmettre des discours de vulgarisation concernant les ressources numériques, de telle sorte que ces expositions se fondent sur des interfaces qui fonctionnent principalement comme un espace qui est écrit et scénarisé (*utilisateur-lecteur*). Nous présenterons cependant deux services Web qui représentent des collections d'images par le moyen d'applications interactives intégrées dans le site Web (*Rikjsstudio* ; *12 Sunset : Exploring Ed Ruscha's Archive*), puis deux autres projets qui font ressortir le potentiel interactif de l'informatique dans le cadre de recherches en histoire de l'art (*VISUALL-tek* ; *Rekall*). Dans ces cas, la stratégie vise à construire un point de vue interactif pour un *utilisateur-observateur*, autrement dit, il s'agit de favoriser

un espace visible plutôt qu'un espace lisible, cela même si le service comprend encore de nombreuses informations sous une forme verbale.

D'une manière générale, les services Web qui seront analysés dans la **troisième partie** indiqueront que l'informatique a d'abord été utilisée pour classer une grande quantité d'images uniquement en fonction d'un langage de contrôle, et non pas à partir des propriétés visuelles des images. Dans la **quatrième partie**, nous présenterons une *troisième tendance* stratégique qui se fonde sur les travaux plus récents en informatique (cf. **chapitre 4**). Cette troisième tendance cherche à éprouver les performances analytiques des ordinateurs dans des tâches de classification automatique et de visualisation d'images, puis à valoriser ces performances à l'aide d'un dispositif de communication en ligne. Le *troisième corpus* rassemble des projets dont le but est d'appliquer des opérations d'analyse sur des corpus d'images numériques à l'aide d'une programmation informatique.

Dans le **chapitre 8**, nous analyserons d'abord deux projets qui s'appuient sur des techniques de visualisation : le premier projet concerne des archives sur des performances artistiques (*The Experiment in Art and Technology Datascape*), tandis que le deuxième projet concerne une collection de photographies artistiques (*Object : Photo*). Ensuite, nous présenterons deux projets dont le but est chaque fois de montrer les performances du traitement informatique dans la classification automatique et la visualisation d'images en fonction de leurs caractéristiques visuelles (*Phototrails* ; *Vy*). Dans la continuité des deux tendances précédentes, l'un des principaux enjeux est de concevoir des interfaces Web qui offrent de nouvelles possibilités pour organiser, pour reproduire et pour manipuler des médias visuels. Sauf que, les *stratégies de visualisation* visent à construire un point de vue interactif pour un *utilisateur-expérimentateur* en apportant une dimension analytique et réflexive dans le dispositif de communication en ligne.

Le **chapitre 9** présentera de manière détaillée la conception d'un moteur de recherche en fonction des caractéristiques plastiques des images (*Replica*). Pour terminer, nous verrons que de plus en plus de services se basent sur la *media visualization* parce que celle-ci met en évidence des relations inédites au sein d'un grand nombre d'images. Les techniques informatiques changent les conditions de visibilité de ces images puisqu'elles offrent un espace virtuel d'expérimentation et de réflexion pour un *utilisateur-expérimentateur*.

Somme toute, notre objectif est de montrer comment (1) la conversion numérique, (2) la communication en ligne et (3) les méthodes d'analyse informatique concourent à mettre en avant de nouvelles conditions de visibilité et de manipulation des larges ensembles de médias visuels. Dans le présent travail, nous insisterons constamment sur un point : le domaine qui concerne l'étude de ces médias est un champ à la fois vaste et paradoxal. C'est dans ce domaine que les images sont *déchiffrées*, c'est-à-dire que nous décelons ce qui, dans ces images, est obscur, difficile à comprendre. Autrement dit, c'est dans ce champ que les significations des images sont élucidées, et que le sens de ces images est traduit en des termes plus clairs. Déchiffrer les images, c'est donc chercher à les interpréter à l'aide des connaissances et des outils disponibles.

D'un autre côté, c'est aussi dans ce domaine que les images sont *chiffrées* à force de théorisations, c'est-à-dire qu'elles s'obscurcissent, comme si elles étaient dissimulées par des systèmes qui restent secrets pour celui qui n'en possède pas la clef. En effet, lorsque nous imposons aux images diverses théories et méthodes d'analyse, ou lorsque nous les convertissons en nombres que seules des machines sont capables de traiter, les significations et le sens des images s'opacifient, parce que les images sont constamment transcrites à travers d'autres systèmes de notations (d'autres codes et codages) ou encore par le moyen de divers langages qui se substituent aux images que nous regardons.

Cette opposition (« chiffrer » *vs* « déchiffrer ») indique selon nous le mouvement interminable qui va de l'observation et de l'interprétation des images à la théorisation et à la méthodologie, et inversement. Parce qu'il ne peut jamais être terminé, ce mouvement résiste à deux tentations extrêmes qui ont toujours menacé le champ de l'étude scientifique des images : d'un côté, la tentation de la transparence de l'image qui s'appuie sur sa ressemblance avec le monde visible, et donc une image qu'il faut uniquement déchiffrer d'après ce qu'elle montre, parce que tout est là, dans cette transparence de l'image même, comme s'il s'agissait de regarder directement et de nommer ce qui est manifesté à travers elle ; de l'autre, la tentation de trop chiffrer les images au risque de les crypter, ou pire, de clôturer les débats sur la compréhension des images à l'aide d'une seule approche théorique, d'une seule méthode ou d'un seul instrument, qui résoudraient tous les problèmes et les difficultés du chercheur.

PREMIÈRE PARTIE

SÉMIOTIQUE DE LA COMMUNICATION EN LIGNE : ÉNONCIATION ET STRATÉGIES MULTIMÉDIAS

CHAPITRE 1
Pour une approche sémio-rhétorique des archives et des collections d'images numériques

Dans cette **première partie,** nous définirons le cadre théorique et méthodologique pour étudier notre problématique (à savoir : montrer comment le numérique favorise une nouvelle forme de médiation des archives et des collections d'images). Le **premier chapitre** présentera d'emblée l'approche sémio-rhétorique qui nous a permis de conduire notre recherche[1]. Nous commencerons par expliquer ce que nous entendons par archives d'images et collections d'images. Plusieurs travaux remarquent que les pratiques sociales et culturelles ont dû s'accommoder d'une production intense de différents supports d'images (**sect. 1.1.**). Ces supports sont généralement organisés dans le cadre d'un traitement documentaire qui permet de décrire, de classer et de diffuser des archives d'images (**sect. 1.2.**). Ensuite, nous considèrerons différents travaux théoriques sur l'archive en relation avec les sciences du langage (**sect. 1.3.**). Cette relation imposera d'éclairer plusieurs notions importantes dans la théorie structuraliste du langage, notamment : les signes, l'énonciation, l'opposition entre discours et textes, la question des objets-supports et des stratégies. Enfin, nous proposerons d'éclairer la relation entre l'archive et la sémiotique à la lumière de travaux sur le support numérique (**sect. 1.4.**) et sur la rhétorique et la pragmatique de la communication (**sect. 1.5.**).

[1] Ce cadre de réflexion sera complété dans le **chap. 2**.

1.1. Accroissement des archives et des collections d'images

Les spécialistes des images et du visuel doivent faire face à une augmentation sans précédent de la quantité et de la variété de supports d'images à prendre en compte. Plusieurs travaux théoriques ont en effet montré que les *mass media* (en français, les « médias de masse », c'est-à-dire l'ensemble des techniques et des supports de diffusion massive de l'information : presse, photographie, radio, cinéma, télévision, etc.) modifient radicalement la perception et les pratiques de communication humaine, ce qui aboutit sur des changements dans l'organisation de nombreuses activités culturelles et scientifiques (Mc Luhan, 1968 [1964])[2]. La prolifération des médias visuels et audiovisuels dans les usages sociaux favorise l'établissement d'une « civilisation des images » (Fulchignoni, 1968), voire d'une « civilisation des médias » où les formats techniques se renouvellent sans cesse (Flusser, 1991). Plus récemment, cette diversité de *mass media* est rassemblée sous l'expression *old media* (« anciens médias »), par opposition aux *new media* (« nouveaux médias ») qui se fondent sur les ordinateurs et le réseau Internet (Manovich, 2015 [2001])[3].

D'autres travaux ont cherché à mettre en évidence les moments-clés de l'évolution des médias visuels et audiovisuels. Certains passages du livre *L'image : communication fonctionnelle* (Moles, 1981) évoquent ces moments-clés. Dans cet ouvrage, même si Abraham Moles distingue d'emblée l'image visuelle fixe de l'image visuelle animée, ou encore l'image figurative de l'image abstraite (cf. Moles, 1981 : 8), l'image est en général comprise comme un processus de « cristallisation du réel ». Ce processus serait passé par trois étapes ou trois « âges » de la production d'images multiples (*ibid.* : 17-19) : (1) le premier correspond à l'âge de la copie multiple à travers la gravure ; (2) le deuxième âge correspond à l'invention de l'image photographique, aboutissant sur une « massification » et une « abondance » des images ; (3) enfin, le troisième âge est celui « d'une prise de conscience », parce que l'image est « un fait de communication

[2] En suivant une autre voie, Theodor W. Adorno et Max Horkheimer avaient déjà proposé de mettre en évidence les dérives esthétiques qu'imposent ces médias de masse, notamment en abordant le problème des « industries culturelles » (Adorno & Horkheimer, 2019 [1947]).

[3] Plus récemment, nous verrons que les travaux de Lev Manovich soulignent que l'informatique offre clairement de nouvelles possibilités, à la fois pour manipuler et pour analyser des images numériques ; cf. **chap. 4, sect. 4.1.** Auparavant, des travaux proposaient déjà d'étudier les techniques de production d'images en mettant en évidence les différences et les points communs entre les images picturales et optiques et les images numériques ; cf. les travaux fondateurs d'Edmond Couchot (1988).

[mais] un fait accepté dans la banalité quotidienne », de telle sorte que les théories de la communication visuelle doivent fournir des concepts qui peuvent être éprouvés en termes d'opérations concrètes et répétables[4].

Nous retrouvons également cette division de l'histoire des images en différentes étapes chez d'autres théoriciens. Dans *Vie et mort de l'image*, Régis Debray (1992 : 283-328) propose trois âges du regard en fonction de trois « charnières » : (1) la logosphère (de l'invention de l'écriture à l'imprimerie), (2) la graphosphère[5] (de l'imprimerie à la télévision) et (3) la vidéosphère (de la télévision aux années 1990). Chaque âge ou *médiasphère*[6] « dessine un milieu de vie et de pensée, aux étroites connexions internes, un écosystème de la vision et donc un certain horizon d'attente du regard » (*ibid.* : 286-287) ; à ces trois « médiasphères », Louise Merzeau (1998) en ajoute par la suite une quatrième, correspondant à l'époque des réseaux numériques : (4) l'hypersphère. De même, dans l'introduction de son *Histoire du visuel au XX^e siècle*, Laurent Gervereau (2003 : 29-40) distingue « trois âges de l'image industrielle » : (1) l'ère du papier (de 1848 à 1916), (2) l'ère de la projection (de 1916 à 1960), (3) l'ère de l'écran (de 1960 à 2000). À ces trois âges, Gervereau ajoute un quatrième temps : (4) l'*ère du cumul*, c'est-à-dire le temps de la constitution de bases de données qui favorise l'accumulation et la diffusion d'images sur Internet. Le même auteur remarque ailleurs que « la grande innovation, à l'ère de la multiplication industrielle des images, c'est le média global, celui qui absorbe tous les autres (sans les tuer pour l'instant) : Internet » (Gervereau, 1994 : 156-157).

Même si ces divisions en périodes de l'évolution des médias visuels ne sont pas exactement semblables, elles mettent en évidence que, depuis plusieurs siècles, des changements importants dans les techniques de production des images et dans les usages pratiques de ces techniques favorisent un *accroissement des images*. Le passage, d'un âge à un autre, ne supprime pas l'usage des anciennes techniques de production. Au contraire,

---

[4] Ces opérations découlent, selon l'auteur, d'une « stratégie de la communication » ; à ce sujet, cf. **infra, sect. 1.3.4. & chap. 2**.

[5] Sur le passage des techniques de l'imprimerie aux techniques de reproduction visuelle, comme la gravure ou la photographie, cf. la contribution de William M. Ivins Jr. (1969 [1953]).

[6] La « médiasphère » est un concept de la *médiologie*. Ce concept concerne un système de transmission culturel qui est associé à un milieu technique. Pour Debray, le « système dominant de conservation des traces (saisie, stockage et circulation) sert de noyau organisateur à la *médiasphère* d'une époque donnée dans une société donnée » (Debray, 1991 : 229).

les anciens médias s'intègrent successivement dans les médiasphères suivantes, ce qui aboutit finalement sur une grande diversité de médias. Pourtant, il est aujourd'hui évident que l'informatique (l'hypersphère, l'ère du cumul) est la principale médiasphère, celle qui s'approprie et englobe toutes les autres.

Bernd Stiegler (2008 : §2) indique lui aussi qu'il est question d'un « déluge d'images » dans les travaux actuels sur le visuel, et qu'il s'agit aussi de prendre en compte la « réflexivité » et le « pouvoir des images ». Le renouvellement des théories sur les images et le visuel s'appuie sur le fait que « les images jouent déjà un rôle central dans de nombreux domaines scientifiques, alors qu'elles n'entrent que peu à peu dans le champ de vision de l'attention théorique » (*ibid.* : §2). Ce renouvellement théorique s'est développé grâce à la collaboration entre les scientifiques et diverses institutions culturelles et patrimoniales. En effet, ce déluge n'est pas apparu de nulle part ; il est surtout constaté dans des archives ou collections d'images dont la portée s'étend « bien au-delà des frontières des domaines classiques de la recherche sur l'image que sont la science de l'art et l'archéologie » (*ibid.* : §7). Les archives et collections d'images sont donc souvent étroitement liées à des institutions patrimoniales (fonds d'archives, musée, institut national ou régional) qui les conservent, les gèrent, les classent[7].

La multiplication des images dans les archives est également liée à la photographie. Dans *Le musée imaginaire*, André Malraux affirme que « l'histoire de l'art depuis cent ans, dès qu'elle échappe aux spécialistes, est l'histoire *de ce qui est photographiable* » (Malraux, 1965 : 123 ; sur l'institution des musées de photographies en France, cf. Challine, 2017). Cette influence de la photographie sur l'histoire de l'art a largement été abordée par les spécialistes du patrimoine culturel. Jean-Pierre Babelon et André Chastel (1994) soulignent l'ampleur des pratiques de reproduction en images du patrimoine dans les institutions en France, mais aussi ailleurs en Europe. Les institutions patrimoniales sont passées « à un plan en quelque sorte technique, scientifique » du patrimoine (*ibid.* : 100), plan qui est étoffé par l'utilisation de plus en plus fréquente du dispositif photographique en tant que moyen de reproduction en images des monuments et des œuvres d'art. L'entreprise bien connue de l'*Inventaire général*[8] illustre bien « la

[7] Cf. **infra, sect. 1.2.**

[8] Service national créé en 1964, à l'initiative de Malraux et Chastel. En utilisant la reproduction photographique, cette initiative a permis de mettre en place des archives d'images qui recensent et illustrent un grand nombre des monuments et d'œuvres d'art en France.

nouvelle stratégie qui convient à une société au patrimoine complexe et encore imparfaitement exploré » (Babelon & Chastel : 100). L'association de la photographie avec une visée stratégique est au cœur de cette entreprise dont le but est de valoriser par le moyen d'images photographiques le patrimoine auquel on attribue un intérêt culturel, artistique, historique ou scientifique.

Même si les procédés photomécaniques et les supports photographiques sont parfois différents selon l'époque considérée[9], la photographie est une technique charnière dans l'évolution des pratiques documentaires. En effet, une partie importante des archives d'images produites durant le XXe siècle comprennent ce type de support, car l'une des « grandes fonctions » de la photographie a été de dresser un « inventaire du réel sous la forme d'albums puis d'archives »[10] (Rouillé, 2005 : 120). Dans le même sens, l'influence de la photographie sur les pratiques muséales[11], sur les pratiques scientifiques et sur le journalisme a assurément été décisive, puisqu'elle a largement été utilisée pour conserver des traces matérielles et visuelles du patrimoine culturel, pour documenter les pratiques de terrain des chercheurs ou encore pour illustrer un discours journalistique[12].

Dans le cadre des pratiques d'archivage et des musées, ces photographies documentaires sont considérées comme le résultat d'un ensemble de « forces créatrices » qui sélectionnent des images, les préservent, les classent, les communiquent pour créer collectivement de « nouvelles ouvertures au monde » (cf. Sohier, Lugon & Lacoste, 2017 : 9-11). Pourtant, la relation esthétique entre les spectateurs (le « public ») et le patrimoine culturel change radicalement. Ce spectateur ne fait plus l'expérience directe des œuvres d'art ou des monuments, il les admire à distance grâce à un dispositif technique[13]. Bien entendu, cette reproduction technique n'épuise pas l'œuvre ou le

[9] À propos des procédés et du support de la photographie, cf. par exemple Cartier-Bresson (2011).

[10] Pour plus de précisions sur le rapport entre photographie et document, cf. Rouillé (2005 : 120-171) ou encore Emerling (2012 : 83-113).

[11] Voir Edwards & Morton (2015) ou encore les contributions dans Sohier, Lugon et Lacoste (2017).

[12] Jean Prinet (1948) a proposé l'une des premières synthèses sur les applications de la photographie dans le domaine scientifique. L'utilité documentaire de la reproduction photographique est constatée dans divers champs de recherches sur le patrimoine culturel (Aa. Vv., 2007) ou dans les pratiques de diffusion des œuvres d'arts visuels (Chirollet, 2013). Nous pouvons en outre mentionner des travaux au sujet de l'utilisation de la photographie en archéologie (Chéné, Foliot & Réveillac, 1999), en histoire (Demeuleneare-Douyère, Plouvier & Souchon, 2010), en architecture (Fanelli & Mazza, 2016) ou encore dans la presse (Fabre, 2020).

[13] Sur la notion de dispositif technique, cf. **infra, sect. 1.4.** Voir notamment la **note 66**.

monument qu'elle représente forcément de manière incomplète. Elle se substitue plutôt à l'objet original et expose celui-ci d'une autre manière[14]. Cette exposition ne vise pas à oublier les monuments et les œuvres, elle a seulement pour fonction d'évoquer l'original à travers des images. Nous verrons que ce principe de substitution est également au centre de la conversion numérique des archives et des collections d'images[15].

1.2. Traitement documentaire des images : conserver, décrire, classer

D'une manière générale, les *archives* sont définies comme un « ensemble de documents hors d'usage courant, rassemblés, répertoriés et conservés pour servir à l'histoire d'une collectivité ou d'un individu »[16]. L'*archivage* se rapporte donc à un but et à une action. Le but est la sauvegarde des informations qui sont inscrites sur des supports matériels. L'activité consiste à recueillir, déposer et classer de manière cohérente ces supports dans une collection ou un fonds d'archives. Ainsi, les archives sont étroitement liées à des supports d'inscription qu'il faut conserver parce que ces supports contiennent des connaissances qui sont considérées comme une attache indéfectible avec le passé et comme des preuves de ce lien.

Dans cette perspective, l'un des objectifs de l'archivage d'images est de mettre en place un traitement documentaire dont le but est d'organiser un ensemble cohérent d'images et de ressources documentaires à destination d'un public ciblé. Ces ressources servent à conserver des connaissances par le moyen d'images, comme d'autres ressources consignent des connaissances à travers des documents écrits ou d'autres types d'objets matériels. Les principaux usagers des archives d'images sont souvent des historiens,

[14] Rappelons-nous de la mise en garde de Walter Benjamin à propos de l'impact de la reproduction technique sur la perception des œuvres d'art : cette reproductibilité technique a tellement pour but d'*exposer* l'œuvre que « la prépondérance absolue » de cette valeur d'exposition de l'œuvre « assigne des fonctions tout à fait neuves » à l'œuvre, « parmi lesquelles il se pourrait bien que celle dont nous avons conscience – la fonction artistique – apparaisse par la suite comme accessoire » (Benjamin, 2012 : 172). Pour un point de vue plus récent, voir par exemple Genette (2010 : 348-353).

[15] **Chap. 2, sect. 2.2.**

[16] https://www.cnrtl.fr/definition/archives. L'histoire des archives a traversé toutes les époques, depuis « l'invention de l'écriture » jusqu'à nos jours. Cette histoire est pour la plupart indissociable de la progressive « affirmation des autorités publiques », et donc de « l'histoire des archives publiques » (Galland, 2016 : 7-45). Sur l'archivage, cf. entre autres : Schellenberg (1956) ; Favier (1958) ; Couture & Rousseau (1982) ; Gagnon-Arguin (1998) ; Couture (1999) ; Delsalle (1998) ; Chabin (2000) ; Cœuré & Duclert (2001).

surtout des spécialistes de l'histoire de l'art ou du patrimoine culturel qui trouvent dans ces ressources des documents utiles pour corroborer leurs travaux.

Dans l'ouvrage sur l'image cité précédemment, Moles suggère en effet que l'iconothèque sert avant tout au « stockage de trésors d'images » (Moles, 1981 : 58). Cette expression (« iconothèque »), nous pouvons la rapporter à deux autres expressions plus courantes. D'un côté, une *collection d'images* est un regroupement d'images qui correspond, par exemple, au travail d'un artiste (ou d'un groupe d'artistes), à des œuvres d'art visuel d'une certaine période ou encore à un certain thème. De l'autre, une *archive d'images* est aussi le résultat d'un regroupement d'images, sauf que, comme nous l'avons suggéré, les images concernent plutôt divers médias visuels auxquels des institutions patrimoniales accordent une valeur documentaire. Cette activité de stockage peut autant être réalisée par un collectionneur amateur, qui constitue une collection privée, que par les conservateurs de musée, les conservateurs de bibliothèque ou des archivistes.

Moles affirme qu'il est plus facile de comprendre des images « quand elles sont *en groupe* que quand elles sont isolées » (*ibid.* : 60), mais il précise en outre qu'une analyse de contenu[17] des images est nécessaire lorsque les images sont nombreuses (*ibid.* : 58). Pour Moles (*ibid.*), analyser le « contenu d'une image », c'est décrire le « contenu sémantique » de cette image ; cette description d'images se base soit sur la « valeur dénotative de l'image » (ce que l'image manifeste explicitement ; p. ex. : une photographie d'une forêt sera décrite par les mots-clés « forêts » ou « arbres »), soit sur les « significations évoquées ou suggérées par une image » (p. ex. : la même photographie d'une forêt peut éventuellement être décrite par le mot « paysage »)[18].

Il en résulte que l'organisation d'un large ensemble d'images[19] correspond à une classification systématique de descriptions verbales qui sont projetées sur chaque image.

---

[17] Sur ce sujet, voir aussi Bardin (1997 [1977] : 30-51 & 125-206), qui explique notamment plusieurs techniques d'analyse du contenu (cf. 207 *sq.*) qui se basent sur le modèle linguistique (analyse catégorielle, analyse de l'évaluation, analyse de l'énonciation, analyse propositionnelle du discours). En outre, Bardin remarque que la systématisation des opérations du traitement documentaire s'accélère à la suite de l'informatisation croissante des ressources documentaires et des progrès en cybernétique pour traiter automatiquement l'information.

[18] Cette conception fait notamment écho au point de vue de Roland Barthes (1961) ; cf. **chap. 3, sect. 3.2.**

[19] Dans le présent travail, nous employons souvent l'expression « larges ensembles d'images », car celle-ci est moins exclusive que les expressions « archives d'images » et « collections d'images » ; un large ensemble d'images peut comprendre des œuvres d'art, des images documentaires, voire les deux.

Cependant, il ne faut pas croire que ce type de classification des images n'est pas problématique pour Moles, car « tout mode de classification analytique d'une iconothèque est en soi une *théorie de l'image* » (*ibid.* : 221-222). Aucune méthode de classification n'est optimale, puisque chaque méthode est contrainte de laisser de côté « un certain nombre d'éléments exclus ou rejetés, inadéquats à la structure de pensée particulière qui est projetée sur eux » (*ibid.* : 224-225). S'il insiste sur la spécificité des images par rapport aux documents écrits, l'auteur retient malgré tout comme pertinente la classification des images par mots-clés (ou, selon lui, la classification par « projection de thèmes ») ; ce qui n'empêche pas que d'autres types de classification soient pertinents, notamment la classification à partir de la ressemblance visuelle entre des images. Moles envisage cette possibilité, notamment lorsqu'il indique que certaines approches consistent à analyser les images selon la méthode des « matrices de similarité » (cette méthode de classification des images repose sur des critères de ressemblance visuelle parmi un ensemble d'images, même si nous ne savons pas expliquer pourquoi ces images se ressemblent). L'auteur reste cependant très évasif sur cette méthode qui vise à regrouper des images à partir de ce qu'il nomme leur « voisinage » (*ibid.* : 224-225). Comme nous le remarquerons, ce type de méthodes de classification est plus récemment développé par des travaux en informatique qui proposent des techniques pour analyser et classer automatiquement de grands ensembles d'images numériques[20].

Cela étant dit, la classification d'images par mots-clés est encore largement utilisée par les institutions patrimoniales. Dans un ouvrage sur la gestion des archives d'images de Cécile Kattnig, nous pouvons lire que la classification d'une importante quantité d'images dans des archives se base sur des descriptions verbales résumant les principales informations relatives à ces images (Kattnig, 2005 : 23 & 73 *sq*.). Ce manuel de gestion des fonds d'archives d'image présente toute une série d'opérations standardisées pour réaliser le traitement documentaire de ces images[21]. Parmi ces opérations de traitement des archives d'images, les opérations du « traitement matériel »

[20] **Chap. 4**. Ces méthodes informatiques sont actuellement testées sur des images conservées par des institutions patrimoniales ; cf. **chap. 8 & 9**.

[21] Au sujet du traitement des documents photographiques, cf. Charbonneau & Robert (2001 : 199 *sq*.). À propos du traitement des documents et des archives audiovisuels, cf. les travaux de Stockinger (2011 & 2012). Pour les documents iconographiques, cf. par exemple Lavalleye (1947) ou les contributions plus récentes dans Bideault (2010).

(numérotation, légendage et étiquetage des supports) sont distinguées de celles du « traitement intellectuel » (catalogage, analyse, classification des ressources).

La première phase se focalise sur la collecte, le tri et la conservation de supports d'images, le but étant de les différencier et de les conserver. La seconde phase se focalise sur le classement de l'ensemble des informations relatives à ces supports, afin de constituer un ensemble de ressources documentaires exploitables. Ces opérations visent à dresser un inventaire cohérent et le plus complet que possible des images disponibles dans un fonds d'archives.

Pour l'instant, ce que nous voulons mettre en évidence, c'est que les archives d'images sont « filtrées » et « tamisées » par des opérations de traitement documentaire, qu'il s'agisse de la collecte et de la sélection des supports d'images ou de l'analyse du contenu des images. D'une part, ce traitement comprend différents processus standardisés pour conserver des supports d'archives (traitement matériel). D'autre part, il vise aussi à appliquer systématiquement une description verbale des images pour attribuer une valeur stable à des images et pour les classer en fonction de cette description (traitement intellectuel). Si toutes les opérations du traitement documentaire sont importantes, notons que l'*indexation*[22] est une opération fondamentale pour effectuer la catégorisation systématique des contenus d'une archive par rapport à un ensemble de termes linguistiques. L'indexation rend possibles l'exploitation et la recherche de documents dans un ensemble cohérent d'images, parce que toutes les ressources documentaires sont reliées les unes aux autres à l'aide de notions et de concepts en fonction du (des) domaine(s) qui concerne(nt) ces documents (Kattnig, 2005 : 78-90).

Par conséquent, le but du traitement documentaire et de l'indexation n'est pas de proposer une analyse approfondie des archives d'images. L'indexation n'est pas un instrument d'analyse d'images, mais un outil documentaire pour organiser et classer ces images en fonction de mots-clés. Si cette description peut éventuellement être complétée par des analyses portant sur les aspects techniques ou morphologiques des images, l'objectif est surtout de caractériser les contenus informatifs véhiculés par les images en utilisant des questions simples (qui ? quoi ? où ? quand ? comment ?), voire une grille d'analyse (cf. *ibid.* : 76). Pour réaliser des analyses approfondies d'images, Kattnig (*ibid.* : 74) renvoie aux travaux d'Erwin Panofsky (1969 [1955]) et d'Ernst Gombrich

[22] À ce propos, voir aussi **chap. 3, sect. 3.3.**

(1996 [1960]), pour l'histoire de l'art, et à ceux de Charles Sanders Pierce (1958) pour la sémiotique et de Roland Barthes (1964) pour la sémiologie.

Nous présenterons plus précisément ces recherches en histoire de l'art et en sciences du langage lorsque nous considèrerons les méthodes d'analyse d'images[23]. Cependant, nous pouvons déjà remarquer que les auteurs cités par Kattnig appartiennent sans aucun doute à une première génération d'approches issues de l'histoire de l'art ou des sciences du langage. La remise en question de cette première génération d'approches par une seconde génération est une étape importante pour les travaux en sciences humaines et sociales[24].

1.3. Point de vue sémiotique sur l'archive

L'un des deux objectifs du présent travail est de s'interroger sur les dispositifs de communication en ligne des archives et des collections d'images selon un point de vue sémiotique[25]. Sans plus attendre, nous expliquerons ce choix. Ces explications nous imposent cependant une digression théorique. En premier lieu, nous verrons que deux travaux éclairent le rapprochement entre l'archive et les pratiques discursives. Dans la foulée, nous rassemblerons plusieurs remarques importantes des travaux en sémiotique sur l'analyse du sens. Ces travaux participent à la construction d'une théorie du langage qui s'appuie principalement sur la sémiotique « poststructuraliste », et qu'il ne faut pas confondre avec la sémiotique « peircienne » qui est fondée sur la pragmatique et la logique[26].

---

[23] Cf. **deuxième partie**.

[24] À titre d'exemple, partant des acquis et des limites de certaines approches en histoire de l'art (iconologie), les recherches relevant de l'*iconic turn* (Mitchell, 1994) ou de la *Bildwissenschaft* (Boehm, 2007) proposent de renouveler les théories des images afin de comprendre les mutations médiatiques et technologiques en cours. D'un autre côté, les premières approches en sciences du langage sur les images sont continuées par plusieurs travaux importants en sémiotique visuelle (Greimas, 1984 ; Groupe μ, 1992 ; Hénault & Beyeart, 2004), ou encore des travaux sur l'énonciation visuelle (Fontanille, 1989 ; Bordron, 2013 ; Dondero, 2016 & 2020). Nous examinerons en détail ces deux perspectives dans le **chap. 3**.

[25] Cf. **introduction**.

[26] Voir Charles S. Peirce (1958) ou encore les discussions dans Eco (1972 [1968] : 21-36), dans Benveniste (1974 : 43-66), ou encore dans Deladalle (1979). Nous ne discuterons pas ici cette deuxième voie, parce que cela nous éloignerait considérablement de notre sujet. Néanmoins, nous verrons que des travaux en sémiotique prennent en compte la dimension pragmatique de l'énonciation, cf. **infra, sect. 1.5.**

1.3.1. L'archive en tant que pratique énonciative

Dans la deuxième partie de *L'archéologie du savoir*, Michel Foucault (1969 : 31-106) considère en profondeur les apports de la linguistique dans les conceptions scientifiques de la connaissance[27]. Ensuite, dans les troisième et quatrième parties, l'auteur défend une approche archéologique qui consiste à mettre en évidence les couches sédimentaires des connaissances et surtout à discuter la méthode pour interpréter ces connaissances selon plusieurs niveaux discursifs. Nous ne voulons pas détailler ici toute l'approche du philosophe ni la méthode proposée dans cet ouvrage. Ce que nous voulons remarquer, c'est comment Foucault définit l'archive. À côté de la notion de « formation discursive », la notion d'archive est définie comme « le système général de la formation et de la transformation des énoncés », mais aussi comme « la loi de ce qui peut être dit », c'est-à-dire les règles qui permettent à ces énoncés d'apparaître comme des « événements singuliers » (Foucault, 1969 : 177). Ces règles ne concernent pas uniquement des normes linguistiques, car elles s'étendent évidemment au domaine plus vaste des pratiques discursives.

Pour Foucault, la langue est comprise comme le « système de construction pour des énoncés possibles », même si c'est uniquement à travers l'observation et l'étude d'énoncés réels que ce système peut être mis en évidence (*ibid.* : 117-118) ; tandis qu'un énoncé est fondé sur des relations entre des variables, et surtout il dépend d'une « fonction énonciative » qui permet « de dire, à propos d'une série de signes » si diverses unités sont manifestées ou non dans cette série (cf. *ibid.* : 120). Le mode d'existence des énoncés ne doit pas être confondu avec le mode d'existence des signes en tant qu'éléments d'une langue. L'énoncé est ce qui fait exister des ensembles de signes et ce qui permet d'actualiser des formes discursives en fonction de certaines règles de construction des énoncés (*ibid.* : 121). Dans la section suivante, nous verrons que ces considérations sont proches de certains travaux en linguistique[28].

Du reste, l'auteur rapproche apparemment l'archive et la langue à l'aide de l'énonciation. L'archive est définie comme une *pratique* « qui fait surgir une multiplicité d'énoncés comme autant d'événements réguliers, comme autant de choses offertes au

[27] Voir aussi Foucault (1966 : 262 *sq.*, 307-313 & 346-351) ou, pour un autre point de vue contemporain, Ricœur (1969 : 31-97).

[28] Cette conception du langage comme système est issue de l'approche structuraliste du langage ; cf. **infra, sect. 1.3.1.1.**

traitement et à la manipulation » (*ibid.* : 178-179). En effet, c'est toujours dans le cadre de pratiques discursives qu'il s'agit d'interpréter ce qui a été dit, écrit ou énoncé, puis conservé dans une archive. Ce rapprochement théorique entre archive et énonciation fait ressortir que « ce qui est dit », « écrit » ou « énoncé » peut être sauvegardé sous forme de discours que l'archéologue doit, selon Foucault, analyser en profondeur. Comme nous l'avons vu, cette sauvegarde est réalisée dans le cadre de pratiques d'archivage. L'une des missions principales des pratiques d'archivage est de maintenir les conditions d'interprétation pour que les connaissances du passé restent compréhensibles. Pourtant, même si *ce qui est dit*, *écrit* ou *énoncé* est interprété en fonction de signes et de règles linguistiques et discursives, l'archive existe aussi à travers des objets matériels qui sont transformés, échangés, combinés, décomposés, recomposés, voire détruits (*ibid.* : 145). C'est pourquoi l'archive doit être rapportée à des communautés d'individus, à des pratiques sociales et culturelles qui produisent, manipulent et interprètent des discours, tout en prenant en compte les supports, les contextes et les normes qui entourent la production des énoncés.

Ce rapprochement entre l'archive et le discours mis en évidence par Foucault a ensuite été étudié par Dominique Maingueneau (1991). Celui-ci remarque que l'objectif de la définition foucaldienne de l'archive est de « dégager l'importance d'une instance largement occultée, celle des communautés que présuppose l'énonciation d'une archive » (*ibid.* : 9-28). Concernant les pratiques d'archivage, l'auteur fait ressortir le problème des médiateurs, c'est-à-dire le problème de la médiation en acte entre un groupe social et des discours. Maingueneau met en évidence deux dimensions interdépendantes de l'archive foucaldienne : (1) l'archive et (2) la communauté discursive, à savoir : « le groupe ou le réseau de groupes à l'intérieur desquels sont produits, gérés, les textes relevant de cette archive » (*ibid.* : 190), car « l'espace d'énonciation lui-même, loin d'être un simple support contingent, un "cadre" extérieur au discours, suppose la présence d'un *groupe spécifique*, sociologiquement caractérisable, qui n'est pas un assemblage fortuit de "porte-paroles" » (*ibid.* : 188).

Plus récemment, Gabriella Giannachi (2016 : 182) remarque que les archives ont tenu un rôle fondamental dans l'administration des individus et des groupes sociaux, car les archives influencent « la manière dont nous stockons, préservons et générons la connaissance », mais aussi « la manière dont nous partageons la connaissance et utilisons ce partage de la connaissance » (*ibid*.). En suivant la définition foucaldienne de l'archive,

Giannachi affirme elle aussi que les archives impliquent une relation complexe du présent avec le passé, relation qui se base sur des connaissances à interpréter et dont il s'agit de comprendre le fonctionnement, de sonder les strates qui sont formées par les archives. Pour l'auteure, les archives constituent un véritable *instrument* qui permet de transformer la société. Cet instrument, il faut désormais le rattacher aux stratégies des institutions culturelles qui combinent différents types de médiation culturelle, notamment grâce à l'hypersphère et aux nouveaux médias, tout en s'appropriant les caractéristiques d'anciennes formes et pratiques discursives.

En effet, l'informatique et le réseau Internet transforment les approches et les outils documentaires pour organiser et diffuser des connaissances liées aux cultures et sociétés humaines[29]. Dans *Patrimoine et numérique : technique et politique de la mémoire*, Bruno Bachimont (2017b : 134-135) remarque que l'interprétation des connaissances consignées dans des archives oscille entre deux modalités fondamentales. D'une part, le savoir qui est construit à partir des documents. D'autre part, la compréhension herméneutique de ce savoir, puisque l'historien doit chercher à saisir un « vécu passé » à travers ces documents.

Cette distinction fait ressortir deux points importants que nous discuterons dans la suite de cette **première partie**.

- Les documents conservés dans une archive sont considérés comme des traces physiques, des supports d'inscription qui consignent des connaissances. Les archives désignent alors un ensemble d'objets matériels qui sont organisés en fonction d'un traitement documentaire[30].
- Ces connaissances sont seulement accessibles à travers une pratique interprétative visant à déchiffrer le sens qui est exprimé par le moyen de ces objets. L'archive, en tant qu'ensemble cohérent de connaissances, se situe à la jonction des pratiques documentaires et d'une pratique herméneutique[31].

---

[29] Sur l'apport de l'informatique dans les pratiques documentaires, voir notamment les travaux de Fayet-Scribe (2000), de Chaumier & Gicquel (2003) et de Chaumier (2006 : 11-99 & 107-114). Nous considèrerons également l'impact du traitement informatique sur la théorie des médias et les méthodes d'analyse d'images dans le **chap. 4, sect. 4.1.2.**

[30] Cf. **supra, sect. 1.2.**

[31] En simplifiant, l'herméneutique désigne la science et les théories de l'interprétation des signes, des symboles et des textes ; sur ce sujet, cf. l'ouvrage de Paul Ricœur (1969) ou, plus récemment, François Rastier (2001).

À ce stade, il faut également remarquer que l'*archive* ne doit pas être confondue avec le *corpus*. Dans ses travaux sur la linguistique des corpus[32], François Rastier propose une réflexion critique importante sur la question du corpus. Simplement, un corpus est la sélection de ressources documentaires dans un ensemble d'archives en vue d'une utilisation pratique de ces ressources[33]. La composition du corpus dépend fortement d'un point de vue et de l'usage des ressources. Pour Rastier, la structure du corpus dépend (1) soit d'une conception documentaire qui « ne retient que des variables globales caractérisant les documents, sans tenir compte de leur caractère textuel, ni de leur structure » ; (2) soit d'une conception philologique-herméneutique qui prend en compte les rapports entre des textes. Dans le premier cas, le corpus est un « échantillon de la langue », « un réservoir d'exemples ou d'attestations ». Dans le second, la sélection aboutit sur des corpus réflexifs, parce que leur composition est « réfléchie selon le principe critique qui est traditionnellement celui de la philologie, numérique ou non » (cf. Rastier, 2002).

Il faut donc bien distinguer l'archive du corpus, et le *corpus documentaire* du *corpus réflexif*. L'archive est un ensemble de ressources accessibles qui n'est pas constitué selon un point de vue et dans le but d'une recherche spécifique. Le corpus documentaire est « un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages », tandis qu'un corpus réflexif est une sélection de textes qui sont rassemblés « en tenant compte des discours et des genres » et en vue d'une utilisation scientifique (*ibid.*). Ce point est important, car cette distinction subsiste dans les archives numériques[34]. Pour Rastier, le traitement informatique des langues doit être considéré comme un instrument et non pas comme un modèle théorique, parce que, selon l'auteur, « la linguistique appartient pleinement aux sciences de la culture ». Par conséquent, le traitement automatique de l'information « n'est pas un objet scientifique, mais un objectif » (*ibid.*). Il faut donc envisager que le traitement

[32] Sur les liens entre linguistique et sémiotique des corpus, voir Rastier (2010 : 13-38) et surtout Rastier (2011 : 33-52).

[33] Voir par exemple : https://www.cnrtl.fr/definition/corpus.

[34] Les *archives numériques* sont comprises comme un ensemble de « ressources numériques » accessibles, tandis que les *corpus numériques* sont le résultat d'une activité d'organisation de ces ressources. Même si l'informatique est sans doute une aide précieuse pour constituer des corpus selon une conception documentaire, ce sont le point de vue et les choix assumés par l'utilisateur de ces corpus qui permettent de considérer un ensemble de ressources comme un corpus réflexif. Voir aussi les discussions sur le patrimoine numérique dans le **chap. 2, sect. 2.3.**

informatique (des langues, des textes, des corpus) concerne principalement une dimension instrumentale et que ce traitement n'est pas un équivalent de l'herméneutique, qui est la science d'interprétation des textes. Autrement dit, traiter et sélectionner automatiquement des ressources à l'aide d'un instrument informatique, ce n'est pas les interpréter.

Les paragraphes précédents indiquent que les *discours* et l'*énonciation* occupent une place importante dans certaines réflexions théoriques sur l'archive. Ceci n'est pas un hasard, car la définition foucaldienne de l'archive est contemporaine de travaux importants en sciences du langage qui ont beaucoup discuté ces deux notions.

Des explications sur ces travaux, qui s'étendent des origines de la linguistique à la sémiotique des pratiques, sont donc indispensables pour comprendre ce que nous désirons retenir de cette définition théorique de l'archive.

Les sections suivantes présenteront dans les grandes lignes la perspective suivie par certains travaux en sciences du langage et la théorie du langage qui en résulte. Si cette théorie s'est d'abord concentrée sur l'analyse du sens à travers les notions de signe et de fonction sémiotique, nous verrons que des sémioticiens se sont ensuite intéressés aux structures et aux opérations qui sont utilisées pour produire et pour manifester des discours dans un contexte de la communication (verbale ou écrite).

Ensuite, nous rapprocherons la définition théorique de l'archive, en tant que pratique énonciative en lien avec une communauté discursive, des travaux de Jacques Fontanille sur les pratiques, les objets-supports et les stratégies. Ce rapprochement servira à définir le volet « sémiotique » de notre approche sémio-rhétorique.

1.3.1.1. Théorie du langage (1) : langage et analyse des signes

Ferdinand de Saussure est à l'origine des travaux en sémiologie qu'il définit comme la « science qui étudie la vie des signes au sein de la vie sociale » (Saussure, 1995 [1916] : 33). Dans le *Cours de linguistique générale*, Saussure aborde en effet plusieurs principes fondamentaux de la linguistique[35] :

[35] Voir par exemple Soutet (2017 [1995]) et Martin (2018 [2002]).

- La distinction entre langue, langage et parole (cf. *ibid.* : 23-43).
- La distinction entre signifiant et signifié (cf. *ibid.* : 97-100).
- Le principe d'arbitrarité du signe (cf. *ibid.* : 100-103 ; les signifiants de certaines unités de langage n'ont pas de lien avec le signifié).
- La distinction entre synchronie (état d'une langue à un moment donné) et diachronie (évolution dans le temps de cet état ; cf. *ibid.* : 114-140).
- Le principe d'axe paradigmatique (ensemble virtuel des éléments de la langue) et d'axe syntagmatique (suite d'éléments du discours ; *ibid.* : 170-175).

Par rapport à la linguistique de son époque, Saussure propose une approche qui considère le langage comme un système. Le « langage » correspond à la *capacité* qui permet à chacun de communiquer et d'interagir avec les autres. Tandis que la « langue » est plutôt l'*outil* qui nous permet de communiquer. Contrairement au langage, la langue est un système complexe comprenant des mots (le lexique) et des règles de fonctionnement (la grammaire) ; la langue n'est donc pas un système commun à tous les humains, mais seulement à un groupe de personnes, et l'utilisation de ce système nécessite un apprentissage. Enfin, la « parole » désigne l'utilisation concrète d'une langue par un individu, autrement dit la manière (la prononciation, l'accent, l'intonation, le rythme et le choix des mots) d'utiliser la langue comme un outil.

Pour Saussure, l'analyse des langues en tant que système est conçue comme l'étude de signes linguistiques. L'idée est d'opposer deux dimensions du signe : le « signifié » et le « signifiant » ; par exemple, il faut distinguer le concept de 'table' de l'image acoustique *table*. En outre, le langage humain est le seul à posséder une double articulation. Plus précisément, les signes linguistiques concernent deux niveaux d'articulation du langage. D'une part, les unités dites de première articulation, les « unités minimales de signification », qui sont constituées d'un signifiant et d'un signifié[36]. D'autre part, les unités dites de seconde articulation, qui sont les « unités minimales distinctives » (par exemple : les sons distinctifs propres à une langue, constitués de signifiants, mais pas de signifiés).

Nous l'avons suggéré, le rapprochement entre l'archive et l'énonciation proposé par Foucault n'est pas le fruit du hasard. Dès la fin des années 1960, l'énonciation

[36] En sémantique, le *sème* est considéré comme l'unité minimale différentielle de la signification (du sens d'un mot). Les signifiés se composent de plusieurs « sèmes » dont la répétition constitue des *isotopies* ; cf. par exemple Rastier (1989).

commence à occuper une place centrale dans les travaux en linguistique. Au même moment, Foucault définissait l'archive comme « le système général de la formation et de la transformation des énoncés » et comme « la loi de ce qui peut être dit » (Foucault, 1969 : 177). Il semble évident pour l'auteur que ces énoncés se fondent à la fois sur des règles linguistiques (la « langue » de Saussure) et sur des actes et des pratiques (la « parole ») qui permettent aux énoncés d'apparaître.

Cette conception structuraliste du langage a ensuite été prolongée par les travaux de Louis Hjemslev (1971) qui propose quant à lui d'opposer « plan du contenu » et « plan de l'expression » ; en simplifiant, les signifiés concernent le plan du contenu et les signifiants correspondent au plan de l'expression , tandis que la « fonction sémiotique » (ou *sémiosis*) est l'articulation de ces deux plans[37].

Il est admis que les propositions théoriques de Saussure et de Hjemslev ont progressivement été discutées et complétées par d'autres travaux. La distinction entre plan de l'expression et plan du contenu est également centrale dans la théorie du langage élaborée par Algirdas Julien Greimas (cf. **infra**). Umberto Eco remarque en outre que la notion de fonction sémiotique ne se limite pas au seul champ de la linguistique.

> La notion de fonction sémiotique a influencé nombre de théories du signe, et a pu trouver des applications en dehors du domaine sémiotique. […] Cette notion de fonction ne se borne pas à définir des signes comme les mots de la langue, […] mais peut être étendue à des domaines bien plus complexes : la relation qui s'établit entre un vaste ensemble textuel (par exemple un livre, ou un tableau) et son contenu constitue une telle fonction (Eco, 1992 [1988] : 121)[38].

Même si plusieurs recherches en sémiotique ont poursuivi et approfondi ces réflexions sur la signification (cf. par exemple Klinkenberg, 1996), il faut souligner que ces recherches ont progressivement remarqué les limites de la notion de *signe*[39]. À ce

---

[37] Sur le « projet saussurien » et sa continuation dans la glossématique, cf. Hénault (1992 [1997] : 26 *sq*. & 55 *sq*.). Sur la théorie du langage de Hjemslev, voir l'ouvrage de Sémir Badir (2014).

[38] Au sujet des approches sémiotiques de la peinture et de l'énonciation visuelle, cf. **chap. 3, sect. 3.2.**

[39] Umberto Eco distingue ailleurs une première génération et une deuxième génération parmi les théories sémiotiques (cf. Eco (1989 [1979] : 13-14). D'une part, « une théorie des codes et de la compétence encyclopédique selon laquelle une langue […] permet (ou devrait permettre) de prévoir toutes ses actualisations discursives possibles, toutes les utilisations possibles dans des circonstances et des contextes spécifiques ». D'autre part « une théorie des règles de génération et d'interprétation des actualisations discursives » ; cf. **infra, sect. 1.3.1.2.**

propos, Rastier affirme que cette notion est un « artefact des sémioticiens ». Si les signes sont compris comme les « unités les moins complexes, cela n'entraîne pas qu'elles soient fondamentales, au sens où toutes les autres se réduiraient à elles sans reste » (Rastier, 2006 : 3). D'abord, parce que l'identification d'un signe est le résultat d'un parcours d'interprétation. Ensuite, parce qu'il semble évident que les signes ne sont pas manifestés de manière isolée, mais qu'ils sont manifestés à travers « des formations complexes dont la segmentation est toujours problématique, parfois impossible » (*ibid.*).

Cependant, il faut quand même rendre compte de la dimension double de la signification que nous avons évoquée (plan de l'expression et plan du contenu). C'est pourquoi, selon Rastier, certains travaux en sémiotique se sont orientés « vers une refondation interprétative de la sémantique, et au-delà, de la sémiotique, à partir de thèses qui intéressent le palier du signe et celui du texte » (*ibid.* : 5)[40]. Cela implique désormais que :

- Un signe se fonde nécessairement sur un parcours interprétatif.
- L'unité fondamentale de l'analyse rhétorico-herméneutique est le texte.
- L'unité linguistique maximale est le corpus de référence[41].
- Le sens est le résultat de « mises en relations internes et externes au texte », c'est-à-dire de la « rencontre d'un contexte et d'un intertexte » (*ibid.*).

Ainsi, les sémioticiens ne se limitent plus uniquement à l'analyse des signes linguistiques, parce qu'il s'agit aussi de faire ressortir les relations d'un texte par rapport aux autres textes qui s'y rapportent (intertextualité), ou encore l'expérience du lecteur qui interprète chaque texte dans un contexte.

Pour mieux comprendre ce passage d'une approche centrée sur l'analyse des *signes* à une approche centrée sur l'analyse des *textes*, il faut encore prendre acte de deux points importants : (1) les propositions de la sémiotique de Greimas et (2) la place qu'occupe l'énonciation dans cette approche sémiotique.

[40] D'après Rastier, la « théorie des parcours interprétatifs permet de rendre compte du lien problématique entre les deux plans du langage » (Rastier, 2006 : 4-5). Pour plus de détails théoriques au sujet des travaux sur la sémantique interprétative, voir aussi Rastier (2009 [1987]).

[41] Cf. **supra, sect. 1.3.1.**

1.3.1.2. Théorie du langage (2) : énonciation, discours et textes

Greimas[42] est souvent considéré comme le continuateur des travaux initiés par Saussure et des aménagements proposés ensuite par Hjemslev, quoique la théorie du langage de Greimas se différencie des travaux de Hjemslev (à ce propos, voir les explications de Driss Ablali, 2001). Dans *Histoire de la sémiotique,* Anne Hénault commente les propositions théoriques de Greimas et remarque notamment que la finalité de cette approche sémiotique « se caractérise par son aptitude à réaliser des analyses concrètes, en rigoureuse cohérence avec la base épistémologique » (Hénault, 1997 [1992] : 99 *sq*. ; nous citons 99)[43]. Selon l'auteure, cette théorie du langage est le fondement d'une méthode pour analyser des discours qui sont actualisés par le moyen d'un acte d'énonciation (*ibid.* : 104).

Cette intégration de l'énonciation dans les sciences du langage est étroitement liée aux travaux d'Émile Benveniste (1974 : 79-88), qui a lui aussi remarqué les limites du projet scientifique de Saussure, notamment lorsqu'il invitait à le dépasser par deux voies : d'une part, « par l'ouverture d'une nouvelle dimension de signifiance, celle du *discours*, que nous appelons *sémantique*, désormais distincte de celle qui est liée au *signe*, et qui sera *sémiotique* » ; d'autre part, « par l'élaboration d'une métasémantique qui se construira sur la *sémantique de l'énonciation* » (Benveniste, 1974 : 66 ; nous soulignons). Concernant la seconde voie, l'énonciation est comprise comme la « mise en fonctionnement de la langue par un acte individuel d'utilisation », dans le sens où l'acte de produire un énoncé « est le fait du locuteur qui mobilise la langue pour son compte » (*ibid.* : 80).

L'énonciation est donc surtout envisagée comme « le fait du locuteur » et « dans les caractères linguistiques qui marquent cette relation [du locuteur à la langue] » (*ibid.*).

---

[42] Voir par exemple Greimas (2002 [1966] & 1970) ou encore Coquet (1982), et surtout le *Dictionnaire raisonné de la théorie du langage* de Algirdas J. Greimas & Joseph Courtés (1993). Cette théorie du langage propose un cadre méthodologique pour encadrer l'analyse sémiotique des discours ; pour un résumé des applications analytiques que propose cette approche sémiotique, voir par exemple Hébert (2007) qui présente douze dispositifs qui sont construits autour de différentes notions issues de la sémiotique (homologation ; carré sémiotique ; carré véridictoire ; schéma tensif ; modèle actantiel ; programme narratif ; schéma narratif canonique ; analyse figurative, thématique et axiologique ; analyse thymique ; analyse sémique ; dialogique ; graphe sémantique).

[43] Notons au passage que pour cette approche, il s'agit de distinguer le « savoir-faire » (compétence cognitive qui rend possible une activité cognitive) et le « savoir-être » (compétence épistémique qui valide le savoir sur les objets) ; cf. Greimas & Courtés (1993 : 321).

Pourtant, Benveniste affirme aussi que l'énonciation est une activité pratique d'où émanent des entités analysables, les énoncés, qui n'existent que par rapport aux expériences des locuteurs et dans les « réseaux d'individus que l'énonciation crée » (Benveniste, 1974 : 84). Ces énoncés sont le résultat d'un acte linguistique, qui a concrètement été accompli dans une situation de communication, et ils sont construits et manifestés par le moyen de signes.

En effet, dans le *Dictionnaire raisonné de la théorie du langage* (*Dictionnaire*), l'énonciation est définie de deux façons : (1) soit comme « la structure non linguistique (référentielle) sous-tendue à la communication linguistique », (2) soit « comme une instance linguistique, logiquement présupposée par l'existence même de l'énoncé (qui en comporte des traces ou marques) » (Greimas & Courtés, 1993 : 126). Cette définition est importante, car elle distingue, dans le premier cas, l'énonciation comprise en tant que situation qui rend possible la communication (verbale ou écrite) et, dans le second, l'énonciation accomplie, qui est comprise comme l'instance de médiation « qui assure la mise en énoncé-discours des virtualités de la langue » (*ibid.*). De plus, étant donné que cet énoncé-discours est le résultat d'un acte linguistique, il « comporte souvent des éléments qui renvoient à l'instance de l'énonciation » (*ibid.*) ; sur le plan linguistique, cette instance se caractérise notamment par l'emploi de déictiques (« ici », « là », « maintenant », « demain », « hier », etc.), le recours à certains pronoms (« je », « tu », « il », etc.) et à certains temps de préférence (présent, passé composé, passé simple, etc.). Ce sont ces *marques discursives* (déictiques, pronoms, temps) que la sémiotique vise à analyser en faisant ressortir l'importance des structures discursives et textuelles qui sont observables dans les résultats d'une énonciation accomplie.

À ce propos, nous pouvons rappeler ici quelques précisions importantes du *Dictionnaire* sur le fonctionnement de ces structures discursives. La *discursivisation* est réalisée au moyen des opérations de *débrayage* qui composent et mettent en œuvre l'instance de l'énonciation – celle-ci étant logiquement présupposée par l'existence de l'énoncé. Cette opération de débrayage comprend trois composantes principales : (1) l'actorialisation institue les acteurs du discours ; (2) la temporalisation organise les successions temporelles du discours ; (3) la spatialisation permet de situer le discours dans une espace pour produire le discours. Dans cette perspective, le débrayage est l'opération par laquelle l'instance de l'énonciation produit dans un énoncé une *représentation d'actants*, mais dans un autre lieu et un autre temps que celui de la

production de l'énoncé. En simplifiant, il s'agit de distinguer : (1) le débrayage actantiel (qui projette dans l'énoncé un « non-je ») ; (2) le débrayage temporel (qui postule un « non-maintenant » qui est distinct du « temps de l'énonciation ») ; (3) le débrayage spatial (qui oppose un « non-ici » au « lieu de l'énonciation »).

En outre, le débrayage est une opération antérieure à l'*embrayage*, qui est le retour à l'énonciation en refusant les projections du débrayage (non-je, non-maintenant, non-ici) et en lui opposant un « je-ici-maintenant » simulé.

Ces travaux sur l'énonciation font ressortir que le producteur d'un énoncé s'approprie la langue en projetant des actants de l'énonciation ou des actants de l'énoncé dans son discours. Dans le premier cas, le débrayage est « énonciatif », et il en résulte un premier type d'unité discursive : les *formes de l'énonciation énoncée* (par exemple : des récits à la première personne, des dialogues). Dans le second cas, le débrayage est « énoncif » et résulte d'un second type d'unité : les *formes de l'énoncé énoncé*, qui s'apparente plutôt à des discours « objectifs » à la troisième personne. Pour Benveniste et pour Greimas, l'utilisation de la troisième personne (« il ») dans un discours révèle que les locuteurs absents le sont toujours par rapport à des locuteurs présents dans une situation de communication « je-tu ». La troisième personne est alors comprise comme un produit dérivé de cet ancrage à une situation originaire « je-tu ». Cette réduction de la troisième personne à la situation « je-tu » laisse entendre que la source des simulacres discursifs provient nécessairement d'un « je-ici-maintenant » et, par conséquent, que l'ancrage principal du discours est toujours une situation originaire qui est externe à l'énoncé.

Dans la section précédente, nous avons vu que, passant de l'analyse des signes à celle des énoncés, la théorie structuraliste du langage accorde une place centrale à l'énonciation. En résumé, nous pouvons remarquer deux points importants. Premièrement, les situations de communication (verbale ou écrite) impliquent nécessairement un *acte d'énonciation* visant à produire un *énoncé-discours* (ces énoncés-discours sont produits par un énonciateur à destination d'un énonciataire). Deuxièmement, l'énonciation est comprise soit comme la *situation de référence* qui sous-tend la communication (énonciation pragmatique) soit comme le *dispositif qui est mis en œuvre dans le discours* au moyen d'*opérations* et de *marques discursives* renvoyant à l'*instance de l'énonciation* ou encore à des simulacres qui sont produits par l'*instance d'énonciation* (énonciation discursive).

Jusqu'à présent, nous n'avons pas clairement différencié les textes des discours lorsque nous affirmions que la sémiotique est passée de l'analyse des signes à celle des textes. Dans le *Dictionnaire*, Greimas et Joseph Courtés (*op. cit.* : 389) opposent le texte au discours, tout en remarquant que le « terme texte est souvent pris comme synonyme de discours » (*ibid.* : 390). La *textualisation* est définie comme l'ensemble des procédures visant à constituer un contenu discursif avant la manifestation du discours (*ibid.* : 391). Le *texte*, en tant que « grandeur considérée antérieurement à son analyse » (*ibid.* : 390), est uniquement constitué par « des éléments sémiotiques conformes au projet théorique de la description » (*ibid.*).

D'un autre côté, le *discours* est le résultat de l'énonciation comprise comme une « mise en discours » (nous l'avons vu, cette mise en discours, ou discursivisation, est la prise en charge des structures sémio-narratives et leur transformation en structures discursives, tandis que l'instance de l'énonciation est le « lieu de génération du discours », cf. *ibid.* : 106). Néanmoins, un passage du *Dictionnaire* souligne que « le problème de savoir ce qu'est le discours – au sens sémiotique – reste entier » (*ibid.* : 105).

Cette opposition entre texte et discours n'est pas acceptée à l'unanimité. Rastier[44] a par exemple remarqué que « l'opposition texte/discours » est problématique : « le mot texte, comme d'ailleurs le mot discours, reste au singulier ; on demeure ainsi dans des caractérisations abstraites, alors même qu'une typologie des textes, des genres[45] et des discours reste nécessaire pour le développement théorique comme pour les applications ». C'est pourquoi l'auteur affirme que ces deux notions ne sont pas séparables

[44] Voir « Discours et texte », cf. http://www.revue-texto.net/Reperes/Themes/Rastier_Discours.html#head.

[45] En effet, les actes concrets d'énonciation aboutissent sur la production de *discours* que le sémioticien rattache à des *genres*. Un « genre discursif » est contraint par le statut des producteurs et des récepteurs des discours, mais aussi par le contexte spatio-temporel, le type de support des discours ou encore les thèmes qui sont considérés ou présents dans ces discours. La question du genre est un problème délicat que nous ne discuterons pas davantage. Sur ce sujet, voir Rastier (2001) ou encore l'entrée « genre » dans le *Dictionnaire de sémiotique* [En ligne] de Louis Hébert : « Tout produit sémiotique (texte, image, etc.) relève […] d'un ou de plusieurs genres. Un genre textuel peut être défini sous différents angles, notamment soit en tant que programme de normes, soit en tant que type associé à ce programme, soit en tant que classe de textes qui relèvent de ce type et/ou de ce programme. Un genre est un programme de prescriptions (éléments qui doivent être présents), d'interdictions (éléments qui doivent être absents) et autres modalités déontiques (facultativités, permissivités, optionnalités, etc.) qui règlent la production et l'interprétation (la réception) des textes. […] Le genre, avec le corpus, le contexte externe, etc., fait partie des unités globales déterminant les unités locales, en l'occurrence ici un produit sémiotique donné » ; cf. https://semiotique.org/wp-content/uploads/2021/07/Dictionnaire-sémiotique.pdf.

théoriquement ni méthodologiquement : « les textes et les discours (au pluriel !) se trouvent [selon Rastier] exactement au même niveau ontologique ».

Per Aage Brandt (2018) a quant à lui formulé cette différence en hiérarchisant les deux notions : d'un côté, le texte est « un signe composé de phrases », où chaque phrase « configure sa propre énonciation » (le sens du texte est alors son « contenu trans-phrastique ») et de l'autre, le « texte s'inscrit à son tour dans un *discours* », le milieu langagier « qui détermine un certain nombre de propriétés des textes ». Les discours sont donc compris comme des « entités plus abstraites que les textes » (*ibid.*), même s'il est possible de reconnaître certains discours institutionnalisés (discours religieux, discours politique, discours scientifique, etc.).

Dans la section suivante, nous présenterons une autre approche sémiotique qui met en évidence deux dimensions fondamentales de la relation entre archive et énonciation, à savoir : (1) la dimension pratique[46] de l'énonciation et (2) la matérialité des énoncés.

En renouvelant la perspective initiée par Greimas, Jacques Fontanille a en effet déplacé l'attention de l'énonciation énoncée à l'énonciation en acte, notamment en accordant une place importante aux différents types d'expérience de l'interprète dans la constitution d'un « parcours génératif »[47] (cf. Fontanille, 2001).

1.3.2. Sémiotique du discours et praxis énonciative

Dans *Sémiotique et littérature* (Fontanille, 1999 : 7), Fontanille adopte la perspective d'une « signification en acte » qui est contrôlée par une « énonciation présente et vivante ». De même, dans l'avant-propos de *Sémiotique du discours* (Fontanille, 2003 [1998]), l'auteur précise :

---

[46] Jean-Marie Floch (2002 [1990]) est l'un des premiers à avoir mis en évidence cette dimension pratique dans une analyse sémiotique.

[47] Pour Greimas et Courtés (*op. cit.* : 157-160), le « parcours génératif » comprend deux composantes principales (syntaxique et sémantique) et diverses sous-composantes (concernant les *structures sémio-narratives* : d'un côté, la syntaxe fondamentale et la sémantique fondamentale, et de l'autre, la syntaxe narrative de surface et la sémantique narrative ; concernant les *structures discursives* : syntaxe discursive, sémantique discursive). Néanmoins, un « point de friction essentiel » persiste entre la sémantique textuelle de Rastier et les « développements post-greimassiens ». Pour Nicolas Couégnas, ce point de friction « réside a priori dans le rôle joué par la générativité, strictement opposable à l'interprétation » (cf. Couégnas, 2022). À ce propos, voir aussi le point de vue de Rastier (2000 & 2017).

> [Choisir] le point de vue du discours en acte, c'est en effet choisir d'observer la manière dont la praxis sémiotique schématise notre expérience pour en faire des langages, plutôt qu'observer et découper des unités minimales. La sémiotique que nous envisageons, dans la perspective définie par Greimas il y a une trentaine d'années, est celle des *ensembles signifiants*, mais des ensembles signifiants en construction et en devenir (Fontanille, 2003 [1998])[48].

Ce déplacement, de l'énonciation énoncée à l'énonciation en acte, est une étape importante dans le renouvellement de la sémiotique, car il étend l'analyse sémiotique au-delà des notions de discours et de textes. Pour ce faire, Fontanille considère la signification mise en œuvre dans les discours par rapport à des tensions en lien avec le sentir et le corps (cf. Fontanille 2004 & 2011)[49]. Selon l'auteur, la mise en œuvre de la signification peut être divisée en trois phases correspondant chacune à un « champ du discours » : la signification est soit « émergente », soit « en cours » ou soit « accomplie » (cf. *ibid.* : 274-275).

Dans sa phase émergente, le champ du discours correspond à un « champ de présence » qui rassemble les tensions qui relèvent du sensible et de l'intelligible ; les relations entre ces tensions forment différents « schémas tensifs »[50]. Tandis que dans la phase durative du procès (la signification « en cours », lorsque les valeurs prennent une « forme »), le champ du discours est un assemblage de « schémas discursifs ». Enfin, lorsque la signification est accomplie, le champ du discours est compris comme un « réseau de différences »[51].

---

[48] Cf. https://www.unilim.fr/pages_perso/jacques.fontanille/textes-pdf/BSemiotiquediscoursintro.pdf.

[49] Notons que les travaux de Maurice Merleau-Ponty sur le corps propre occupent une place importante dans les réflexions sur le corps de Fontanille ; cf. notamment Merleau-Ponty (2010).

[50] Avec Claude Zilberberg, Fontanille a proposé une « structure tensive de la signification » où l'apparition des valeurs intelligibles dépend des valences perceptives de l'*intensité* et de l'*étendue* (cf. Fontanille & Zilberberg, 1998). Ces deux valences sont associées à deux opérations élémentaires : la *visée* et la *saisie*. D'une part, la visée « opère sur le mode de l'intensité : le corps propre se tourne alors vers ce qui suscite en lui une intensité sensible (perceptive, affective) », d'autre part, la saisie « opère en revanche sur le mode de l'étendue : le corps propre perçoit des positions, des distances, des dimensions, des quantités » (Fontanille, 2003 [1998] : 93).

[51] Par ailleurs, Denis Bertrand affirme que trois paramètres définissent le discours : (1) l'énonciation (la « mise en discours »), (2) l'interaction (« action réalisée et effets produits sur l'énonciataire ») et (3) l'usage (« part impersonnelle de l'énonciation que convoque – ou révoque – un énonciateur individuel dans sa mise en œuvre du discours) » ; cf. http://denisbertrand.unblog.fr/glossaire-de-semiotique/6-glossaire-de-termes-de-semiotique-a-e/ (ce glossaire renvoie à Bertrand, 2000).

Pour expliquer pourquoi cet apport est important pour définir notre approche sémio-rhétorique de la communication en ligne, nous nous appuierons sur les propositions de Fontanille dans *Pratiques sémiotiques* (Fontanille, 2008). Dans cet ouvrage de synthèse, l'auteur a formulé une approche qui élargit la théorie du discours et de l'énonciation que nous avons rapidement présentée dans la section précédente. Nous avons vu que, pour Greimas, les discours sont considérés comme le résultat d'un acte d'énonciation accompli. Dans ce sens, le discours désigne la *réalisation du procès sémiotique manifesté sous forme de texte*[52].

Pour Fontanille, l'interprétation est un *processus discontinu* où il ne faut pas confondre le *plan d'expression* qui est décrit ou analysé (*ibid.* : 57-58 & 293-299). L'auteur opte pour une approche *générative*, qui s'oppose ici à l'approche *génétique* (dans cette seconde approche, la genèse de l'objet est située sur une « ligne du temps » et s'accomplit en « une suite de formes successives », cf. Greimas & Courtés, *op. cit.* : 161). En simplifiant, l'approche de Fontanille répartit l'analyse du *parcours génératif du plan de l'expression* en fonction de six niveaux de pertinence (la « pertinence » est une règle de description scientifique qui sert à limiter les « nombreuses déterminations (ou les traits distinctifs) possibles d'un objet », cf. Greimas & Courtés, *op. cit.* : 276) :

- Le niveau des signes.
- Le niveau des textes-énoncés.
- Le niveau des objets-supports.
- Le niveau des scènes pratiques.
- Le niveau des stratégies.
- Le niveau des formes de vie.

Selon Fontanille, l'*énonciation en acte* s'accomplit simultanément à travers différents plans d'immanence. Chaque plan d'immanence particulier engage un type d'expérience. Ainsi, l'expérience de la « figurativité » est caractérisée au « niveau des signes », tandis que l'expérience de la « cohérence et la cohésion interprétative » des discours correspond au « niveau des textes-énoncés »[53].

[52] Les textes sont des « systèmes clos », tandis que les pratiques sont des « systèmes ouverts » ; cf. Paolucci (2010).

[53] Fontanille affirme que les « conversions » entre un niveau de pertinence et un autre s'opèrent par des « mouvements énonciatifs ». Par exemple, dans le mouvement qui va du « niveau de l'objet » à celui de la « scène pratique », l'attention est focalisée sur les simulacres d'action possibles proposés à l'utilisateur,

Par rapport aux définitions du discours et des textes que nous avons présentées, Fontanille affirme en outre que le *texte-énoncé* est un « dispositif d'inscription » comprenant un « ensemble de figures sémiotiques organisées en un ensemble homogène grâce à leur disposition sur un même support ou véhicule » (ainsi, le discours oral est « unidimensionnel », tandis que les textes écrits et les images sont « bidimensionnels » ; cf. Fontanille, 2007a). L'expérience de la « corporéité » concerne donc un autre niveau de pertinence, celui des « objets-supports » (soit au niveau de la matérialité de l'objet soit au niveau des possibilités d'inscription que ces supports permettent d'aménager en vue de la manifestation et d'un usage dans des situations pratiques)[54].

Cela étant dit, le niveau des scènes pratiques de l'énonciation occupe une place centrale dans l'approche de Fontanille, même si cette énonciation ne se réalise pas nécessairement de la même manière à chaque niveau (cf. Fontanille, 2008 : 34, 36-39 & 42-58). Au contraire de Benveniste et Greimas, qui considèrent que l'énonciation discursive présuppose toujours une mise en discours dans une situation de communication « je-tu » (la situation originaire et externe à l'énoncé), Fontanille propose que la singularité de tout acte énonciatif ne devient pertinente que dans un second temps. En effet, pour l'auteur, la signification « émerge de la dynamique d'un cours d'action, et de l'activité d'une instance qui est elle-même en cours d'instauration » (Fontanille, 2014 : 4). Ce changement de perspective implique que la source du débrayage n'est plus un « sujet énonciateur », mais « l'événement en soi » issu de la pratique qui distribue, par la suite, les positions actantielles pertinentes à l'étude de la signification[55].

L'auteur s'appuie pour ce faire sur la théorie de la *praxis énonciative*[56], théorie qui concerne la confrontation de l'interprète à « diverses formes immanentes » en compétition à l'aide d'un processus à quatre mouvements, où chaque mouvement correspond à un mode d'existence sémiotique[57] (actualisation, réalisation, virtualisation,

---

tandis que le mouvement de la « pratique » au « niveau de l'objet » concerne plutôt l'étude des traces de l'action qui résultent des usages des objets.

[54] Sur les objets-supports, cf. **infra, sect. 1.3.3.** Nous ne discuterons pas ici la question des « formes de vie » (cf. Fontanille, 2015).

[55] À ce propos, cf. les précieux commentaires de Maria Giulia Dondero (2017a).

[56] Sur la « praxis énonciative », voir aussi Bertrand (2000 : 54-57).

[57] Nous trouvons une définition de la notion de mode d'existence dans le *Précis de sémiotique littéraire* de Denis Bertrand : « Le mode d'existence définit le statut variable des formes de présence sous lesquelles les objets sémiotiques se manifestent dans le discours (actants, modalités, temporalité, etc.). La tradition

potentialisation). Cette « praxis » agit par des opérations de tri qui hiérarchisent les différentes formes immanentes afin de « les conduire vers la manifestation » (cf. *ibid.* : 5-8), de telle sorte que l'énonciateur et l'énonciataire doivent constamment puiser dans une réserve de formes sémiotiques disponibles qu'il s'agit de réaliser, d'actualiser, de virtualiser et de potentialiser.

Cette relecture de la théorie de l'énonciation fait ressortir que, dans les pratiques discursives, l'activité d'une instance qui est « en cours d'instauration » concerne d'abord une énonciation « impersonnelle, diffuse et en mouvement », et non pas une situation de communication « je-tu ». En mettant l'accent sur l'*énonciation en acte*, l'intérêt de l'approche de Fontanille est en outre de proposer une alternative au point de vue qui considère que les textes sont les seules unités fondamentales de l'analyse sémiotique des discours[58]. En effet, même si la cohérence des discours est comprise au niveau du texte (le texte-énoncé est alors considéré comme le résultat d'un acte d'énonciation), l'énonciation *en cours* se réalise avant tout au sein d'une *scène pratique*, où l'interprète se confronte à plusieurs formes immanentes à travers des *objets-supports* qui sont eux aussi influencés par des normes et des usages pratiques.

1.3.3. Sémiotique de la communication en ligne : objets-supports, énonciation et stratégies de l'archivage numérique

Dans une section précédente, nous avons vu que les pratiques d'archivage consistent à rassembler et à conserver « ce qui a été dit ». L'archive est alors définie comme une pratique énonciative d'où surgit une multiplicité d'énoncés qu'il faut analyser en fonction de compétences discursives et dans le cadre d'une communauté discursive[59]. En retour,

saussurienne distinguait l'existence virtuelle (le système de la langue) et l'existence actuelle (sa réalisation dans la parole). Élargissant son champ d'application, la sémiotique a ajouté à la virtualisation et à l'actualisation un troisième mode d'existence, la réalisation. Ainsi, le contrat ou la manipulation virtualisent le sujet, la compétence l'actualise, l'action et la reconnaissance le réalisent. Les modalités du croire, du vouloir ou du devoir engendrent un sujet virtuel, le savoir et le pouvoir un sujet actualisé, le faire un sujet réalisé. Plus largement encore, on peut considérer que les modes d'existence concernent la modulation des formes de présence du sens au sein du discours […]. Cette approche permet de rendre compte, aussi près que possible de sa réalité fluctuante, des modulations du sens » (cf. http://denisbertrand.unblog.fr/glossaire-de-semiotique/m-p/). Sur les modes d'existence, cf. aussi Dondero (2017b) ou encore Badir (2020) et Colas-Blaise (2022).

[58] Cf. **supra, sect. 1.3.1.2.**

[59] Voir la définition proposée par Foucault et reprise par Maingueneau ; cf. **supra, sect. 1.3.1.**

les pratiques d'archivage visent à maintenir les conditions d'intelligibilité des textes-énoncés (les « dispositifs d'inscription ») qui sont conservés.

Pour terminer cette longue digression sur la théorie du langage et sur la sémiotique du discours, nous remarquerons que Fontanille a aussi mis en évidence que l'expérience de l'interprète est encadrée par des *stratégies* dont le but est d'aménager les situations pratiques où divers objets-supports sont manipulés.

Fontanille (2005 : 183-200) a présenté ailleurs la distinction entre (1) *support formel* et (2) *support matériel* dans les usages des objets d'écriture. Le support matériel est une proposition qui est limitée à un ensemble défini de possibles et qui présente des « lignes de force », des « tendances substantielles », tandis que le support formel « sélectionne et systématise », « opère un tri formel » et « fournit une syntaxe pour le plan de l'expression de l'écrit, qui décidera de la valeur et du fonctionnement syntagmatique des caractères » (*ibid.*). Le support formel est le résultat d'une extraction de propriétés qui émanent du support matériel. Le premier est à la fois la « structure d'accueil » des inscriptions et l'ensemble des « règles topologiques d'orientation, de dimension, de proportion et de segmentation » qui permettent de « faire signifier les caractères inscrits » sur le second (*ibid.*).

Comme nous l'avons vu, les textes-énoncés sont inscrits sur des objets-supports qui s'intègrent dans des scènes pratiques. À partir du support matériel, le support formel « engage, contraint et modalise en effet un type d'échange, une structure de communication idéale, et des rôles d'énonciation » (*ibid.*). Les objets d'écriture sont manipulés dans des « situations sémiotiques » que l'auteur comprend comme des « macro-objets » qui, d'une part, imposent des règles d'implémentation, d'accrochage ou de montage aux objets d'écritures, et d'autre part, proposent des parcours de lecture à l'observateur par le moyen des objets-supports. Ces situations sont analysées par Fontanille selon deux dimensions : celle de la *scène prédicative*[60] et celle de la *stratégie*, où se mettent en place, dans l'espace et dans le temps, les manipulations déictiques, modales et affectives.

[60] En linguistique, le prédicat désigne la partie d'une phrase qui caractérise ou dit quelque chose à propos du sujet grammatical de cette phrase. Pour Greimas & Courtés (*op. cit.* : 290), le prédicat est « la relation constitutive de l'énoncé », une fonction « dont les termes-aboutissants sont les actants ». Les auteurs distinguent deux types d'énoncés élémentaires (« deux sortes de relations-prédicats, constitutives de ces énoncés ») : (1) les énoncés de faire et (2) les énoncés d'état.

C'est dans ce sens que le niveau hiérarchiquement supérieur des stratégies régule celui des objets-supports et des scènes pratiques qui sont spatialement et temporellement en concurrence entre elles[61]. Plus précisément, les situations sémiotiques considérées par l'auteur comprennent deux dimensions complémentaires (Fontanille, 2007b) :

- La situation-scène est organisée autour d'une dimension prédicative.
- La situation-stratégie est organisée autour d'une dimension d'ajustement, puisque cette situation est plus ou moins prévisible et programmable, et parce que chaque scène prédicative est contrainte de s'ajuster, dans l'espace et dans le temps, aux autres scènes concomitantes.

Ces deux dimensions constituent une *situation dans son ensemble* comprenant des actes d'énonciation et des modalités de l'énonciation, puisque la « situation scène » comporte généralement une distribution de rôles et de valeurs. Pour Fontanille, ce sont justement ces actes et ces modalités d'énonciation qui « surdéterminent la scène prédicative de base (celle de l'objet-support) » et qui donnent à cette scène un « caractère stratégique ».

L'auteur prend pour exemple le cas des stratégies d'affichage. Chaque affichage (chaque scène) est contraint de s'ajuster à d'autres affichages et, plus généralement, aux autres dispositifs figuratifs présents dans un même environnement. L'auteur distingue un premier dispositif d'énonciation : (1) la *scène de l'affichage*, dispositif qui comprend « une ou plusieurs prédications » et des « actes d'énonciation » impliquant des rôles actantiels qui sont joués par l'affiche ou certains de ses éléments, ou encore par son support, par des éléments de l'environnement, par l'observateur, etc. Tandis que le second dispositif, (2) la *stratégie de l'affichage*, comprend le déploiement « figuratif et thématique » et le déploiement « spatial et temporel de l'affichage », ainsi que « toutes les contraintes (modales et isotopiques) issues de l'environnement » (Fontanille, 2007b ; cf. aussi Fontanille, 2008 : 179-215).

Les objets-supports participent aux deux dispositifs. D'un côté, les supports participent à la scène de l'affichage, parce que ces supports jouent souvent un rôle « dans les actes d'énonciation propres à chaque implantation ». De l'autre, ils participent aussi à la stratégie, étant donné que ces supports « sont un des éléments essentiels de la

[61] Pour Fontanille (2005 : 183-200), l'énonciation mise en œuvre par l'implantation stratégique d'un objet d'écriture est impersonnelle, car « les objets, qu'ils portent des écritures ou qu'ils n'en portent pas, agissent sur leurs usagers sans qu'il soit nécessaire que ces derniers reconnaissent en eux une identité subjective ». De plus, cette énonciation est également un « site », car elle « opère d'abord par référence déictique ».

récurrence, de la distribution et de la succession propres à chaque campagne, ainsi que des relations stratégiques entre les campagnes concomitantes » (*ibid.*). Plus précisément, l'affichage relève d'une sémiotique de la manipulation, où cette manipulation se décline en deux types : (1) une « manipulation restreinte », qui repose sur la scène prédicative et qui exploite les interactions entre les différents rôles qui composent cette dernière ; (2) une « manipulation étendue », qui repose sur la stratégie, dans la mesure où l'ajustement ou la programmation de l'affichage contribuent à capter l'intérêt du passant. Dans certains cas, il faut selon l'auteur envisager un « réglage tactique » entre ces deux types de manipulation, surtout dans les cas où ces deux types de manipulations apparaissent « incompatibles ou dissonantes ».

Ainsi, il ne s'agit pas seulement de choisir une affiche et les lieux où l'implanter. Pour Fontanille, il faut avant tout concevoir et programmer une « campagne d'affichage » qui décidera du nombre d'affiches, du type de variantes, des modalités de leur récurrence et de leur distribution dans l'espace et dans le temps. Fontanille suggère en effet que la dimension stratégique doit prendre en compte « la manière dont les différentes scènes spécifiques se conjuguent, se combattent ou se renforcent, sur chaque site ou dans chaque laps de temps », puisque chaque campagne d'affichage doit se confronter avec toutes les autres campagnes concomitantes qui occupent « les mêmes lieux et les mêmes périodes », et qui visent « à capter l'attention des mêmes passants-observateurs » (*ibid.*).

Ces dernières considérations sont essentielles pour notre approche. Les travaux de Fontanille que nous avons rapidement présentés mettent en évidence que les textes-énoncés sont souvent manifestés au croisement de différents « points de vue » et au sein d'un « réseau de positions » qu'il faut prendre en compte[62]. C'est en suivant ces travaux que nous étudierons les *stratégies de la communication en ligne des larges ensembles d'images*. Comme nous l'avons suggéré[63], de plus en plus d'archives et de collections d'images sont converties en ressources numériques par les institutions patrimoniales qui les gèrent, de telle sorte que ces images sont désormais diffusées sur Internet. Cette diffusion s'appuie sur l'expérience pratique d'une scène d'énonciation qui prévoit

[62] La dynamique du conflit est présente dans les travaux de Fontanille parce que, selon lui, l'énonciation nous positionne toujours au cœur d'un conflit de visions en concurrence. Dans *Les espaces subjectifs. Introduction à la sémiotique de l'observateur* (1989), l'auteur affirmait déjà que l'énonciateur se définit en tant qu'actant du savoir qui cache quelque chose à l'énonciataire et que l'énonciataire est toujours dans une position d'exploration et de recherche de ce savoir ; cf. **chap. 3, sect. 3.2.3.**

[63] Cf. **introduction**. Voir aussi **chap. 2, *passim***.

nécessairement une mise en scène qui implique des choix stratégiques visant à préparer et à manifester une *interaction* entre un utilisateur humain et des ressources numériques.

Dans le travail cité sur les objets d'écritures, Fontanille a également considéré le cas particulier du *support des textes informatiques* (cf. Fontanille, 2005 : 191-192 ; voir aussi Anis & Zara, 2005). Fontanille suggère que le support des textes qui sont manifestés par le moyen d'un environnement informatique correspond à « deux objets d'écriture différents et complets ». D'une part, un *support imperceptible* (le support « interne ») constitué à la fois de composants électroniques et d'un « support langagier », celui du codage informatique. D'autre part, un *support perceptible*, la « page-écran », qui s'appuie sur un support matériel « externe » (l'écran) et un support formel, dans ce cas-ci le logiciel qui génère la page-écran.

À ce propos, Nicole Pignier (2009 : 108-109) a déjà suggéré que la « sémiotique de la communication en ligne » doit s'articuler sur différents niveaux[64] qui correspondent chacun à une étude spécifique : (1) l'étude de l'interface-machine (design d'interface) ; (2) l'étude de l'interface des applications informatiques (les navigateurs) ; (3) l'étude des interfaces graphiques des sites Web (webdesign). Selon l'auteur, « ce qui fait sens pour l'usager », c'est l'*orchestration* de toute l'information présente sur le site Web et des applications (les fonctionnalités) de l'interface graphique (*ibid.* : 109). En effet, par rapport au support imperceptible (les composants électroniques et le langage informatique), ce support perceptible implique une *relation coénonciatrice* entre l'usager et le producteur du service informatique (*ibid.* : 108). L'agencement de l'interface, qu'il ne faut pas confondre avec le support matériel (l'écran) et le support formel (le logiciel), détermine la mise en scène de cette relation énonciative. À un autre niveau, cette énonciation est encadrée par des choix stratégiques dont le but est, dans le cas qui nous intéresse, l'aménagement d'un dispositif de communication en ligne.

1.4. Approche théorique du support numérique

À ce stade, nous pouvons remarquer que la conversion numérique impacte les pratiques d'archivage d'images pour plusieurs raisons. Cet impact concerne au moins trois niveaux de pertinence (parmi les six niveaux proposés par Fontanille) :

[64] Nous considèrerons plus précisément ces niveaux dans le **chap. 2, sect. 2.3.**

- Le niveau du support des images qui sont conservées dans le cadre des pratiques d'archivage, puisque ce sont ces supports qui sont actuellement transformés en archives numériques.
- Le niveau des scènes pratiques (où des textes-énoncés sont manifestés dans des situations concrètes, et où l'usager prend place pour manipuler des supports), étant donné que la manifestation et la manipulation des archives numériques s'appuient désormais sur la combinaison d'un support imperceptible (les composantes du système informatique) et d'un support perceptible (l'écran et la page-écran).
- Le niveau des stratégies qui (1) aménagent, dans l'espace et dans le temps, l'expérience de l'usager en prenant en compte la concurrence entre plusieurs services Web, et qui (2) encadrent le positionnement et le point de vue des usagers à travers un dispositif de communication en ligne qui prépare une mise en scène au sein de la page-écran.

La visée stratégique de la communication en ligne est donc de fournir un support intermédiaire pour manipuler indirectement des archives et des collections d'images à l'aide des fonctionnalités prévues dans une interface. Nous verrons bientôt qu'il faut envisager une « tension sémio-rhétorique » entre ce support et les pratiques d'usage du patrimoine numérique[65]. Auparavant, nous présenterons certaines caractéristiques de la conversion numérique à travers la théorie du support de Bruno Bachimont (2010 ; voir aussi Bachimont, 2007a & 2017b).

Partant des acquis de la sémiotique, de l'ingénierie des connaissances et de la philosophie de la technique, Bachimont (2010) défend une théorie du support qui prend en compte l'impact de l'informatique sur la manipulation des connaissances. L'auteur examine cette question à travers la notion de *dispositif technique*[66]. L'auteur distingue trois types de dispositifs pour définir la relation entre la *technique* et le *sens* (*ibid*. : 34-41) :

[65] **Chap. 2, sect. 2.3.**

[66] Le « dispositif » peut être compris comme « une organisation matérielle et spatiale capable de produire et déterminer un devenir » (Bachimont, 2010 : 42). Cette notion se base d'ailleurs sur une large réflexion scientifique au sujet de l'évolution des techniques médiatrices et l'influence de ces techniques sur les sociétés humaines ; sur ce sujet, cf. entre autres : Wiener (1948), Leroi-Gourhan (1945, 1964 & 1965), Bateson (1984 [1979]), Séris (1994 [2013]), Stiegler (1994, 1996 & 2001), Mattelart (2018 [2001]), Simondon (2014 & 2015 [2010]).

- L'*outil* est un objet technique dont la structure matérielle conditionne et prescrit le geste à entreprendre en réalisant une fonction que l'humain ne peut réaliser avec ses moyens physiologiques naturels.
- L'*instrument* n'assiste pas directement une action (ou un geste) comme l'outil, il est plutôt un dispositif technique qui élargit notre perception en proposant une médiation de cette perception à travers ce dispositif.
- Le *contenu* est considéré comme un dispositif technique qui manifeste des éléments à interpréter par celui qui l'utilise.

L'essence du dispositif est de déterminer un comportement temporel par la configuration d'un dispositif matériel, et par là même, de réaliser une action. Dans chaque cas, le dispositif est caractérisé par des actions à entreprendre : l'outil vise une *action physique*, l'instrument concerne une *prise d'information*, et le contenu est lié à la *formulation d'une pensée*. Pour que ces possibilités s'actualisent, il faut nécessairement qu'un utilisateur s'empare de l'objet technique et l'active, autrement dit, il faut que le dispositif s'intègre dans des usages pratiques où il est manipulé et utilisé par des humains.

De plus, la technique est influencée par trois « logiques intrinsèques », chacune renvoyant à une « figure typique » (Bachimont, *ibid.* : 98). La *cohérence interne*, la loi interne du dispositif qui assure son fonctionnement répétable et prédictible. Cette logique renvoie à la « figure du chercheur » (celui qui « vise l'universel et le nécessaire, le répétable et le mesurable »). La *cohérence concrète*, c'est-à-dire la mise en œuvre effective du dispositif et les ajustements pour réaliser cette mise en œuvre. Cette logique renvoie à la « figure de l'ingénieur » (celui qui « sort de la nécessité du monde idéal du chercheur pour se confronter à la contingence du réel »). Enfin, la *cohérence externe*, à savoir : la relation du dispositif avec l'univers des valeurs (économiques, juridiques, sémiotiques, sociales, culturelles). Cette logique renvoie à la « figure du rhéteur » (celui qui argumente et persuade, celui qui cherche l'accord pour agir).

En parallèle de ce questionnement sur la relation entre technique et sens, un point central de l'approche théorique de Bachimont est sans doute un regain d'intérêt envers la notion d'écriture[67]. Ce questionnement sur l'écriture est également un enjeu important

[67] Sur l'écriture voir notamment l'ouvrage bien connu de Jack Goody (1979 [1977]), ou, dans une perspective sémiotique, ceux de Roy Harris (1994) ou encore de Jean-Marie Klinkenberg (2018). Nous renvoyons en outre aux travaux sur les objets d'écriture de Fontanille (2005) et à la contribution de Klinkenberg & Polis (2018) qui propose un éclairage sur la « scripturologie ».

pour des travaux en sciences de l'information et de la communication, surtout lorsque ces travaux cherchent à mettre en évidence la complexité des inscriptions dans les pratiques qui apparaissent avec l'émergence des « nouveaux médias »[68].

Bachimont insiste sur le fait que le troisième type de dispositif, le contenu, est inscrit sur un « support d'enregistrement » qui immerge l'utilisateur de ce dispositif « dans une rationalité donnée, dans une manière de penser et de concevoir le monde » en suivant une *logique*[69] qui « amène la pensée qui l'interprète à dégager certains concepts, à élaborer des structures cognitives, à constituer une rationalité propre à ce type de contenu » (cf. *ibid.* : 38-39). Selon l'auteur, ce processus d'inscription est fondamental dans l'élaboration d'une théorie de la connaissance, car une connaissance est « la capacité d'exercer une action pour atteindre un but », surtout la capacité de répéter cette action (*ibid.* : 110). Cette capacité d'exercer une action et de la répéter est toujours en relation avec (1) un « objet de connaissance » et (2) avec le support d'enregistrement qui est sa condition de possibilité, de telle sorte que « la connaissance n'est pas un objet, une chose, mais seulement une dynamique, un processus interprétatif » (*ibid.* : 106-107).

En effet, la thèse centrale de la théorie du support de Bachimont est que les propriétés du substrat matériel d'inscription et le format physique de l'inscription « conditionnent l'intelligibilité de l'inscription » (*ibid.* : 122). En tant que dispositif technique, les contenus entraînent une rationalité particulière qui permet d'interpréter des contenus à travers les modalités de manipulation propres à un support. Chaque type de support d'inscription favorise en quelque sorte une rationalité appropriée en fonction de la structure matérielle et des possibilités d'interprétation des contenus. Cette rationalité dépend aussi d'habitudes sociales et culturelles qui stabilisent les significations d'une inscription par rapport à « un sens possible selon un point de vue normé » (*ibid.* : 126). Cependant, Bachimont affirme que le sens qu'on peut attribuer à un dispositif n'apparaît

[68] Voir les contributions et surtout les conclusions dans Tardy & Jeanneret (2007). Les médias informatisés sont compris comme des *machines à suggérer*, c'est-à-dire des machines qui gèrent et propagent d'anciennes formes écrites qui peuvent être saisies, mais des machines qui favorisent aussi « la création constante de nouvelles formes écrites qui suggèrent des possibilités d'expression et de pratiques » (*ibid.* : 214). Ces médias « s'intègrent à un processus de communication pensé sur le mode l'implication et non du contrat » (*ibid.* : 214-215), de telle sorte que « les écrits d'écran et de réseau […] ne se contentent pas d'exposer cette implication, ils la reproduisent en permanence comme un substrat de nos pratiques. En somme, *l'écrit informatique est une forme de texte dont le support n'est pas seulement un objet matériel mais toute une configuration possible d'écritures* » (*ibid.* : 215 ; nous soulignons).

[69] Pour Bachimont, « le contenu suit une logique graphique, écrite ou autre » (Bachimont, *op. cit.* : 38).

pas toujours clairement et distinctement. À ce propos, l'auteur distingue ce qui relève d'un *savoir-faire* et ce qui concerne un *savoir-penser* (cf. *ibid.* : 116 & 118).

- Le « savoir-faire » concerne des « instruments » qui donnent à percevoir par le moyen de la programmation d'un geste technique.
- Le « savoir-penser » concerne des « contenus » qui programment la pensée.

En suivant les termes de l'auteur, les *instruments* sont définis comme des *inscriptions instrumentales* et les *contenus* comme des *inscriptions sémiotiques.* Les instruments sont considérés comme des dispositifs prescrivant « une action dans le monde des choses », tandis que les contenus prescrivent quant à eux « des actions dans le monde des représentations » (*ibid.* : 124).

Nous avons vu que Fontanille (2005 & 2008) propose d'interpréter la manifestation du sens à partir du parcours génératif du plan de l'expression. Les textes-énoncés sont compris comme des dispositifs d'inscription qui sont matérialisés à travers des objets-supports et manipulés dans le cadre de scènes pratiques[70]. Dans tous les cas, la dimension matérielle du support est fondamentale pour que les contenus soient manipulables. C'est pourquoi Bachimont explique justement que le *numérique* bouleverse les propriétés de l'inscription et de manipulation des contenus. En modifiant ces propriétés, le support numérique modifie par conséquent les conditions d'interprétation des connaissances (cf. Bachimont, 2007 : 243-245). En effet, lorsque les contenus inscrits sur des supports sont transformés en « contenus numériques », la manipulation (le geste d'appropriation du dispositif) est prolongée par des « systèmes techniques où ni le corps ni la main n'interviennent directement » (Bachimont, 2010 : 149).

Nous touchons ici un point essentiel. Comme nous l'avons déjà suggéré[71], l'expérience des textes écrits informatiques ne repose pas sur les mêmes modalités d'inscription et de manipulation que dans le cas de la rationalité graphique, parce qu'un support intermédiaire remplace le corps et ses actions sur les contenus. En revanche, le numérique augmente considérablement les possibilités de manipulations des contenus parce que le support numérique permet de transformer à plusieurs reprises des contenus. Ce principe de transformation des supports est au fondement du numérique.

[70] Cf. **supra, sect. 1.3.3.**

[71] Voir les discussions sur le support imperceptible et la page-écran ; cf. **supra, sect. 1.3.3.**

Le numérique se base, d'une part, sur un système de données numériques déterminées et dénombrables, et d'autre part, sur un ensemble de règles de fragmentation et de combinaison de ces données. Par contre, ces données numériques ne sont pas directement manipulées par l'utilisateur, car elles sont uniquement manipulées par le système informatique, le support imperceptible. Bachimont souligne ainsi le caractère ambivalent du numérique :

> Le numérique devient désormais le principe technique des différentes écritures que nous permettent de façonner et exprimer nos contenus. En jouant sur le fait que le code numérique est à la fois une commande physique pour un système et une expression symbolique pour une interprétation, *le numérique modifie profondément l'économie de nos supports, dispositifs de pensée et d'action* (*ibid.* : 162 ; nous soulignons).

Si le numérique bouleverse notre manière de penser en permettant d'inscrire, de manipuler et d'agir sur des connaissances, il faut « savoir si ces modifications remettent en cause les structures cognitives et les catégories conceptuelles à travers lesquelles nous pensons le monde et nous rapportons à lui » (*ibid.* : 162). Selon l'auteur, « le changement du statut de la manipulation entre le graphique et le computationnel […] amène à s'interroger sur leurs conséquences cognitives éventuellement distinctes » (*ibid.* : 148). Plus le numérique sera employé dans diverses pratiques en lien avec des connaissances, plus ces pratiques seront confrontées à un problème : la raison computationnelle affecte notre façon de percevoir, de manipuler et de comprendre des connaissances. Dans le cas des « objets d'écriture », la manipulation porte sur l'expérience des dispositifs d'inscription qu'il s'agit d'interpréter. Dans le cas des « écrits numériques », une partie de la manipulation est réalisée par le système informatique qui traite uniquement des données numériques, tandis que l'autre partie est simulée à travers une page-écran.

S'il est évident, comme l'affirme Bachimont, que le numérique *modifie profondément l'économie de nos supports, dispositifs de pensée et d'action* en lien avec diverses connaissances, nous nous concentrerons seulement sur les modifications qui touchent les institutions patrimoniales qui gèrent des archives et des collections d'images.

Dans *Patrimoine et numérique*, Bachimont (2017b : 18-50, 76-77 & 81-104) a proposé un travail important sur la question du patrimoine numérique. Ce travail offre un panorama théorique sur plusieurs problématiques au sujet du passage des « archives » et de la « documentation » au « patrimoine numérique » :

- La question de l'organisation des institutions patrimoniales.
- La question de l'expérience des usagers.
- La question de l'importance des considérations technologiques qu'il s'agit désormais d'intégrer dans les pratiques d'interprétation des archives.

Pour passer de l'écriture au numérique, il faut ajouter à la rationalité graphique « une manipulation automatique, aveugle, mais dont la systématicité apporte une ouverture au sens » (*ibid.* : 65). Puisque la manipulation des contenus numériques « est désormais fondée sur le calcul au lieu d'avoir un rapport à la culture reposant sur la médiation de la langue et la compréhension de l'humain » (Bachimont, 2017a : 23), nous avons suggéré qu'il faut également concevoir un « dispositif intermédiaire » entre l'humain et le système informatique.

Par rapport à la théorie du support que nous avons présentée (cf. **supra**), Bachimont remarque en outre une autre caractéristique générale des dispositifs d'inscription : la *donnée*. Celle-ci est définie par l'auteur comme « la plus petite partie d'information pourvue d'une signification » ; cette signification est uniquement observable dans des « inscriptions formatées logiquement, selon une correspondance formelle entre forme et sens, entre syntaxe et sémantique » (*ibid.* : 33-34 & 237). Nous pourrions alors penser que nous avons toujours « affaire à des signes » comme ceux des langues naturelles[72], et donc que leurs significations dépendent d'un « contexte sémiotique, où les signes font sens à travers l'usage qu'on en fait » (*ibid.* : 76). Si ce rapprochement entre les *signes linguistiques* et cette notion de « donnée » peut paraître séduisant pour caractériser la signification du support numérique, il comporte un risque de confusion, parce que l'utilisateur n'utilise pas, à proprement parler, des *données numériques* lorsque des *contenus numériques* sont manifestés sur une page-écran. Même si la programmation informatique se base sans aucun doute sur un langage, il ne faut pas considérer les « *bits* »[73] traités par l'ordinateur comme des significations ayant un sens pour un humain.

En effet, l'usager banal d'un système informatique n'est jamais confronté aux données binaires ou aux opérations de calcul de ce système ; à la limite, les informaticiens

[72] **Chap. 1, sect. 1.3.1.1.**

[73] Le *binary digit* est une unité de mesure qui représente une séquence d'information, et qui peut ne prendre que deux valeurs : 0 ou 1.

sont les seuls usagers qui observent ce type de données lorsqu'ils utilisent des langages de programmation. Le but de cette dissimulation des données numériques est simple : un ordinateur ne manifeste pas les données qu'il traite, car ces chiffres n'auraient aucun sens pour un humain ; l'ordinateur *est* le système qui réalise le traitement de l'information, tandis que l'écran est le dispositif périphérique où sont manifestés les contenus numériques. Du côté « imperceptible », les contenus numériques sont conditionnés par une dimension matérielle et un format de codage[74] et, du côté « perceptible », ils sont uniquement manifestés par « une forme sémiotique donnée qui introduit sa cohérence et ses contraintes interprétatives » (*ibid.*).

Comme les autres types de supports d'inscriptions de connaissances, le support numérique est donc caractérisé par un « axe matériel » et un « axe interprétatif ». Le support numérique s'approprie les caractéristiques expressives du régime graphique *pour manifester des contenus sur la page-écran*, en articulant toujours des formes sémiotiques à interpréter avec un objet matériel à manipuler, cela même si cette articulation est réalisée dans un contexte où la manipulation des données numériques est « commandée et automatisée » (Bachimont, 2017b : 66). Selon l'auteur, c'est pourquoi il faut distinguer deux types d'unités :

- Les « unités techniques de manipulation », c'est-à-dire les unités manipulables en fonction du codage adapté à un support d'enregistrement (graphique ou numérique). Ce « codage » fixe le cadre préalable pour réaliser un contenu.
- Les « unités sémiotiques d'interprétation » que l'usager « reconnaît et dégage dans un contenu via son parcours interprétatif ». Ces secondes unités ne sont pas la condition, mais plutôt le résultat d'un processus d'interprétation (cf. Bachimont, 2017b : 18 *sq.* & 245-246).

Nous pouvons pousser plus loin ce rapprochement entre sémiotique et numérique. Dans *Semiotic of Programming*, Kumiko Tanaka-Ishii (2010) affirme que l'application de la sémiotique à la programmation informatique peut aider à situer la technologie informatique « dans un cadre de réflexion propre aux sciences culturelles », parce qu'on retrouve des problèmes généraux propres aux systèmes de signes dans les langages de programmation (*ibid.* : 5 ; nous traduisons de l'anglais). En s'appuyant sur la théorie

[74] En informatique, le « format » est l'agencement, la disposition, la manière de représenter un type de données sur un support d'inscription.

saussurienne du signe (modèle dyadique), mais aussi sur la théorie peircienne (modèle triadique), Tanaka-Ishii (2010 : 180) a rapproché les langages de programmation de la langue, en tant que « système de signes »[75], à travers le principe de réflexivité (cf. *ibid.* : 193).

En remarquant que les signes fonctionnent comme un *médium* qui peut acquérir des significations hétérogènes, l'auteure affirme qu'un « système linguistique réflexif » est un processus qui utilise une expression en entrée (*input*) et qui en produit une autre en sortie (*output*). Dans cette perspective, l'auteure envisage qu'un système informatique composé de signes « peut naturellement aboutir sur une réflexivité, et que la stratégie pour manipuler la réflexivité détermine le système de signes en devenir » (*ibid.* : 196). En effet, dans les systèmes linguistiques comme dans la programmation informatique, certaines parties des résultats du processus réflexif forment des expressions interprétables par le système lui-même. L'auteur parle alors d'une réflexivité par « auto-augmentation » (ou *self-augmentation* ; cf. Tanaka-Ishii, 2010 : 180). En tant que systèmes complexes de signes, les langages informatiques seraient donc, comme les langues naturelles, intrinsèquement réflexifs, bien que nous ne sachions pas encore « comment exploiter la réflexivité des systèmes informatiques dans une direction spécifique » (cf. Tanaka-Ishii, *ibid*. : 189-191). Pour Tanaka-Ishii, cette auto-augmentation ne manque pas d'un cadre technique ou théorique, mais le défi pour les ingénieurs en informatique concerne plutôt la recherche du « bon équilibre » entre les possibilités réflexives et les usages de l'informatique.

Andrea Valle et Alessandro Mazzei (2017) ont également proposé d'étudier les langages de programmation, mais afin de révéler « comment l'identité formelle sous-jacente de tous les langages de programmation laisse la place à une variété énorme et "sauvage" de réalisations en acte » (*ibid.* : 507-509). Puisque la réalisation du langage nécessite un acte d'énonciation, les auteurs envisagent les langages de programmation comme des *systèmes sémiotiques* comprenant différents processus d'énonciation. L'énonciation qui est mise en œuvre dans les langages de programmation peut par exemple être considérée selon un axe *interne*, celui du *système* concernant un « niveau

[75] Tanaka-Ishii utilise dans son ouvrage la notion de signe, probablement parce que l'auteure n'a pas pris acte des travaux plus récents en sémiotique ; cf. **supra, sect. 1.3.1.2. à 1.3.2.** D'autre part, la théorie du signe de Peirce est également pertinente pour étudier les langages de programmation, car cette théorie s'appuie notamment sur la *logique mathématique* (cf. Marty, 1990).

structurel » (le *langage*, dans la théorie de Saussure) ou selon un axe *externe*, celui du *procès*[76], où la subjectivité peut être articulée en relation avec l'usage d'un langage compris comme système de signes qu'il s'agit de s'approprier (la *parole* dans cette même théorie).

Néanmoins, il ne faut pas confondre la réflexivité des langages de programmation et la réflexivité qui s'appuie sur des langues naturelles. En effet, la réflexivité (computationnelle) du système informatique, celle qui concerne les langages de programmation et les données numériques, ne concerne pas le même niveau que les usages pratiques, éventuellement réflexifs, que le numérique rend possibles. Cependant, nous pouvons dire qu'il s'agit bien de deux types de langage, chacun visant avant tout à réaliser des applications et des usages pratiques.

À ce stade, l'approche théorique du support de Bachimont met en évidence au moins trois points importants qui sont complémentaires des propositions de Fontanille sur le support des textes informatiques :

- Chaque type d'inscription sémiotique (ou contenu) entraîne une rationalité particulière permettant de manipuler et d'interpréter ces inscriptions.
- Lorsqu'ils sont transformés en données numériques, les contenus sont manipulés par le moyen d'un système technique qui se fonde sur le calcul et l'automatisme.
- Du point de vue de l'utilisateur, les contenus numériques sont toujours manifestés par des formes sémiotiques issues d'autres régimes médiatiques, en particulier le régime graphique.

Lorsque des documents sont convertis en ressources numériques, cette conversion numérique ne donne pas accès à la matérialité des documents originaux, car le numérique se base en principe sur des opérations de calcul portant sur des nombres et non pas sur des objets matériels. Pourtant, même si le numérique laisse de côté cette matérialité, les pratiques numériques nous engagent dans un processus d'*interprétation* par le moyen d'un *dispositif intermédiaire*. En effet, la matérialité des ressources numériques est remplacée par l'expérience d'un autre objet-support, à savoir : les composants

[76] Dans la terminologie de Hjemslev, le « procès sémiotique » correspond à l'axe syntagmatique du langage, en opposition au « système sémiotique » qui en représente l'axe paradigmatique ; cf. **supra, sect. 1.3.1.1.**

électroniques et l'écran d'ordinateur d'une part, et la page-écran de l'autre. L'une des particularités de la page-écran est qu'elle peut combiner plusieurs types d'unités sémiotiques d'interprétation (icônes, listes, textes écrits, images, vidéos, etc.) dans une même scène d'énonciation.

Dans la section suivante, nous présentons d'autres travaux en sémiotique qui prennent en compte cette dimension pragmatique de l'énonciation, mais d'une autre manière. Les sections précédentes ont éclairé le volet *sémiotique* de notre approche sémio-rhétorique. Cette section introduit le volet *rhétorique* de cette approche.

1.5. Rhétorique et pragmatique de la communication en ligne

Dans la continuité de la sémiotique poststructuraliste, le Groupe μ est connu pour avoir renouvelé le cadre de réflexion sur la rhétorique (Groupe μ, 1982 [1970])[77]. Dans son *Précis de sémiotique générale*, Jean-Marie Klinkenberg (membre du Groupe μ) rappelle que la *nouvelle rhétorique* concerne toujours la question de l'argumentation, en tant que technique de persuasion. En insistant sur la place des « techniques médiatrices » utilisées dans les processus d'argumentation, cette « nouvelle rhétorique » cherche à « décrire les énoncés, quelle qu'en soit la qualité » (*ibid.* : 342-343).

Cette nouvelle rhétorique s'inscrit dans une perspective herméneutique qui a pour but d'analyser « *a posteriori* les faits de parole et de discours et dégager les règles générales de leur production » (Klinkenberg, 1996 : 342). Pour Klinkenberg, l'hypothèse d'une *rhétorique générale* est en effet compatible avec le projet sémiotique, car « s'il y a une rhétorique linguistique et si par ailleurs existent des lois générales de la communication et de la signification […] alors il doit exister une rhétorique générale dont les règles peuvent rendre compte des phénomènes rhétoriques se manifestant dans divers champs sémiotiques » (*ibid.* : 376)[78].

---

[77] La *nouvelle rhétorique* du Groupe μ s'est développée en parallèle des travaux sur la rhétorique et l'argumentation de Chaïm Perelman (cf. Perelman & Olbrechts-Tyteca, 2008 [1958]). De plus, les travaux du Groupe μ (qui étaient d'abord fermement ancrés dans le structuralisme ; cf. notamment Ricœur, 1975 : 173-219) se sont progressivement dirigés vers le pragmatisme et les sciences cognitives (Groupe μ, 2015). Sur les travaux du Groupe μ, voir les contributions dans Badir & Dondero (2010) ou dans Dondero & Sonesson (2010).

[78] Nous verrons que ce point de vue a par exemple proposé d'étudier la rhétorique du signe visuel ; cf. **chap. 3, sect. 3.2.3.** Nous constaterons aussi que l'énonciation occupe également une place centrale dans les recherches en sémiotique visuelle ; cf. **supra, sect. 1.3.1.2. & 1.3.2.**

Dans le même ouvrage, l'auteur a également beaucoup œuvré pour rapprocher la sémiotique des travaux sur les théories de la communication[79] et sur la pragmatique (*ibid.* : 43-116). Ce rapprochement implique pourtant un changement de focalisation, comme le remarque Eco (1988 [1984] : 28-29) : « Opposer une pragmatique des énoncés à une sémantique des unités du signe, c'est déplacer l'attention des systèmes de signification sur les processus de communication ». Pour Eco, ce déplacement confirme que « tout système de signification est élaboré pour produire des processus de communication » (*ibid.* : 29 ; sur la relation entre signification et communication, cf. aussi *ibid.* : 70)[80].

Dans une perspective *sémio-pragmatique*, Roger Odin (1990 & 2011) propose quant à lui d'étudier les contextes de production textuelle en cherchant à dépasser les approches sémiotiques qui accordent une place importante au « principe d'immanence »[81] dans l'élaboration de la théorie sémiotique élaborée par Hjelmslev, puis reprise par Greimas.

D'une autre manière que Fontanille, l'approche sémio-pragmatique affirme également que le sémioticien doit prendre en compte « l'expérience d'un travail de production textuelle en contexte » et pas seulement s'intéresser à l'analyse de textes ou de discours (Odin, 2011 : 123). Odin (cf. *ibid.* : 39-41) propose de remplacer la notion de « contexte » par celle d'*espace de communication*, à savoir l'« espace à l'intérieur duquel le faisceau de contraintes pousse les actants […] à produire du sens sur le même axe de pertinence ». En choisissant un « axe de pertinence », le théoricien sélectionne certaines « contraintes qui régissent le processus de production du sens », (*ibid.* : 40-41). Ce théoricien peut « donner à l'espace de communication sur lequel il vaut travailler le degré et le genre exact de généralité ou de particularité (l'extension) qu'il désire en termes

[79] Sur les théories de la communication, cf. Mattelart & Mattelart (2018 [1995]). Pour un autre point de vue sur les relations entre communication et langage, cf. par exemple Moles & Vallancien (1963) ou, dans une perspective anthropologique, Winkin (1981).

[80] Le rapprochement entre sémiotique et sciences de l'information et de la communication est toutefois une question délicate (cf. par exemple Ablali, 2007, ou encore Carbou, 2013). Sur le rapport entre sémiotique et sciences de l'information et de la communication, voir aussi Boutaud (2004 & 2011), Jeanneret (2007), Boutaud & Berthelot-Guiet (2013).

[81] À propos de ce principe d'immanence, cf. notamment Rastier (1994 : 325-335). Pour un autre point de vue, voir aussi les propositions de Fontanille (2008 : 9-16).

d'objets, d'espace et de temps » (*ibid.* : 41), puisque cet espace est d'abord « une construction effectuée par le théoricien » en vue d'une analyse sémiotique.

Pour l'auteur, ce sont en effet les « expériences de lecture » qui nous engage « dans un mouvement migratoire d'espaces de communication en espaces de communication », même si ces expériences de lectures et ces espaces impliquent parfois « un changement de mode de production de sens et d'affects et plus largement un changement de rôle » (*ibid.* : 140). Plus précisément, les constructions théoriques qui visent à structurer les processus de production de sens en ensembles fonctionnels sont appelées des *modes* ; la construction des modes répond, selon l'auteur, à trois grandes questions : (1) quelle(s) mise(s) en forme discursive(s) ce mode met-il en œuvre ? (2) Quelles relations affectives produit-il ? (3) Quelle relation énonciative, quel(s) énonciateurs(s) invite-t-il à construire ?

Nous avons vu avec Fontanille que le sens produit est manifesté au sein de scènes pratiques qui sont encadrées par des stratégies dont le but est d'aménager l'environnement complexe où les usagers prennent place et où différentes scènes d'énonciation sont en concurrence. Dans son approche sémio-pragmatique, Odin distingue au moins quatre composantes de ce qu'il nomme la « compétence communicationnelle » (cf. *ibid.* : 43-44) :

- La compétence discursive, correspondant au réservoir de « processus discursifs » dans lequel les acteurs de la communication puisent pour produire du sens.
- La compétence sémiolinguistique qui « concerne la maîtrise des langages » et met en œuvre certains « processus en fonction du langage utilisé ».
- La compétence socioculturelle concernant « la dimension interactionnelle et sociale de la communication » et qui « intervient dans la relation entre l'espace de communication et les acteurs de cet espace » et « dans les relations entre les acteurs de cet espace ».
- La compétence référentielle « qui règle la relation aux domaines d'expériences et aux objets du monde, et donc directement le *sens produit* ».

Dans la suite du présent travail, nous verrons que la conversion numérique des archives et des collections d'images aboutit sur la communication en ligne d'importants stocks d'images[82]. Ensuite, dans la **troisième partie** et la **quatrième partie**, nous

[82] **Chap. 2.**

analyserons plusieurs *stratégies de communication en ligne* qui débouchent sur la conception d'*interfaces Web* qui transmettent ces stocks d'images. L'expérience de ces interfaces présuppose des composantes relatives à cette *compétence communicationnelle*. Nous verrons comment ces stratégies construisent un point de vue[83] pour l'utilisateur à l'aide de l'agencement visuel de cette interface.

Nous avons suggéré que cet utilisateur est uniquement confronté à un support perceptible, la page-écran. D'autres travaux mettent en évidence l'énonciation et les opérations de lecture manifestées par cette partie visible de la *médiation numérique* et, par là même, son potentiel rhétorique.

Dans *Rhétorique du texte numérique*, Alexandra Saemmer (2015 : 63) affirme que le texte numérique implique une « pratique sociale de *co*construction du sens par les concepteurs et les lecteurs », et surtout, que des *figures de la lecture* (les potentiels d'action du texte numérique) préfigurent justement une place pour ces « lecteurs ».

> La rhétorique du texte numérique considère le processus de lecture comme une rencontre, toujours forcément partielle, entre l'anticipation des pratiques par les procédés rhétoriques, et les attentes individuelles et socialement partagées du lecteur. Il s'agit donc, d'une part, de prendre en compte les imaginaires du dispositif et de la lecture numériques, des médias et des genres textuels […]. D'autre part, il s'agit d'analyser comment le texte numérique à lire, à regarder et à manipuler, cadré dans une page-écran et ancré dans son dispositif, anticipe sur les pratiques de lecture et les constitue. Ces modalisations de pratiques par le texte numérique sont appelées figures de la lecture (Saemmer, 2015 : 243).

Ces figures de la lecture (en tant modélisations des pratiques textuelles dans un environnement informatique), ou encore l'orchestration de la page-écran (en tant qu'objet d'écriture) évoquée par Pignier, indiquent sans aucun doute qu'une nouvelle forme d'énonciation/interprétation est en jeu dans les pratiques d'écriture numérique, même si nous trouvons, dans ces pratiques, des formes sémiotiques relevant d'autres médias, d'autres pratiques que l'écriture verbale. La question est de savoir comment la communication en ligne *orchestre* cette énonciation multimédia. Quelles formes sémiotiques et discursives sont mises en œuvre dans la communication ligne d'archives

[83] **Chap. 3, sect. 3.2.3.**

et de collections d'images ? Qui est l'énonciateur ? Qui est le destinataire ? La relation énonciative est-elle présentée dans l'interface ou alors est-elle dissimulée ?

Avant d'essayer de répondre à ces questions[84], nous remarquerons que la conversion numérique change radicalement les conditions de visibilité et de manipulation des images en modifiant « profondément l'économie de nos supports, dispositifs de pensée et d'action ». Dans *Image et vérité*, Jean-François Bordron (2013 : 83) remarque que l'économie désigne d'abord « l'ordonnancement qui fonde la possibilité des valeurs et leur éventuelle circulation », ou encore « un jeu d'instances, présupposant une opération de partage qui les institue dans leurs différences et par là dans leurs relations possibles ». Pour l'auteur, comprendre « l'économie d'une image », c'est d'abord « se demander *dans quel ordre global elle s'inscrit*, quelle articulation fondamentale est présupposée pour qu'on puisse la comprendre » (*ibid.* : 82). Les articulations qui rendent possible une « économie du visuel » sont cependant liées à l'une des principales caractéristiques de l'expérience d'images. Selon Bordron, cette expérience d'images ouvre plusieurs scènes qu'il s'agit de comprendre par rapport à une « économie iconique dans laquelle elles doivent être perçues » ; économie qui « fonde un style de visibilité et plus généralement, un style de perception » (*ibid.* : 55)[85]. Plusieurs économies des images sont évidemment possibles, chacune permettant de mettre en évidence le rôle attribué au visuel dans une culture donnée.

De plus, Bordron remarque que la *technologie* est actuellement un des éléments essentiels de l'*épaisseur rhétorique de l'image* (*ibid.* : 172), comme il affirme que les relations entre la technique et la rhétorique ne doivent pas nous amener à sous-estimer l'importance de l'une comme de l'autre : « Au contraire, la technologie est devenue la pièce essentielle de l'art de persuader et d'argumenter dans le registre de la connaissance » (*ibid.*)[86].

[84] Cf. nos analyses dans la **troisième partie** et la **quatrième partie**. Notons déjà que notre recherche ne se limite pas à la communication en ligne d'images. En effet, de plus en plus de travaux suggèrent que le traitement informatique peut être utilisé pour analyser, classer et visualiser des corpus d'images.

[85] Voir aussi les discussions dans le **chap. 3, sect. 3.2.3.**

[86] Sur les relations entre image et connaissance, voir aussi les contributions dans Dondero (2019a).

Somme toute, les travaux que nous avons présentés dans les sections précédentes font ressortir trois constatations importantes pour la suite de notre propos.

(1) Les pratiques d'archivages visent à sauvegarder les connaissances inscrites sur différents types de supports. Pour Bachimont, la rencontre de la technique et du sens est caractérisée par deux axes complémentaires. D'une part, la *matérialisation*, parce que les connaissances doivent être physiquement structurées et implémentées dans des dispositifs matériels. D'autre part, l'*interprétation*, puisque ces supports et ces connaissances se fondent sur une rationalité (par exemple, graphique, audiovisuelle ou numérique) et s'intègrent dans des pratiques d'usage, où ils sont manifestés à travers des formes sémiotiques qui sont interprétables par l'usager.

(2) Le numérique modifie profondément les conditions de manipulation des supports et des contenus parce que ceux-ci sont convertis en *nombres*. En plus d'être un support, le numérique est un *principe de transformation des contenus* entraînant les connaissances dans un nouveau type de rationalité qui se base sur le *calcul* et l'*automatisme*. Dans la **deuxième partie**, nous remarquerons en effet que des travaux en histoire de l'art[87] et en sémiotique visuelle[88] s'intéressent aux possibilités d'analyse qu'apporte le numérique, surtout depuis que des méthodes assistées par des ordinateurs sont utilisées pour analyser des larges ensembles d'images et pour mettre en évidence des relations inédites entre ces images[89].

(3) Des langages de programmation (le support imperceptible) et une page-écran (le support perceptible) sont utilisés pour mettre en scène un dispositif de communication en ligne qui implique une *dimension stratégique* et des *compétences sémio-pragmatiques*. La dimension stratégique vise, d'une part, à aménager un objet-support comme une scène d'énonciation qui offre un *point de vue* à l'utilisateur et qui lui permet d'interagir avec le système informatique qui traite automatiquement des contenus numériques. D'autre part, la dimension stratégique orchestre cette scène et cette interaction à l'aide de l'épaisseur

[87] **Chap. 3, sect. 3.1.4.** En s'inspirant d'autres approches en histoire de l'art, l'idée est notamment de favoriser la comparaison d'une grande quantité d'images, afin de chercher à comprendre, dans la production d'images d'une certaine période, les continuités et les ruptures dans l'usage de certaines compositions visuelles ; cf. les discussions dans le **chap. 3, sect. 3.1.2. & 3.4.**

[88] **Chap. 3, sect. 3.2.4.**

[89] Cf. les discussions sur les techniques de l'*image processing*, de la *data visualization*, de la *computer vision* et du *deep learning* dans le **chap. 4**, ainsi que la présentation de certaines applications possibles de ces techniques dans les **chap. 8 & 9**.

rhétorique de la page-écran et des figures de la lecture qui sont programmées à l'avance. Ainsi, nous proposons d'envisager cette orchestration comme un *dispositif rhétorique* qui a pour but de construire une relation d'énonciation à travers la mise en scène d'une interface Web.

1.6. Conclusions

Dans ce **premier chapitre**, nous avons vu que les archives d'images rassemblent des supports d'images qui sont les preuves du passé. Concernant ces supports d'images, nous avons également montré que différents régimes médiatiques se sont succédé dans l'histoire du visuel et des images. Même si les passages, d'un régime à un autre, peuvent être différents selon le point de vue de tel ou tel auteur, nous avons constaté un même mouvement : le régime graphique (l'écriture) est remplacé par le régime audiovisuel (les images) et, plus récemment, le régime numérique remplace ces régimes graphique et audiovisuel. Nous avons constaté que ces changements de régime débouchent actuellement sur un accroissement considérable des archives et des collections d'images.

Ensuite, nous avons montré que les archives d'images sont organisées à l'aide d'un traitement documentaire. En simplifiant, le traitement matériel concerne la conservation du support d'inscription des images, tandis que le traitement intellectuel concerne la description verbale du contenu et l'organisation administrative de ces images. Par conséquent, l'archivage d'images consiste surtout à classer des images en fonction de mots-clés qui décrivent le plus simplement possible les images. Cette constatation nous a permis de mettre en évidence une relation entre l'archive et la langue.

Nous avons rapproché la définition foucaldienne de l'archive de considérations sur les discours et sur l'énonciation. Après avoir défini l'archive comme une pratique d'où surgit une multiplicité d'énoncés, nous avons présenté plusieurs travaux en sémiotique qui soulignent que l'analyse des énoncés est indissociable de la dimension pratique des situations de communication où ces énoncés sont manifestés. Selon Fontanille, ces situations se rapportent à un parcours génératif comprenant différents niveaux de pertinence (signes, textes-énoncés, objets-supports, scènes pratiques, stratégies, formes de vie ; ces niveaux forment ensemble un parcours d'expression où différentes dimensions de l'interprétation sont imbriquées les unes dans les autres).

L'énonciation est ainsi un processus qu'il s'agit de considérer à différents niveaux, chaque niveau correspondant à un plan d'expression spécifique à prendre en compte.

C'est pourquoi nous avons vu que le sémioticien ne doit pas seulement prendre en compte la cohérence des signes qui composent le plan de l'expression d'un texte. Il doit aussi considérer les caractéristiques de l'objet-support du texte, c'est-à-dire la relation entre certaines propriétés du support matériel et la formalisation de ce support en vue de la manifestation d'un énoncé. Comme le montre l'approche de Fontanille, ou encore la théorie du support de Bachimont, le support d'un dispositif d'inscription n'est pas constitué de façon aléatoire. Au contraire, ce support est toujours intégré matériellement dans un contexte (ou une scène pratique) et il est conçu en vue d'un usage prévu en fonction des caractéristiques matérielles et formelles de ce support. En effet, selon Bachimont, chaque type de contenu entraîne une rationalité particulière qui dépend des caractéristiques matérielles du support d'inscription. D'un côté, cette matérialité est fondamentale pour que les contenus soient physiquement manipulables. De l'autre, les habitudes sociales et culturelles stabilisent les significations et modalités d'interprétation des inscriptions.

En outre, les travaux sémiotiques que nous avons présentés insistent sur quatre points importants : (1) l'usage du support numérique n'évacue pas la question de l'interprétation des connaissances qui sont transformées en contenus numériques ; (2) la communication en ligne implique la mise en œuvre de stratégies qui aménagent un objet-support comme un lieu de rencontre entre des utilisateurs et un large ensemble d'images ; (3) la différence entre le numérique et d'autres types de supports d'inscription réside surtout dans les opérations de calcul appliquées sur des données numériques et dans les possibilités de manipulation de ces données ; (4) en s'appuyant sur des langages de programmation, l'avantage de la raison computationnelle est qu'elle permet d'augmenter considérablement les possibilités de manipulations des contenus, parce que ceux-ci sont manipulés automatiquement par un ordinateur sous la forme de chiffres.

Le support numérique impose par conséquent un nouveau type de rationalité aux contenus en modifiant profondément l'économie des supports. Nous avons vu que l'utilisation du numérique nécessite la conception d'un dispositif intermédiaire que l'utilisateur mobilise pour observer et manipuler des ressources numériques. C'est également ce qu'indiquent les travaux récents en sémiotique du discours, lorsque ceux-ci remarquent que les textes numériques se basent sur deux objets d'écriture

différents : d'une part, le support imperceptible, les composants électroniques et les langages de programmation ; et d'autre part, le support perceptible, l'écran et la page-écran. Les travaux présentés indiquent que ce dispositif intermédiaire se fonde sur des expériences de lecture qui sont soumises à des contraintes linguistiques, discursives ou socioculturelles. Néanmoins, la lisibilité de ce dispositif dépend d'une dimension stratégique qui aménage l'environnement virtuel où les utilisateurs prennent place.

Dans le **chapitre suivant**, nous présenterons d'autres travaux sur cette dimension stratégique de la communication en ligne des archives d'images. Si ce type de communication vise à préparer une relation d'énonciation par le moyen d'une interface Web, nous verrons cependant que l'utilisateur doit participer activement à la production du sens qui est mis en scène dans cette interface.

CHAPITRE 2
Stratégies de communication en ligne d'images et dématérialisation de la culture visuelle

Le passage du graphique au computationnel marque un tournant dans les conditions de visibilité des larges ensembles d'images. C'est pourquoi l'impact de l'informatique dans les pratiques d'archivage d'images impose de nombreuses réflexions critiques.

Dans ce **deuxième chapitre**, nous indiquerons d'abord qu'un contexte de recherche transdisciplinaire (*digital humanities* ; cf. **sect. 2.1.**) favorise des réflexions innovantes sur la relation entre les sciences humaines et le numérique. Ensuite, nous passerons en revue plusieurs travaux qui mettent en évidence l'impact de l'informatique dans des pratiques de médiation en lien avec la culture visuelle, notamment les pratiques qui ont pour objectif de diffuser des archives et des collections d'images (**sect. 2.2.**). Nous verrons que ces travaux insistent sur l'importance des choix stratégiques des institutions patrimoniales en ce qui concerne la reproduction numérique des larges ensembles d'images (**sect. 2.2.1.** & **2.2.2.**).

Dans un second temps, nous présenterons ensuite une approche sémio-rhétorique du patrimoine numérique qui théorise et met en évidence différentes tensions concernant cette transformation des ressources matérielles en ressources numériques (**sect. 2.3.**). Couplée à d'autres travaux sur la conception des sites Web et des bases de données, ainsi que sur la dimension énonciative des interfaces (**sect. 2.3.1.** & **2.3.2.**), cette approche

sémio-rhétorique indique que les stratégies de communication en ligne du patrimoine numérique s'appuient principalement sur la conception d'un « dispositif de lecture » ; c'est ce dispositif qui permet à l'internaute d'accéder aux ressources numériques et de les utiliser (**sect. 2.3.3.**).

2.1. Humanités numériques : une réaction transdisciplinaire

Les motivations scientifiques qui poussent les sciences humaines à aménager cette transition des archives aux archives numériques se fondent sur un contexte particulier.

À mi-chemin entre les sciences humaines et l'informatique, les travaux en humanités numériques (*digital humanities*) participent à la formation d'un nouveau champ de recherches qui a pour objectif d'étudier l'impact de l'informatique sur les disciplines en sciences humaines[90]. D'une manière générale, les travaux des humanités numériques s'organisent principalement autour de *trois usages* du numérique qu'il s'agit de questionner : le numérique peut être utilisé (1) en tant qu'instrument de recherche, (2) en tant qu'outil de communication ou (3) en tant qu'objet de recherche. Parmi ces recherches, nous verrons qu'il s'agit notamment d'examiner de nouvelles hypothèses épistémologiques et méthodologiques dans le domaine de l'histoire de l'art à la suite du « tournant numérique »[91].

Même si la définition de ce nouveau champ ne fait cependant pas encore l'objet d'un consensus, un ensemble de chercheurs a toutefois proposé en 2011 un « Manifeste des Digital Humanities »[92] qui nous permettra de faire le lien avec les considérations sur le numérique que nous avons déjà présentées[93]. Par exemple, ce manifeste invite à constater que le « tournant numérique pris par la société modifie et interroge les conditions de production et de diffusion des savoirs ». En effet, durant les dernières décennies, des « communautés particulières issues de l'intérêt pour des pratiques, des

[90] Pour plus de détails sur les discussions autour des *Digital Humanities*, voir par exemple Brossaud et Reber (2007a, 2007b), Mounier (2012), Dacos & Mounier (2014) ou encore Hooland *et al.* (2016). En outre, cf. les travaux en anglais : Svensson (2009) ; Berry (2012) ; Gold (2012) ; Burdick *et al.* (2012) ; Liu (2013) ; Vanhoutte, Nyhan & Terras (2014) ; Bode & Arthur (2014).

[91] **Chap. 3, sect. 3.1.4.**

[92] Ce manifeste est publié par Marin Dacos à la suite du *THATCamp Paris 2010*, une « non-conférence » sur les humanités numériques ; voir https://tcp.hypotheses.org/318.

[93] **Chap. 1, sect. 1.4.**

outils ou des objets transversaux divers » convergent et constituent le champ actuel des humanités numériques. Ce champ est *transdisciplinaire* parce qu'il s'appuie « sur l'ensemble des paradigmes, savoir-faire et connaissances propres [aux sciences humaines], tout en mobilisant les outils et les perspectives singulières du champ du numérique » (même s'il ne s'agit pas de faire « table rase du passé »).

Dans cette perspective, l'objectif des humanités numériques est de faire progresser les connaissances à travers « la qualité de la recherche dans [les disciplines des sciences humaines], et l'enrichissement du savoir et du patrimoine collectif » à l'aide de l'informatique. Pourtant, si le numérique offre des avantages, il impose aussi des contraintes techniques, pour les auteurs du manifeste, ces contraintes sont aussi une opportunité qui, dans une perspective collective, devrait pousser les spécialistes en sciences humaines et les informaticiens à améliorer les usages du numérique dans une perspective scientifique. Pour ce faire, il s'agit de favoriser le libre accès aux données et aux métadonnées qui concernent des savoirs ou le patrimoine culturel, mais aussi d'intégrer des formations aux humanités numériques dans les cursus en sciences humaines, ou encore de développer les compétences dans ce domaine à travers des infrastructures informatiques qui répondent à des besoins réels pour les chercheurs.

En résumé, nous pouvons mettre en évidence que :

- Les humanités numériques étudient l'évolution des pratiques en lien avec des outils computationnels utilisés pour diffuser et valoriser des connaissances.
- Ce champ vise à fournir un cadre de réflexion théorique pour comprendre l'utilisation de ces méthodes informatiques.

Par conséquent, il ne faut pas seulement utiliser aveuglément ces outils sans chercher à les comprendre. Pour les humanités numériques, il s'agit surtout de constituer une communauté scientifique qui interroge l'impact du numérique dans diverses disciplines en sciences humaines (sciences historiques ; littérature et linguistique ; sociologie ; histoire et sciences de l'art ; science de l'antiquité ; etc.) et dans différents secteurs d'activité (archivistique ; documentation ; bibliothèque ; musées) en lien avec ces disciplines.

Comme nous l'avons déjà souligné[94], les archives ou les collections photographiques ont été utilisées afin de reproduire des œuvres d'art ou des monuments. Cette importante quantité d'images constitue évidemment des ressources documentaires qui sont conservées parce qu'elles sont utiles pour étudier et pour mieux comprendre le passé. Dans le même sens, la conversion numérique de ces photographies aboutit sur des larges en ensembles d'images numériques qui sont intégrées dans un processus de communication en ligne où l'expérience de l'usager doit être programmée et anticipée. Cependant, s'il est évident que la numérisation intensive des archives et des collections d'images a favorisé l'émergence de nombreux stocks de substituts numériques sur le patrimoine culturel, à ce stade, il faut apporter une distinction importante au sujet de la documentation numérique.

En effet, un document numérique est une « unité documentaire, nativement numérique ou numérisée, qui a été créée ou numérisée par la bibliothèque ou acquise sous forme numérique comme élément de la collection de cette bibliothèque »[95]. D'une part, il existe des ressources documentaires matérielles qui ont été numérisées ; chaque unité documentaire concerne une information enregistrée ou un objet matériel qui peuvent avoir des formes et des caractéristiques différentes. D'autre part, d'autres ressources sont à l'origine des ressources numériques, dans le sens où elles ont été générées[96] à l'aide d'un système informatique (p. ex. : fichier informatique en format texte ; photographie numérique ; image de synthèse). Les archives numériques comprennent donc des ressources *numérisées* et des ressources *nativement numériques* ; cette différence est importante en ce qui concerne la production des *métadonnées*[97]. Dans le premier cas, toutes les informations relatives aux documents originaux doivent être encodées par un

[94] **Chap. 1, sect. 1.1.**

[95] Voir la norme ISO 11620:240 ; cf. https://www.iso.org/obp/ui/#iso:std:iso:11620:ed-3:v1:fr.

[96] Nous n'aborderons pas la question de la génération informatique d'image ; sur ce sujet, voir toutefois les exemples qui sont présentés dans le **chap. 9, sect. 9.2.**

[97] Nous avons vu que les pratiques d'archivage d'images visent à traiter *physiquement* et *intellectuellement* des documents. D'un côté, le but du traitement matériel est d'anticiper et de prendre en charge la dégradation propre aux supports originaux, c'est-à-dire qu'il s'agit de mettre en place un lieu et des conditions de stockage appropriés pour conserver différents types d'images. De l'autre, le traitement intellectuel a pour but de rendre compte des contenus d'une archive et de classer ces contenus. Nous avons suggéré que l'indexation est le processus qui permet de faire correspondre à des images d'archive un ensemble de notions et de concepts (thésaurus), ainsi que différentes informations sur les images (métadonnées) ; voir aussi **chap. 3**, **sect. 3.3.**

humain dans un système informatique. Dans le second, certaines informations sont éventuellement produites par un système informatique lorsque le fichier numérique est créé ou modifié, tandis que d'autres informations doivent également être encodées.

2.2. Archives d'images numériques et musée virtuel

Divers travaux ont déjà approfondi cette réflexion sur les usages de la technologie numérique dans les pratiques d'archivage d'images. Nous présenterons rapidement quelques prises de position sur ce sujet afin de faire ressortir deux points importants : (1) l'apparition d'un nouveau dispositif de médiation dans l'étude de la culture visuelle ; (2) la place centrale qu'occupent les principes de reproduction visuelle et d'interactivité (entre l'utilisateur et l'ordinateur) dans cette médiation.

En effet, plusieurs travaux sur les archives et collections d'images constatent que de plus en plus d'images sont intégrées dans un environnement et des réseaux virtuels de communication. Dans *Visual Culture*, Richard Howells (2003) remarque la place importante de la communication en ligne dans la *culture visuelle*[98]. Pour l'auteur, le *design*, c'est-à-dire la conception des sites Web insiste sur les possibilités d'interactivité entre l'homme et le système informatique, parce que c'est un « spectateur-opérateur » qui « navigue à travers le site pour en révéler de nouveaux aspects et dimensions » (*ibid.* : 232). Ainsi, le site Web est une innovation technique qui permet de transmettre des images par de nouveaux moyens de communication ; en outre, cette transmission est aussi comprise comme une *nouvelle forme* qui est développée pour rendre visible (spectateur) et pour agir (opérateur) sur la culture visuelle (*ibid.*). Howells (*ibid.* : 232-233) considère cette « forme » selon trois caractéristiques : (1) l'intégration, car les sites Web sont présentés comme des sortes de « vestibules d'accueil » à travers lesquels on initie une navigation, un parcours ; (2) l'interaction, étant donné que les choix de navigation affectent l'apparence (visuelle) du site Web ; (3) l'impermanence, parce que le site n'a jamais d'état fixe ou stable et qu'il peut disparaître à tout moment, ou encore parce que le navigant peut voir des événements changer fréquemment et « en direct », même s'il n'y a pas d'action en direct sur le site.

[98] Sur la culture visuelle, cf. aussi **chap. 3, sect. 3.1.2.**

Cette nouvelle forme est, comme nous l'avons vu[99], le résultat visible d'un dispositif complexe qui comprend, d'une part, un support imperceptible, et d'autre part un support perceptible. Ce dispositif s'intègre dans le projet, plus général, de transformation des connaissances en ressources numériques. Dans l'un des *Cours de médiologie générale* (celui sur « la dynamique du support »), Régis Debray (1991) suggère que le dessein de cette conversion numérique est notamment de *dématérialiser* les documents et les contenus qui sont liés à ces documents ; cette dématérialisation est le résultat d'un mouvement qui précède de loin le numérique.

> La tendance générale, la plus ostensible, de l'évolution des supports de traces est à la dématérialisation. Elle domine et conditionne toutes les autres. Elle traduit et accompagne logiquement le mouvement général de désincarnation qui, par révolutions industrielles successives, abouti […] à la langue sans mots (les bits) […] aux bibliothèques sans livres, modèle idéal futuriste où le document physique ne sera plus directement manié, mais consulté à distance, sur un écran, sous la forme d'un sosie numérisé (Debray, 1991 : 209).

Le choix de l'expression « sosie » est intéressant, car elle souligne vraisemblablement que la copie numérique *ressemble*[100] à ce qu'elle reproduit ou copie. En effet, nous remarquerons dans les paragraphes suivants que, lorsqu'il est appliqué aux œuvres d'art et aux collections d'images, ce principe de dématérialisation est nécessairement complété par un principe de reproduction visuelle. Autrement dit, l'expérience de la matière supportant les connaissances est remplacée par l'expérience du visible (sans cette matérialité).

S'il est évident que l'ordinateur est le dispositif de prédilection pour faire circuler la plupart des archives matérielles qu'il est possible de transformer en ressources numériques[101], certains travaux révèlent parfois leurs anxiétés sur la capacité de l'ordinateur à fabriquer des images, cela même si l'ordinateur permet « de les disséminer électroniquement aux quatre coins du globe » (Batchen, 2002 : 165 ; nous traduisons de l'anglais). Hans Belting (lorsqu'il considère l'impact de la conversion numérique sur

---

[99] **Chap. 1, sect. 1.4.**

[100] Pour Jean-François Bordron, la *ressemblance* est « une relation *symétrique*, fondée sur l'*analogie de proportionnalité* et supposant une certaine *stabilité phénoménale* » (Bordron, 2011 : 157).

[101] À propos de l'informatisation croissante de la culture visuelle, voir aussi Gervereau (2003 : 341 *sq.*) ou encore Melot (2007 : 132-135).

l'expérience des archives photographiques) remarque que l'informatique semble surtout résoudre le « problème de l'entreposage et de la présentation des photos » sans pour autant remettre en question cette logique de stockage (Belting, 2004 : 281). Si, d'un côté, l'informatique modifie radicalement la vitesse de production et de transmission des images, de l'autre, la conversion numérique tend aussi à remettre en cause « l'ancienne analogie corporelle de l'image » (Belting, 2004 : 145). L'informatique et la conversion numérique des archives peuvent susciter la méfiance des approches qui affirment l'importance de cette relation corporelle aux images[102], parce que ce passage de la consultation d'images matérielles à l'exploration d'un *stock virtuel de copies numériques* transforme radicalement l'expérience des archives d'images.

Dans *Le musée virtuel*, Bernard Deloche explique le déplacement entre, d'un côté, le « musée traditionnel » (ou « institutionnel ») et, de l'autre, ce que l'auteur nomme le « musée parallèle » (Deloche, 2001 : 215). Selon l'auteur, le présupposé initial du musée institutionnel est que les œuvres originales y sont sacralisées. La fonction de ce premier type de musée est de conserver et de « montrer » les œuvres originales dans une perspective de « réappropriation patrimoniale » (l'œuvre est sacralisée en tant que vestige du passé) et surtout de les entourer d'un « culte sacralisant » (l'œuvre est souvent unique et rare). Dans le cas du musée parallèle, la fonction du musée et le statut des collections d'œuvres d'art changent ; il ne s'agit plus de présenter les œuvres originales, mais de diffuser des reproductions en images de ces œuvres (des « simulacres »). Néanmoins, le rôle de ces simulacres n'est plus, selon l'auteur, de « montrer l'original ». Il s'agit plutôt de « monter du sensible ».

Ainsi, plusieurs projets visent à mettre en place des *stocks d'images numériques*, mais aussi des *musées virtuels*, c'est-à-dire un *musée de substituts* « généralement sans lieu et sans murs, et cependant tout aussi réel et certainement plus efficace que ne l'est le musée institutionnel » (*ibid.* : 161-162). Pour Deloche, dans les musées virtuels, l'œuvre d'art n'est plus « sacralisée », au contraire, l'œuvre est « banalisée » et, progressivement, elle « s'assimile à l'image ». Si le musée traditionnel se fonde sur la présence de l'observateur en face des œuvres, le musée virtuel se fonde sur une relation à distance aux œuvres par la présence de substituts (photographies ou images numériques). Pourtant, Deloche (2001 : 240) nous invite judicieusement à ne pas réduire trop vite la virtualité de

[102] **Chap. 3, sect. 3.1.2. & 3.1.3.**

ces nouveaux musées à des « images numériques » ou à un « cybermusée », même si l'accès à distance d'images est possible à travers divers environnements virtuels permettant aux internautes de consulter d'énormes quantités de substituts visuels. Une image numérique n'est pas plus ou moins virtuelle qu'une peinture ou une photographie, même s'il s'agit « d'un autre procédé de production et de reproduction, les pixels calculables se substituant aux touches de pinceau ou aux grains de la pellicule argentique » (*ibid.* : 147-148)[103]. Pour l'auteur, la virtualité qui ressort de ces musées parallèles met en évidence que c'est la notion même de musée d'images qui est déracinée de son environnement, qui est déstabilisée dans ses normes et dans ses fonctions[104].

Dans la même perspective que la reproduction photographique[105], la reproduction numérique est employée pour dématérialiser et se substituer à des objets trop fragiles ou trop précieux pour être manipulés, ou encore pour les *représenter à distance* ; cela favorise de plus en plus une expérience « indirecte » des œuvres, du patrimoine, etc. parce que le spectateur est de moins en moins *en présence* de ces œuvres, des monuments, etc. En s'appuyant sur la reproduction numérique de la reproduction photographique, la diffusion informatique amplifie l'idée de la reproduction technique et surtout la représentation visuelle des ressources, de telle sorte que la présence des substituts numériques « vient donner un sens à l'absence des objets réels » (Roustan, 2016 : 182). C'est pourquoi des recherches sur les archives et le patrimoine affirment que le support numérique transforme « les façons de penser et de faire de la recherche dans les bibliothèques patrimoniales, archives et musées » (Roustan & Monjaret, 2016 : 11)[106]. La question est toutefois de savoir dans quelle mesure le support numérique permet de

[103] Sur ce sujet, voir aussi **chap. 4, sect. 4.1.**

[104] Ce changement d'environnement est le résultat « d'une nouvelle mutation qui accroît le pouvoir de l'homme sur lui-même et sur les choses » (Deloche, *op. cit.* : 244). Pour réagir à cette transformation, différentes voies d'étude sont possibles, comme le suggère Deloche (*ibid.* : 148) : (1) « Développer la connaissance et la maîtrise technique de l'image et de ses effets » ; (2) « Construire une éthique des nouvelles images » ; (3) « Maîtriser la civilisation des images ». C'est surtout la première voie que nous considèrerons dans le présent travail. Voir aussi Laforet (2011 : 47-102).

[105] **Chap. 1, sect. 1.1.**

[106] Les professionnels de la documentation et du patrimoine envisagent par conséquent l'archivage en lien étroit avec l'actualité récente, notamment en discutant les apports de l'informatique, mais aussi les défis pour l'avenir de l'archivage numérique. Voir notamment Hamard (2011 : 161-174) ou, plus récemment, les conclusions présentées par Bruckmann (2016), les contributions dans Aa. Vv. (2017) ou encore dans Clavert & Muller (2019).

fournir une expérience à l'utilisateur sans que l'absence de matérialité entrave l'expérience des images.

De plus, ces constatations sur la reproduction numérique mettent au jour que cette reproduction vise à représenter l'original par le moyen de copies qui ressemblent aux originaux. Cependant, la « ressemblance » est distincte de la « représentation », même si ces deux notions sont proches et parfois complémentaires. La ressemblance se fonde sur l'analogie et une stabilité phénoménale. Tandis que la représentation s'appuie sur des conventions (Bordron, 2011 : 157). Même si les images reproduites par la conversion numérique et un site Web ressemblent aux images originales, ces reproductions sont toujours représentées en fonction de conventions, de normes et de règles qu'il reste à définir.

Du reste, la section précédente met en évidence au moins deux arguments en faveur de l'utilisation du support numérique dans les pratiques d'archivage d'images : (1) l'objectif est d'homogénéiser la diversité des supports et des informations descriptives sur la base d'un même format technique et d'un seul support (le numérique) ; (2) il s'agit de favoriser l'accès et l'interaction à distance avec des substituts numériques qui ressemblent à ce qu'ils reproduisent. Comme nous l'avons suggéré (cf. **supra**), l'accès à des copies dépend d'une transmission via Internet s'appuyant sur l'expérience d'une page-écran qu'il faut *regarder* et à travers lesquels il faut agir (ou opérer). Cette transmission est encadrée par une stratégie de communication en ligne qui propose une *manière de voir*, *de manipuler* et *d'explorer des ressources numériques*.

À ce propos, François Brunet (2017 : 346) constate que les musées de photographies prennent « une ampleur nouvelle avec l'économie numérique d'une part, et d'autre part l'élargissement de l'histoire (artistique) de la photographie en une histoire (culturelle) des images »[107]. Cette situation de « renaissance des archives à l'ère du numérique » aboutit pour l'auteur sur une « double logique de concentration et de dispersion » qui transforme les archives photographiques en « une ressource inépuisable de redécouverte de passés plus ou moins méconnus, plus ou moins restreints, plus ou moins locaux, au service de propos se définissant aussi bien "pour" que "contre" l'histoire » (Brunet, 2017 : 273 & 337 *sq.*). Les principaux acteurs de la valorisation de l'histoire culturelle de la photographie (les musées et institutions patrimoniales) sont

[107] Sur l'histoire de cette culture photographique, voir par exemple Frizot (2018).

tenus d'accompagner cet engouement pour les archives d'images (*ibid.* : 347). Comme le remarque Brunet, les archives d'images numériques relèveront principalement « de *l'économie de leurs usages à venir et des opérations de lecture qui leur donneront sens et valeur* » (Brunet, 2017 : 361 ; nous soulignons).

Étant donné que les ordinateurs ont considérablement modifié les conditions de visibilité et de manipulation des larges ensembles d'images, l'hypothèse du présent travail est justement de montrer comment les choix stratégiques des institutions patrimoniales influencent ces « opérations de lecture » évoquées par Brunet, notamment à travers des « figures de la lecture » (Saemmer, 2015)[108]. Comme nous l'avons dit dans l'**introduction**, nous distinguerons différentes stratégies de construction du point de vue de l'utilisateur à travers trois tendances générales. Dans chaque tendance, des sites Web sont conçus comme une forme de médiation multimédia qui a pour but de diffuser la culture visuelle via Internet en s'appuyant sur un support qui structure et manifeste ces opérations et ces figures.

2.2.1. Contrecoup de la dématérialisation et de la communication en ligne des archives d'images

À ce stade, nous constatons que la communication en ligne des archives et de collections d'images se base sur plusieurs principes :

- Le principe de dématérialisation.
- Le principe de reproduction visuelle (ou de substitution par ressemblance).
- Le principe d'interactivité.

La conversion numérique des ressources conservées dans les archives est indissociable de la mise en place d'une stratégie qui, comme le suggèrent les humanités numériques, doit prendre en compte les spécificités du champ numérique par rapport à celui des sciences humaines et sociales ; en effet, « [l]'*espace ouvert par les réseaux*, c'est-à-dire l'ensemble des lieux structurés par les pratiques de consultation et de "navigation", devient un *espace élastique* dont le cadre de référence n'est plus seulement défini par le territoire culturel » (Caune, 1997 : 120)[109]. Pour le spécialiste de la conservation des

[108] **Chap. 1, sect. 1.5.**

[109] Pour, Jean Caune (*op. cit.*), si les réseaux informatiques offrent de nouvelles techniques pour communiquer, d'un autre côté, ces nouvelles technologies « ne modifieront durablement, et de manière significative, les conditions de la communication sociale, qu'à condition de prendre en compte la réalité de

photographies Bertrand Lavédrine, il faut principalement « mettre à la portée de tous des recommandations normalisées conçues pour de grandes collections patrimoniales » (Lavédrine, 2016 : 12), c'est-à-dire des recommandations techniques concernant la collecte, le tri, l'identification, le classement, l'indexation et la conservation des collections de supports photographiques[110]. L'auteur remarque aussi qu'il faut apporter « autant de soin et d'effort à la numérisation des collections qu'à leur conservation » (*ibid.* : 311), à tel point qu'il est nécessaire d'« intégrer la numérisation dans une *stratégie globale de conservation*, en y incluant une *réflexion sur la manipulation et la préservation des originaux* » (*ibid.* : 314 ; nous soulignons).

Comme nous l'avons déjà suggéré, la maîtrise des nouvelles technologies révèle « un enjeu stratégique pour l'ensemble des professionnels de l'image », parce que ces technologies marquent « le passage d'une économie matérialisée et cloisonnée de l'image argentique à une économie dématérialisée et mondiale de l'image numérique » (cf. la contribution d'Antoine Raulin, dans Martinez, 2007 : 43). Pour compléter ce que nous avons déjà dit sur ce sujet, remarquons que des travaux en sciences de l'information et de la communication proposent d'introduire la notion de « médiatisation »[111] (cf. Davallon, 2006 : 35-44 ou encore Davallon, 2016). L'objectif est notamment de comprendre comment les institutions patrimoniales sont confrontées à repenser la place des documents photographiques au sein d'un environnement informatique[112]. L'usage de la photographie dans un environnement informatique débouche sur deux activités pratiques différentes : (1) la médiation documentaire (le stockage et la consignation de documents)

---

la communication interpersonnelle » (*ibid.* : 122). Le processus de cette communication interpersonnelle se réalise en effet « dans le jeu des imaginaires et de l'intersubjectivité », mais la maîtrise de l'informatique dans la communication se base aussi sur « la production d'un sens rendu commun par la maîtrise des effets sensibles des nouvelles technologies » (*ibid.* : 123).

[110] Sur la gestion des archives photographiques, voir Charbonneau & Robert (2001 : 47 *sq.* & 199 *sq.*) ou encore Martinez (2007 : 23-32, 57-76 & 77-90). Pour une réflexion critique sur l'archive photographique, voir aussi Wilder (2009 : 79 *sq.*). Voir aussi **chap. 1, sect. 1.2.**

[111] Nous pouvons comprendre cette notion à partir de sa définition la plus simple : « Fait de rendre ou de devenir médiat ». La notion de *médiatisation* se fonde sur le terme *médiatiser*, qui signifié « rendre médiat (quelque chose) par l'introduction d'un intermédiaire », voire : « Servir d'intermédiaire, de moyen à quelque chose pour le faire connaître ou apparaître » ; *médiat* s'oppose ici à *immédiat*, en ce sens que : est médiat ce « qui n'existe ou n'agit qu'indirectement, qui n'est en relation avec autre chose que par un ou plusieurs intermédiaires ». Voir : https://www.cnrtl.fr/definition/médiat ; https://www.cnrtl.fr/definition/médiatiser ; https://www.cnrtl.fr/definition/médiatisation.

[112] Voir entre autres : Casemajor Loustau & Gellereau (2009) ; Tardy (2010 & 2014) ; Casemajor Loustau (2012 & 2013).

et (2) la médiatisation (la diffusion et la valorisation des activités de l'institution ; cf. Saouter, 2014). Diffuser des documents photographiques dans le domaine public par le moyen du réseau Internet nécessite donc de prendre en compte au moins deux fonctions possibles de la photographie au sein de ce processus de médiation informatique. En effet, le *pôle référentiel* concerne le stock d'images d'archives où il est possible de « cueillir des informations ». Le *pôle médiatique* vise à valoriser en images les activités et les projets réalisés par une institution patrimoniale, voire par un groupe de chercheurs, en intégrant ces images dans des services Web. Dans les deux cas, ce changement d'environnement où les photographies sont diffusées apporte des contraintes qu'il ne faut pas négliger.

En effet, des travaux en sémiotique de la photographie ont montré que la conversion numérique de photographies met de côté la spécificité matérielle de l'image, car l'image numérique « est supportée par un instrument de production (le logiciel) et de visualisation, l'écran, qui devient son support matériel et qui en fait un "objet" » (cf. Basso Fossali & Dondero, 2011 : 132-136). L'écran ne renvoie plus à la technique originale de production de l'image, mais à une manipulation à travers un instrument de production et de visualisation qui, nous l'avons vu, comprend à la fois un support perceptible et un support imperceptible. Cette conversion numérique n'est pas une « traduction neutre », et l'on « risque de perdre l'intégrité écologique de la photographie avec son histoire matérielle » (*ibid.* : 133). Des propriétés du support photographique risquent d'être perdues (information inscrite sur le verso ; contexte de production, de collection ou de mise en série ; manipulations et retouches possibles sur le support physique original ; etc.) ou alors d'être remplacées par des métadonnées[113], qui sont uniquement lisibles à l'aide de logiciels et de programmes informatiques. Il faut donc « se demander ce qui survit – ou devrait survivre – de la signification inhérente à l'original, et donc au fait qu'elle ne soit pas qu'une image, mais un objet » (*ibid.* : 136)[114].

Les paragraphes précédents font ressortir que la conversion numérique favorise sans aucun doute un effacement des caractéristiques du support des documents photographiques (spécificité de support d'enregistrement, format et taille du support,

[113] Cf. **supra, sect. 2.1.**

[114] Voir aussi les discussions dans le **chap. 4, sect. 4.1.2.**

inscription au verso, retouches sur le support, etc.) parce que cette conversion contraint de transformer toutes ces caractéristiques en *informations descriptives*.

Nous avons déjà étudié cette problématique dans deux travaux où nous considérons l'impact de la numérisation sur les pratiques d'archivage photographique[115]. Même si le passage des inscriptions matérielles (documents écrits ou photographies) aux archives numériques entraîne des changements dans les pratiques d'archivage d'images, nos travaux suggèrent que tout se passe comme si ces changements n'impactaient pas l'importance accordée à la description verbale de ces images, notamment en ce qui concerne les outils pour les classer et les diffuser en ligne[116]. En effet, dans un premier temps, la conversion numérique n'a pas changé les outils d'indexation d'images ; ce n'est que dans un second temps que la conversion numérique aboutit sur des approches analytiques en mobilisant des outils informatiques pour traiter des images en fonction des pixels qui composent les images.

Somme toute, les sections précédentes mettent en évidence trois points qui rejoignent les constatations sur la reproduction numérique évoquées dans le **premier chapitre**.

- Le numérique est un principe de conversion d'autres médias (textes, images, vidéos, fichiers audios, etc.) en ressources numériques.
- Les spécificités matérielles de ces médias sont soustraites durant cette conversion numérique. Par conséquent, les informations sur les supports originaux sont uniquement accessibles sous la forme de métadonnées ou de descriptions verbales (principe de dématérialisation).

[115] Voir Honba Honba (2018 : 567 *sq*.) et Honba Honba (2022). Ces travaux ont permis de montrer que, d'un côté, la photographie documentaire est une pratique dont l'enjeu est de mobiliser la photographie pour produire des ressources documentaires, tandis que, de l'autre, l'archivage photographique rassemble et organise ces ressources pour en tirer des connaissances. Ces connaissances sont fondées sur des supports photographiques qui sont produits dans le cadre de diverses pratiques photographiques et qui sont conservés et valoriser dans le cadre d'une autre pratique (les pratiques d'archivage). Concernant ces supports, la photographie comprend à la fois une phase d'inscription (le support matériel de l'image) et une phase d'expression (le « texte-énoncé » ; cf. Basso Fossali & Dondero, 2011). La production d'une photographie peut être comprise comme le résultat d'un acte d'énonciation (chaque photographie est un texte-énoncé à interpréter en considérant les pratiques de production et de réception du texte ; tandis que le support photographique est compris à la fois comme un objet matériel et comme un support formel qui accommode cette matérialité en vue d'un usage pratique).

[116] Cf. également **chap. 1 (sect. 1.2.)**, et les discussions sur ce sujet dans le **chap. 3, sect. 3.3.**

- Enfin, la diversité des médias qui sont convertis apparaît toutefois au sein d'un même dispositif intermédiaire. Ce dispositif est manifesté par une page-écran permettant de visualiser et de manipuler des substituts (les « copies informatiques des médias » ; principe de substitution et principe d'interactivité).

Dans les paragraphes et sections suivantes, nous présenterons encore plusieurs recherches en sémiotique pour compléter notre approche sémio-rhétorique.

2.2.2. Accessibilité à distance et médiation multimédia

Dans *Qu'est-ce que le patrimoine numérique ? Une sémiologie de la circulation des archives*, Matteo Treleani (2017 : 6) remarque en outre que la gestion des archives est désormais soumise à un « principe de variation » et surtout à un « principe de conversion des documents et des formats ». Dans la continuité de la multiplication des médias qui ont considérablement augmenté les ressources dans les archives, le numérique apparaît comme une solution technologique qui a surtout apporté des changements importants dans la manière d'archiver et de stocker les documents : « Les contenus culturels *circulent* et l'évolution des médias semble être placée sous l'égide de l'*accessibilité immédiate à distance* » (*ibid.* : 6).

Cependant, Treleani souligne que la diffusion des ressources numériques ne se limite pas seulement à cette « circulation de contenus dans l'espace public », car prendre excessivement en compte le potentiel du numérique dans la circulation des ressources « pourrait paradoxalement s'assimiler à une forme de présentisme », une volonté « de *faire présent* le passé » (*ibid.* : 6-8). Ainsi, la conversion numérique et la publication sur Internet des archives favoriseraient l'*objectivation* (la *chosification*) des archives : « l'idée de *voir le patrimoine comme un stock* où l'on dépose des objets afin de les préserver de l'œuvre du temps et *que l'on ouvre pour permettre à tous de les voir* » (*ibid.* : 82-83 & 85 ; nous soulignons).

Comprendre les activités de médiation entre le public et le patrimoine numérique en suivant cette idée d'un accès direct aux connaissances revient pourtant à dissimuler cette médiation derrière l'objectivité du dispositif technique qui rend possible cette accessibilité à distance (cf. *ibid.* : 59 *sq.* & 86-87). Or le numérique n'est pas un « dispositif transparent » qui donne directement accès aux ressources sans rien y

changer[117]. Au contraire, il s'agit d'un « dispositif à deux faces », l'une étant « perceptible », l'autre étant « imperceptible ». Cette « face imperceptible » permet de programmer le support et les fonctionnalités de communication en ligne qui sont uniquement perceptibles à l'aide de la mise en scène de la page-écran.

Dans *Les nouveaux produits d'information : conception et sémiotique du document*, Peter Stockinger a déjà mis en évidence que le « changement de support devient un *changement stratégique* dans le sens où le support électronique et numérique affecte l'*organisation structurale* et *fonctionnelle* du *document* » (Stockinger : 100-101, nous soulignons). L'auteur a en effet montré que la conception de ressources documentaires peut être comprise comme un « produit d'information » (cf. *ibid.*, 1999 : 141), c'est-à-dire qu'il s'agit de produire un ensemble cohérent de ressources documentaires qu'il s'agit d'indexer, d'éditer, de mettre à jour, d'aménager. Pour Stockinger, « la production et la communication d'une information thématisée et développée dans un produit documentaire reposent également sur un choix entre différents codes et objets d'expression organisés en "scènes visuelles" » (*ibid.* : 52).

À ce propos, Alina-Elena Romescu encourage à renouveler l'« herméneutique de la connaissance des produits multimédias » (Romescu, 2010 : § 23) à travers la prise en compte d'une dimension interprétative dans la communication multimédia. Pour l'auteure, le site Web prend la forme d'un espace à explorer et avec lequel il faut interagir, mais « un espace produit par la pratique d'un lieu qui est *écrit* » (*ibid.* : § 22). L'analyse sémiotique doit par conséquent prendre en compte ces activités de navigation, et surtout considérer les produits multimédias « comme un tout multimédia » afin de mettre en évidence le « contexte » qui lie l'utilisateur et le produit multimédia (*ibid*.). Même s'il est attendu que ces scènes visuelles et ces contextes s'approprient des formes sémiotiques et des habitudes perceptives qui sont issues d'autres régimes médiatiques (notamment certaines caractéristiques de l'écriture à travers les nombreuses inscriptions verbales qui sont manifestées dans la page-écran[118]), il est donc réducteur d'envisager de circonscrire uniquement l'analyse de ces scènes et de ces contextes à des règles

[117] Cf. **supra, sect. 2.2.1.**

[118] **Chap. 1, sect. 1.3. & 1.4.** À ce propos, nous avons présenté une approche théorique qui prend en compte la dimension stratégique qui vise à programmer des scènes pratiques en fonction de l'implémentation dans l'espace et dans le temps d'objets-supports ; cf. **chap. 1, sect. 1.3.3.** Nous constatons désormais que cette dimension est l'un des principaux enjeux de la conversion numérique et de la communication en ligne du patrimoine et des archives, plus généralement, des connaissances.

grammaticales. L'expression « scène » doit être entendue dans un sens large qui prend en compte la dimension pratique de la communication.

Comme nous l'avons vu avec les travaux de Jacques Fontanille et de Roger Odin, la production du sens est manifestée au sein d'une *scène pratique* qui implique un objet-support, une relation d'énonciation et une compétence communicationnelle. Ce point est important, car toutes les stratégies de communication en ligne que nous analyserons dans la **troisième partie** et la **quatrième partie** engagent une relation d'énonciation avec l'utilisateur en fondant cette relation sur un *dispositif multimédia* qui sert d'intermédiaire entre des ressources numériques et l'expérience de l'utilisateur.

Marie Cambone (2019) remarque à ce sujet que les institutions patrimoniales participent souvent à la « conception de dispositifs numériques de médiation, en tant qu'expertes de la médiation patrimoniale, comme elles le font pour tout dispositif de médiation » (*ibid.* : § 41). Cambone (*ibid.* : § 6) suggère trois axes pour analyser ce type de dispositifs : (1) « l'attention portée sur les concepteurs du dispositif » ; (2) « la scénarisation de l'information », c'est-à-dire « la mise en scène de la plateforme pensée comme un récit » ; (3) « l'implication de l'internaute dans les scénarios d'usage des dispositifs numériques ». Ces trois axes indiquent clairement que le but de la conception est de scénariser et de rendre lisible ces dispositifs de médiation. Ils suggèrent aussi que ces dispositifs s'appuient sur l'*implication* de l'utilisateur (dans le même sens, Howells remarque que le site Web est une « nouvelle forme » qui est caractérisée par l'*interaction* et les *choix de navigation de l'utilisateur*).

Dans une section suivante[119], nous présenterons un premier point de vue sur la conception esthétique (le *design*) des interfaces Web, puisque celles-ci occupent une place centrale dans la migration des ressources documentaires vers l'« hypersphère ».

Pourtant, il faut d'emblée remarquer que l'informatique évolue, de même que les choix stratégiques pour organiser des ressources documentaires à l'aide de l'informatique. L'aspect et les fonctionnalités des interfaces Web ne sont pas les mêmes au début des années 2000 qu'à la fin des années 2010 ou au début des années 2020[120].

[119] Cf. **infra, sect. 2.3.1.**

[120] Voir, entre autres, l'intervention de Peter Stockinger aux journées d'étude de l'équipe *Communication, Pratiques du Sens et Textualités* du *Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales,* en juillet 2015, à l'Université Toulouse Jean Jaurès, dont l'enregistrement est disponible à l'adresse : https://www.canal-u.tv/video/universite_toulouse_ii_le_mirail/....

L'un des enjeux actuels dans les activités de conception des sites Web est par exemple de s'adapter (en tant que « lieu personnalisable ») aux « souhaits, besoins ou désirs de ses utilisateurs, de ses parties prenantes » (cf. Stockinger, 2018). Cette adaptation concerne autant la conception visuelle que l'organisation du contenu du site. De plus, les progrès de la « science des données » (*data science*) entraînent « un renouvellement profond de la médiasphère numérique » développée depuis le début des années 2000, de telle sorte que « tout site web peu importe s'il relève de la génération du web 1, du web 2 ou encore du web 3 […] constitue pour son écosystème de communication une solution plus ou moins réussie, toujours perfectible à une classe donnée de problèmes » (*ibid.*).

Nous limitons le présent travail à l'étude de quelques « solutions provisoires » concernant la communication en ligne de larges ensembles d'images numériques. Même si ces solutions sont apparemment diverses, elles répondent toutes à un même besoin, à savoir : *créer un dispositif de médiation, scénarisé et interactif, entre les internautes et les substituts numériques des archives et des collections d'images*.

2.3. Patrimoine numérique : tensions entre technique et interprétation

Dans *Patrimoine et numérique*, Bruno Bachimont (2017b : 73-76) affirme que le numérique instaure plusieurs *tensions* entre les fonctionnalités prévues par le système informatique et l'usage effectif du numérique. Ces « tensions » sont réparties sur trois niveaux.

Au niveau théorique, le numérique est compris comme un « principe calculatoire » qui réduit les contenus à des « combinaisons automatiques » (des opérations computationnelles sont implémentées dans un support matériel).

Au niveau techno-applicatif, il s'agit de formater les données binaires pour construire un « espace des possibles » (les contenus sont structurés sous une forme calculable et manipulable à travers des formats).

Enfin, au niveau sémio-rhétorique, « le contenu numérisé s'accompagne d'une manière de faire sens héritée de son instrumentation numérique liée à l'interaction proposée, les manipulations rendues possibles, les parcours interprétatifs induits ».

Ces trois niveaux de tensions évoquent les trois cohérences (interne, concrète et externe) que Bachimont propose dans sa théorie du support[121]. Le point important est que si le numérique est nécessairement matérialisé à travers une implémentation, ce support prend toujours une consistance sémiotique à travers la manifestation des contenus numériques à travers des formes sémiotiques qui doivent être perceptibles et interprétables pour l'usager. Si, à un certain niveau, le numérique fait abstraction du sens au profit du calcul, en programmant des opérations informatiques relatives à des données chiffrées, ce sens resurgit nécessairement à un autre niveau, à savoir au niveau des possibilités d'usage et d'interaction avec des contenus numériques. C'est pourquoi les calculs du système informatique sont nécessairement couplés à un support perceptible, à savoir une interface-utilisateur qui conduit le sens vers la manifestation, tout en dissimulant le support imperceptible, à savoir le système informatique qui traite les données numériques.

2.3.1. Conception des services et des interfaces Web

Si la matérialité de l'*écran* et la *page-écran* constituent ensemble le support perceptible qui sert d'intermédiaire entre l'humain et les fonctionnalités prévues par des langages de programmation, le fait que les ordinateurs sont reliés entre eux à travers des systèmes et des réseaux informatiques *imperceptibles* est un point central de la communication en ligne. Celle-ci exige toutefois une standardisation des processus techniques de communication qui sont mis en œuvre pour manipuler et interagir avec des ressources numériques. Par exemple, le modèle OSI (*Open Systems Interconnection*) est un modèle de communication dont le but est de mettre en place un réseau de systèmes informatiques en suivant les normes qui sont proposées par l'*International Organization for Standardization* (« Organisation internationale de normalisation » ou ISO). Ce modèle comprend trois étapes principales : (1) la caractérisation du *service* à réaliser à l'aide du réseau informatique, (2) le *protocole*, c'est-à-dire la conception et la programmation informatique pour réaliser le service et (3) la conception de l'*interface* qui permet aux utilisateurs de manipuler les ressources numériques.

De plus, ce modèle définit l'architecture typique d'un réseau informatique en distinguant sept couches :

[121] **Chap. 1, sect. 1.4.**

- Couche physique
- Couche liaison de données
- Couche réseau
- Couche transport
- Couche session
- Couche présentation
- Couche application

Le modèle ne propose donc pas de service particulier ou des protocoles informatiques, mais il spécifie un cadre général pour la création de normes ultérieures cohérentes.

La couche n°1 est chargée de la transmission effective des signaux entre les interlocuteurs (émission et réception de données) ; p. ex. : des câbles.

La couche n°2 gère les communications entre deux machines directement connectées entre elles, ou connectées à un équipement qui simule une connexion directe ; p. ex. : une liaison Ethernet.

La couche n°3 gère les communications entre différentes machines qui créent le réseau ; p. ex. : adresses IP (*Internet Protocol*).

La couche n°4 concerne les communications entre les programmes en cours d'exécution, tandis que la couche n°5 concerne la synchronisation des échanges entre machines (ouverture et fermeture des sessions, c'est-à-dire la période limitée de communication entre l'utilisateur et le service).

La couche n°6 (présentation) concerne la conversion entre données manipulées au niveau applicatif et données numériques effectivement transmises et, finalement, la couche n°7 (application) est le lieu d'accès aux services du réseau informatique.

Jusqu'à présent, nous avons parlé de la page-écran et des interfaces sans expliquer comment celles-ci sont produites. Dans les paragraphes suivants, nous mettons en évidence quelques principes de base concernant la conception de la dernière « couche », celle qui nous étudions dans le présent travail.

Premièrement, la conception des interfaces Web[122] se base sur un ensemble cohérent d'éléments standardisés (les « infrastructures de développement ») qui servent

[122] Sur la conception des sites Web, voir par exemple Nogier (2008).

à créer les structures informatiques de base d'un logiciel, autrement dit, son architecture. Pour faciliter la création d'un logiciel, une « architecture informatique » est en quelque sorte un cadre de travail, un ensemble d'outils et de composants logiciels qui sont organisés conformément à des standards de programmation et qui servent de canevas de base pour les développeurs ; une architecture est toujours spécialisée dans le cadre d'un langage de programmation déterminé et en vue de réaliser des tâches spécifiques à l'aide de l'informatique. Un *framework* est un exemple d'architecture générique de structures informatiques, éventuellement composée de plusieurs bibliothèques logicielles. Le principe des *frameworks* est de proposer des modèles de référence en offrant une aide au développeur qui utilise ces modèles dans la conception de l'architecture d'un logiciel applicatif particulier. L'avantage est que ces modèles sont standardisés, réutilisés et surtout adaptés en fonction du projet et des objectifs du logiciel. Différents types de *frameworks* sont utilisés en fonction des besoins du développeur :

- Des *frameworks d'infrastructure système* pour développer des systèmes d'exploitation, des interfaces graphiques, des outils de communication.
- Des *middlewares*, pour rassembler et contrôler des applications hétérogènes au sein d'une interface unique.
- Des *frameworks pour les entreprises*, concernant des applications spécifiques au secteur d'activité de l'entreprise.
- Des *frameworks de gestion de contenu* qui mettent en place des systèmes pour créer, collecter, classer, stocker et publier des données numériques.

Deuxièmement, les interfaces Web prennent généralement la forme d'un site Web comprenant plusieurs pages Web. Ces pages présentent toujours un espace graphique limité que l'utilisateur fait « défiler » de haut en bas (plus rarement de gauche à droite) et inversement ; en outre, ces pages sont plus ou moins évidemment liées les unes aux autres. Le parcours de l'utilisateur s'effectue donc *dans l'espace d'une page* ou *de page en page*. Il faut utiliser un éditeur de texte spécialisé pour programmer les composants graphiques des pages Web, et par la suite enregistrer les pages du site Web sur un serveur qui héberge les données numériques relatives à chaque page.

Cette programmation est réalisée à l'aide d'une programmation en langage *JavaScript* en utilisant aussi des feuilles de style en cascade (*Cascading Style Sheets* ou CSS) et d'un document rédigé selon le format *Hypertext Markup Language* (HTML). À partir d'un serveur Web, le langage HTML permet d'intégrer et de disposer des

ressources dans la page (textes, images, vidéos, applications, etc.) ; ce langage sert aussi à définir l'adresse Web dans le réseau Internet (la page d'accueil correspond à l'adresse racine du site Web et propose les « hyperliens »[123] vers les principaux services et informations accessibles par le moyen de l'interface).

En combinant ce langage HTML avec les langages CSS et *JavaScript*, les navigateurs Web adaptent le rendu visuel de chaque page Web (p. ex. : dimensions de la fenêtre, caractères typographiques, préférences de l'utilisateur) par rapport aux dispositifs de consultation de l'utilisateur (écrans d'ordinateurs, tablettes, télévisions, smartphones). Étant donné que ce sont ces langages qui permettent au *webdesigner* de structurer et de mettre en forme le contenu de ces pages, d'intégrer des applications ou d'afficher plusieurs ressources multimédias, ces langages (HTML, CSS, *JavaScript*) sont donc les principaux outils de programmation informatique utilisés pour produire les différents éléments constituant les pages Web qui sont affichées[124] sur l'écran de l'utilisateur.

Troisièmement, dans le même sens que les *frameworks*, les interfaces de programmation ou API (*Application Programming Interface*) sont des outils utilisés pour la conception des interfaces Web. En simplifiant, il s'agit d'une bibliothèque logicielle qui vise à faciliter la mise en place d'autres logiciels ou applications informatiques. Une application Web est généralement hébergée et installée sur un serveur Web et non pas sur l'ordinateur de l'utilisateur. Avec les *liens hypertextes*, les *applications* sont les principaux moyens pour interagir avec l'interface Web, notamment en actionnant des *widgets* ou des *applets* par l'intermédiaire d'un navigateur Web connecté à un réseau informatique. D'une part, un composant d'interface graphique (ou *widget*) peut être défini comme un outil graphique de base intégré dans une interface utilisateur ; cet outil permet à un utilisateur d'interagir avec l'ordinateur (pour ce faire, le développeur doit assembler

[123] L'hyperlien est l'élément fondamental du « système hypertexte » inventé par Ted Nelson en 1965 dans le cadre du projet Xanadu. Il s'agit d'un système de renvois via des hyperliens qui permettent de passer d'une partie d'un document à une autre, ou d'un document à d'autres documents. Depuis les années 1990, ce système est largement utilisé sur le réseau Internet pour passer d'une page Web à une autre.

[124] Concernant l'affichage des pages du site, nous pouvons souligner que la conception de site Web prend en compte différents types d'affichage. Un *affichage fluide* repose sur des méthodes qui adaptent proportionnellement la mise en page à la largeur d'écran, tandis que l'*affichage adaptatif* vise à proposer différents types de mises en page en fonction la largeur de l'écran. D'autre part, un *affichage fixe* correspond à une méthode dont le but est de figer la taille de chaque élément de la page (le rendu optimal correspond à une mise en page fixe, avec des composants dont la taille ne change pas en fonction de la dimension l'écran).

tous les composants de l'interface pour que l'écran et le support numérique se transforment en un espace interactif). D'autre part, les *applets* visent à mettre en place une application qui fonctionne par le moyen d'un navigateur Web plutôt que d'être installé sur un serveur Web ; cela afin de proposer une interactivité plus fluide (ces applications sont programmées dans un langage différent du langage HTML).

Dans les grandes lignes, c'est ainsi que l'accès au patrimoine numérique s'appuie sur une programmation dissimulée pour créer l'*espace virtuel* sur lequel se fonde l'expérience de navigation de l'usager[125]. Tous les projets que nous étudierons dans le présent travail se basent sur un site Web qui met les internautes en rapport avec des archives ou des collections d'images numériques[126].

En outre, la plupart des sites Web conçus par des institutions patrimoniales pour rendre accessibles des ressources numériques sont couplés à une *base de données*. Celle-ci correspond à un protocole qui tient un rôle central dans les stratégies institutionnelles, parce que le système informatique sélectionne uniquement des ressources en fonction de *mots-clés* correspondant par exemple à des informations descriptives (auteur ; date ; localisation ; type de support ; etc.)[127].

Pour comprendre comment les institutions patrimoniales se sont approprié les outils de la communication en ligne dans une perspective stratégique, des explications sur le fonctionnement des bases de données sont indispensables.

2.3.2. Bases de données et interaction humain-machine

Un système de gestion de base de données (SGBD) est un logiciel qui est utilisé pour stocker, gérer et partager différents types de contenus, mais en cachant la complexité des opérations informatiques qui sélectionnent, relient et manifestent ces données. Ce type de

[125] **Chap. 1, sect. 1.3.3.**

[126] Cf. **troisième partie & quatrième partie**.

[127] La classification des archives d'images dépend des opérations du traitement documentaire des images, traitement où la description verbale des images occupe une place centrale ; cf. **chap. 1, sect. 1.2.** La spécificité des supports d'images originaux est par là même éliminée au profit des métadonnées descriptives ; sur ce sujet, voir aussi **chap. 3, sect. 3.3.** Cette procédure de sélection des images en fonction de mots-clés est un élément essentiel des premières stratégies pour rendre accessibles de larges ensembles d'images numériques ; cf. **chap. 5**.

logiciel regroupe toutes les informations sur ces contenus dans un seul logiciel[128]. Un SGBD est souvent utilisé dans le cadre de différentes applications informatiques, notamment la gestion des bibliothèques ou des inventaires numériques. Les informaticiens réutilisent telles quelles les informations produites par les spécialistes des archives et du patrimoine, mais sans prendre en compte l'interprétation subjective et la dimension réflexive à l'origine de ces informations[129].

En principe, ces logiciels de gestion de données sont indépendants des contenus et de l'usage des bases de données. Sans trop entrer dans le détail, nous pouvons remarquer que la plupart des SGBD manipulent surtout des *bases de données relationnelles*.

En informatique, une base de données relationnelle est une base de données où l'information est organisée dans des tableaux à deux dimensions représentées par des lignes et des colonnes (les lignes sont aussi appelées des *nuplets*, tandis que les colonnes sont des *attributs* ; cf. Codd, 1970). Il s'agit donc d'un instrument pour représenter des données dans des tables. Cependant, cet instrument nécessite des « données structurées », c'est-à-dire des données précises et converties dans un même format[130] ; à l'inverse, des données non structurées concernent différents types de données qui sont stockées dans leurs formats d'origine. Une base de données relationnelles est ainsi composée d'une ou plusieurs relations entre des données structurées dans des tableaux.

En simplifiant, nous pouvons dire qu'un SGBD est un instrument de structuration intermédiaire entre les utilisateurs et les données numériques, structuration établissant des relations déterminées entre plusieurs fichiers qui sont systématiquement organisées par l'ordinateur. La combinaison d'un SGBD à des bases de données comprenant à la fois des ensembles d'images numériques et des informations descriptives sur ces images permet de relier ces images et ces informations au sein d'un même logiciel, et surtout de les relier à travers un même format technique (le numérique). Étant donné que plusieurs usagers et logiciels utilisent simultanément les données, le système informatique vérifie

[128] Le logiciel peut être différent (par exemple : *MySQL* ; *Access* ; *FileMaker*, etc.), mais le principe du SGBD reste globalement le même d'un produit commercial à un autre : il permet à un administrateur de convertir et d'intégrer des informations textuelles dans le fonctionnement automatisé de la base de données.

[129] Cette séparation des tâches n'implique pas forcément une absence de dialogue entre les informaticiens et les spécialistes.

[130] Cf. **note 74.**

qu'aucune personne non autorisée n'accède à des données confidentielles contenues dans la base de données. Ainsi, les données sont accessibles, mais elles ne peuvent pas être modifiées par les utilisateurs.

Lorsque toutes les ressources sont structurées et encodées dans un SGBD, il faut coupler ce système à un site Web qui est la partie visible des services de communication en ligne. Pour ce faire, nous avons vu que des outils et des modèles standardisés (*frameworks*, *APIs*) facilitent la programmation de ces fonctionnalités, ce qui permet au programmeur et au *webdesigner* de se concentrer davantage sur l'organisation des ressources numériques et les fonctionnalités des interfaces Web[131]. À ce propos, l'ouvrage *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (Shneiderman *et al.*, 2016) propose huit « règles d'or » applicables dans la plupart des systèmes informatiques dont le but est l'interaction personne-machine[132] :

- Cohérence de l'interface (actions, terminologie, couleurs, mise en page, polices).
- Ergonomie (prendre en compte les besoins et les profils différents des utilisateurs, qui peuvent être expérimentés ou novices).
- Feed-back (retour d'informations sur utilisations de l'interface).
- Limiter les séquences d'actions (début, milieu, fin ; confirmation des actions par des séquences de dialogues limités).
- Anticiper les erreurs des utilisateurs.
- Inversion facile des actions (afin d'annuler d'éventuelles erreurs de l'utilisateur).
- Permettre aux utilisateurs de garder le contrôle (efficacité des réponses de l'interface aux actions de l'utilisateur afin de produire le résultat souhaité).
- Réduire la charge de la mémoire à court terme de l'utilisateur (les capacités limitées des utilisateurs à traiter des informations imposent de concevoir des interfaces où les utilisateurs ne doivent pas se souvenir de trop d'information).

L'interface porte bien son nom, puisqu'elle fonctionne justement comme un dispositif intermédiaire entre, d'un côté, une face orientée vers l'utilisateur, et de l'autre, une face orientée vers la machine. Comme nous l'avons remarqué, *l'utilisateur interagit uniquement avec cette interface*. C'est à travers elle qu'il passe d'une page à une autre, qu'il agit sur les éléments graphiques (liens, menus, icônes, etc.) présents dans l'espace

[131] Cf. **supra, sect. 2.3.1.**

[132] Sur les interfaces, cf. aussi : Laurel (1990) ; Johnson (1997) ; Galloway (2012) ; Spencer (2014).

de la page, qu'il écrit dans le moteur de recherche de l'interface. En retour, l'interface réagit automatiquement aux actions et requêtes de l'utilisateur, tout en masquant la complexité des opérations informatiques qui rendent cette interaction possible.

La relation d'énonciation en jeu dans les interfaces a déjà été considérée par Johanna Drucker. L'auteure apporte une nuance importante à propos de l'utilisateur de l'interface[133].

> [Une] interface est un espace dans lequel un *sujet*, et non un utilisateur, est invoqué. L'interface est un système énonciatif. Les textes et les locuteurs sont situés au sein de circonstances pragmatiques d'utilisation, de rituel, d'échange et de communautés de pratique. Les dimensions énonciatives d'une page identifient qui parle à qui et dans quels ordres de langage et de modalités énonciatives (Drucker, 2020 :148).

Drucker a notamment pris en compte la théorie de l'énonciation de Benveniste dans son approche[134]. Selon l'auteur, l'*écran parle* : « le spectateur est positionné comme le "tu" qui reçoit l'énonciation du "je" de l'écran "parlant" » (*ibid.* : 154) ; cette relation d'énonciation est même « envahissante » parce que l'écran « absorbe le regard ». Par exemple, les diagrammes et les graphiques qui sont inscrits dans le champ discursif de l'écran incarnent « un mode implicite d'adresse » (*ibid.*). Cependant, le fait que les « expressions de données », les « formats d'information et leurs visualisations » sont des instruments énonciatifs est une constatation assez récente. En effet, l'auteure précise :

> « [L]'apparente neutralité et objectivité de telles expressions graphiques les alignent sur l'idéologie des sciences empiriques, comme si le simple énoncé d'un fait quantitatif, ou d'un ordre et d'un arrangement, n'avait aucune valeur rhétorique en propre et ne servait que de support de communication pour transmettre l'information » (*ibid.* : 155).

Selon Drucker, il faut pourtant s'éloigner de cette conception « non marquée des expressions d'information » et favoriser la lecture critique de ces « expressions ». Les conséquences que l'auteure tire sont particulièrement éclairantes.

---

[133] Voir Drucker (2011, 2013 & 2017) ; ces travaux ont été traduits en français et rassemblés dans Drucker (2020).

[134] **Chap. 1, sect. 1.3.1.2.**

> Si nous élargissons cette discussion à l'analyse des interfaces interactives, nous commençons à voir les façons dont un système encore plus explicite de dispositifs visant à engager et à produire un sujet parlant et parlé est à l'œuvre. La conception des interfaces interactives et le mode répété d'utilisation des données de type recherche/interrogation/filtrage, offrent l'illusion que la structure donne le contrôle à l'utilisateur. […] Pouvons-nous concevoir des modèles d'interface qui seraient des instruments authentiques de recherche ? Qui ne seraient pas simplement des interrogations au sein de données prédéfinies qui cherchent et trient selon un programme immuable ? Comment pouvons-nous imaginer une interface qui permet la modélisation de contenu, l'argument intellectuel et l'engagement rhétorique ? Dans une telle approche, la matérialité formelle et graphique de l'interface pourrait enregistrer les dimensions performatives mais aussi les soutenir. De telles approches seraient fondamentalement distinctes de celles de la communauté IHM [interaction humain-machine]. À la place de la transparence et de la clarté, elles mettraient au premier plan l'ambiguïté et l'incertitude, des multiplicités insolubles au lieu de singularités et de certitudes. Un engagement interprétatif soutenu, et non l'exécution efficace de tâches, serait le résultat souhaité. Une telle approche de la conception, ancrée dans des principes d'interprétation et une théorie de la subjectivité, reste encore à développer. Mais elle exposerait le processus de la pensée plutôt que d'afficher les résultats fixes de l'activité intellectuelle comme s'ils étaient des produits (Drucker, 2020 : 156 & 160).

Même si nous avons vu que les interfaces sont conçues en suivant un protocole standardisé qui a tendance à homogénéiser l'aspect et les fonctionnalités des interfaces[135], ces interfaces reposent donc aussi sur un « mode d'adresse », une énonciation, notamment à travers la distribution de rôles énonciatifs.

Dans la section suivante, nous verrons que cette énonciation des interfaces s'appuie à la fois sur un principe de *visibilité* et sur un principe de *lisibilité*.

2.3.3. Point de rencontre entre rhétorique et numérique : le dispositif de lecture

Nous avons vu que le numérique est un principe de manipulation automatique appliqué à des contenus[136]. Pourtant, cet automatisme reconfigure les conditions de manipulation et

[135] Cf. **supra, sect. 2.3.1.**

[136] Cf. **supra (sect. 2.3.) & chap. 1 (sect. 1.4.)**.

d'interprétation des supports, parce que les inscriptions sont en quelque sorte vidées de leurs matérialités. Par conséquent, Bachimont insiste beaucoup (lorsqu'il considère le cas des pratiques d'archivage et du patrimoine numérique) sur le fait que plus les documents sont numérisés, plus ceux-ci sont progressivement « dématérialisés », parce que le numérique sépare les documents de leurs supports d'inscriptions originaux[137]. L'usager n'accède plus aux documents authentiques, il accède à des ressources « que l'on peut exploiter de multiples manières selon le *dispositif de lecture* que l'on emploie pour reconstruire le contenu » (Bachimont, 2017b : 102 ; nous soulignons). Déjà dans l'enregistrement et l'usage de certains médias audiovisuels (par exemple, la photographie et le cinéma), le format d'enregistrement est rarement directement lisible (négatif, film celluloïd), mais doit d'abord être reconstruit par le moyen d'une forme visible (tirages, projection). Dans ce cas, la « lisibilité » des inscriptions et des formes sémiotiques est en relation avec une dimension spatiale et temporelle sur laquelle repose l'identification d'un parcours de lecture.

En suivant Bachimont, nous pouvons distinguer deux caractéristiques concernant ce dispositif de lecture :

- Le dispositif impose une lecture qui se déploie toujours dans le temps, mais cette « temporalité peut être librement reconstruite par la lecture du contenu décodé, ou imposée par le décodage ».
- Le dispositif aboutit sur une forme visible qui peut être temporelle ou spatiale « quand elle restitue une forme directement lisible *via* une médiation ou lecture culturelle qui temporise le contenu » (cf. Bachimont, 2017b : 102).

Le support numérique permet ainsi de dupliquer infiniment des contenus et de les rendre accessibles et consultables à travers un dispositif d'*énonciation-interaction* qui impose des *modalités de lecture*[138]. C'est dans ce sens que l'utilisation du support numérique

[137] Voir **supra, sect. 2.2.**

[138] Rappelons à ce propos que la lecture est une performance qui présuppose des compétences comparables à celle de la production d'un texte (cf. Greimas & Courtés, 1993 : 206). Un texte est souvent aménagé pour guider le processus d'interprétation en vue d'un usage pratique. Umberto Eco (1979) a proposé la notion de *lecteur modèle* : le producteur d'un texte organise une « stratégie textuelle » sur la base d'un ensemble de compétences partagées avec le lecteur potentiel. Implicitement, ce « lecteur modèle » est capable de « coopérer à l'actualisation textuelle de la façon dont lui, l'auteur, le pensait et capable aussi d'agir interprétativement comme lui a agi génératívement » (cf. *ibid.* : 67-69). Ainsi, ces stratégies textuelles permettent de construire le texte avec son contexte, d'insister sur telles ou telles compétences interprétatives, de créer au fond toute l'épaisseur rhétorique d'un texte. Dans le même sens, nous voulons

dans les pratiques d'archivage aboutit sur une nouvelle économie des archives et de nouvelles « opérations de lecture »[139]. C'est là tout l'enjeu des « humanités numériques », comme nous l'avons suggéré dans une section précédente.

Puisque les pratiques patrimoniales et les approches documentaires visent à rassembler et à organiser différents types de supports au sein d'une archive, Bachimont a raison d'affirmer que les archives sont une « totalité organique » ; ou encore de souligner que si cette organicité acquiert une « valeur de témoignage », c'est justement parce que ces pratiques respectent, d'une part, un « principe de provenance » (ce principe garantit l'origine des documents), et d'autre part, un « principe de respect du fonds », c'est-à-dire : respecter l'ordre de dépôt des documents (cf. Bachimont, 2017b : 93-95). Pour l'auteur, le numérique a d'abord été considéré comme une « promesse » permettant de résoudre « les deux principaux enjeux » de la gestion des contenus : la *conservation* des contenus et la *diffusion* dans le temps des contenus (*ibid.* : 183). Pourtant, ces deux principes de préservation technique des documents ne suffisent pas pour garantir la transmission des contenus du passé ; c'est du moins ce que l'auteur met en évidence.

Afin d'élaborer « des principes adéquats tant pour les enjeux de la mémoire que pour les propriétés des supports numériques » (*ibid.* : 106), ces considérations amènent l'auteur à défendre la nécessité d'une réflexion critique sur les stratégies de conservation et de diffusion du patrimoine numérique. D'une manière plus générale, Bachimoint met en évidence trois autres points importants en ce qui concerne le patrimoine numérique.

Premièrement, les pratiques d'archivage ne se limitent pas à stocker les supports et à renouveler les formats techniques. Pour Bachimont, l'un des enjeux de l'archivage numérique est de mettre en place une « mémoire dynamique » des archives (*ibid.* : 78). L'exercice de cette mémoire doit autant se réaliser par la conservation des supports d'inscriptions que grâce aux discussions critiques sur les contenus[140]. En effet, il faut prendre en compte que la clé de cette transmission des contenus est une « pratique rhétorique de la mémoire » qui rend possible « une réinvention critique du souvenir et des

---

mettre en évidence que d'autres types de stratégies influencent la manière d'organiser les contextes de communication en ligne.

[139] Cf. **supra** (fin de la **sect. 2.2.**) **& chap. 1 (sect. 1.4. & 1.5.)**.

[140] Cet « exercice », Bachimont le considère comme « une possibilité du souvenir à entretenir et à maintenir par une pratique qui, si elle tombe en désuétude, conduit à la perte de l'intelligibilité des contenus » (*ibid.* : 196).

discours » (*ibid.* : 172-173). Cette pratique rhétorique de la mémoire ne répond pas aux questions de l'historiographie, « la bonne manière d'écrire l'histoire », comme elle ne décide pas quelle est la « bonne méthode d'investigation historique (l'épistémologie de l'histoire) » (*ibid.* : 173). Selon l'auteur, cette pratique doit plutôt laisser « la place au fait et à sa critique positive tout en intégrant l'appropriation subjective », puisque l'archive s'enrichit « constamment au gré de la pratique » qui accumule les « médiations entretenant l'intelligibilité » de l'archive (*ibid.*)[141]. Ainsi, l'archivage numérique doit s'intégrer dans une *pratique rhétorique* « permanente et entretenue » qui vise à s'approprier les archives « dans les enjeux du moment » (*ibid.* : 174). Si les pratiques de conservation du patrimoine rassemblent les principes techniques pour conserver la mémoire à travers des objets du passé (*ibid.* : 196), les contenus des archives numériques resteront en effet intelligibles tant que les connaissances et les compétences nécessaires à l'interprétation de ces objets et de ces contenus feront partie d'un arrière-plan culturel commun et discutable[142].

Deuxièmement, l'auteur distingue quatre niveaux qui conditionnent la lisibilité et la visibilité des ressources numériques (*ibid.*, 2017b : 184) ; dans l'ensemble, ces niveaux sont comparables au « modèle OSI » que nous avons présenté[143] :

- Le codage du contenu pour transformer un codage binaire en une forme physique de restitution du contenu.
- Le codage des métadonnées et des formats fichiers dont le but est d'encadrer l'usage et la manipulation du contenu.
- Les applications, qui permettent de « décoder le contenu » et de le restituer de manière perceptible et compréhensible.

[141] Voir **chap. 1, sect. 1.3.**, en particulier les discussions sur la définition de l'archive de Foucault. Notons également les travaux sur la mémoire de Maurice Hallbwachs, qui ont mis en évidence que le passé peut être réinventé à travers une reconstruction, un travail documentaire sur les sources (*mémoire historique*), mais peut aussi être recomposé de façon magique à travers un ensemble de représentations partagées par un groupe social (*mémoire collective* ; cf. Halbwachs, 1968). Voir aussi la préface de Jean Duvignaud : http://classiques.uqac.ca/.../memoire_collective_preface.html.

[142] Dans les pratiques d'archivage, nous pouvons notamment faire ressortir l'importance de l'*éditorialisation*, qui concerne les opérations pour structurer et surtout pour rendre accessibles et visibles des ressources sur le Web via un environnement numérique.

[143] Cf. **supra**, **sect. 2.3.1.**

- Les plateformes informatiques, c'est-à-dire les interfaces qui constituent les environnements où les ressources et les applications informatiques sont exploitées par un utilisateur.

Cette « lisibilité » et cette « visibilité » de la médiation informatique dépendent d'une série de processus intermédiaires qui s'appuient sur des systèmes automatisés traitant les données à la place de l'utilisateur. Il est important de souligner que si un (ou plusieurs) de ces niveaux est dysfonctionnel, alors c'est l'ensemble du processus de médiation informatique d'un contenu qui est compromis.

Troisièmement, l'auteur (*ibid.* : 195) met en évidence que la gestion des archives (matérielles ou numériques) ne se résume pas seulement à des problèmes organisationnels ou technologiques, mais nécessite surtout l'élaboration d'un rapport collectif et délibératif à la mémoire qu'il s'agit de faire dialoguer avec cette culture du numérique[144]. En effet, le support numérique et la communication en ligne ne se limitent pas seulement à un dispositif technique intermédiaire, où les pratiques d'interprétation du sens et la réflexion humaine sont absentes. Au contraire, le numérique « reconfigure l'existence sémiotique du sujet dans son ensemble », en proposant « une pluralité multi-sémiotique de normes et de pratiques » (cf. Ablali & Bertin, 2020 : 12-13). Si ces normes et ces pratiques sont nécessairement fondées sur une implémentation matérielle, elles visent aussi à mettre en œuvre une énonciation non négligeable, puisqu'il faut toujours rapporter la manipulation des ressources numériques à un dispositif intermédiaire qui impose, globalement et localement, des parcours d'énonciation et d'interprétation, ainsi que des modalités d'interaction qui se fondent principalement sur la complémentarité de l'écriture et de la lecture[145].

Les sections précédentes soulignent l'importance des aménagements techniques qu'impose la communication en ligne des archives et des collections d'images. Nous

[144] Notons que différentes perspectives sont envisageables : (1) « constituer, maintenir et renforcer la communauté universitaire associée aux contenus à préserver » ; (2) « créer une compétence spécifique pour les contenus dont on a la charge, alors qu'ils sortent progressivement du champ scolaire ou savant » ; (3) « susciter une appétence des communautés potentiellement concernées pour créer une animation, un besoin et donc une activité culturelle autour des contenus » (cf. Bachimont, *op. cit.* : 198).

[145] Dans la **troisième partie**, nous verrons comment les stratégies de communication en ligne observables sur Internet se sont progressivement éloignées d'une conception de l'interface comme un « espace à lire » (dans ce cas, l'interface se fonde en grande partie sur une « rhétorique verbale ») au profit d'une conception de l'interface comme un « espace multimédia » qui propose diverses fonctionnalités et applications pour explorer, visualiser et manipuler des ensembles d'images numériques.

avons vu que les bases de données permettent aux institutions patrimoniales qui gèrent des archives d'images d'encoder et de relier toutes les informations disponibles sur les images. Lorsqu'un utilisateur explore, par exemple, une base d'images numériques en relation avec un SGBD, cet usage se fonde sur un « dispositif de lecture » (l'interface Web) qui masque les opérations de traitement et de sélection des données (dans cet exemple, le système informatique classe et sélectionne des images, tandis que le dispositif de lecture concerne uniquement la présentation automatique des ressources). Ainsi, l'expérience interprétative de l'usager ne s'appuie pas sur la manipulation directe des données numériques, elle repose entièrement sur les possibilités de manipulation qu'offre l'interface ; c'est sans doute pourquoi Lev Manovich remarque ailleurs que l'interface fonctionne à la fois comme un « code », parce qu'il repose sur un « langage de programmation », et comme un dispositif technique qui conditionne les gestes de l'usager (cf. Manovich, 2015 [2001] : 157)[146]. Pour ce faire, les informaticiens programment les opérations techniques nécessaires pour retrouver des contenus dans un ensemble de ressources numériques en fonction des choix stratégiques de ces institutions. Ces choix orientent la conception des fonctionnalités du service Web qui sont en tension avec les pratiques d'usage du patrimoine numérique.

Nous avons souligné à plusieurs reprises que les changements dans les pratiques de communication en ligne des larges ensembles d'images numériques sont encadrés par les *choix stratégiques* des institutions, notamment à travers la *place* et les *possibilités de navigation* que le site Web prépare pour les utilisateurs. Pour étudier ces choix, cette place et cette navigation, nous mobiliserons (en la détournant un peu) la notion d'*ouverture* proposée par Umberto Eco dans *L'Œuvre ouverte* (cf. Eco, 2015 [1965]).

Rappelons-le, en considérant la problématique de l'interprétation des œuvres d'art, Eco oppose les *œuvres classiques*, où la signification est fortement déterminée par des codes et des règles d'expression, aux *œuvres contemporaines*, qui sont plus difficilement réductibles à une interprétation à partir des significations manifestées par l'œuvre. D'une manière plus générale, si les lecteurs ou les spectateurs construisent toujours leurs propres interprétations lorsqu'ils sont confrontés à des textes ou à des œuvres, cet apport interprétatif est souvent plus important dans les œuvres

[146] À ce propos, nous avons déjà remarqué que la conversion numérique des photographies n'est pas un « dispositif transparent » ou une « traduction neutre », puisque les versions informatiques ne donnent pas accès à la matérialité du support original ; cf. **supra, sect. 2.2.1.** Voir aussi **chap. 4, sect. 4.1.2.**

contemporaines (que l'auteur nomme « œuvres ouvertes ») parce que ces œuvres échappent de plus en plus aux normes artistiques, aux règles de décodage.

Ainsi, la différence entre les œuvres classiques et les œuvres contemporaines concernerait surtout les conventions et les habitudes socioculturelles qui permettent de les comprendre. Autrement dit, l'idée est que les œuvres classiques sont régies par des règles et des normes de significations qui sont clairement établies, tandis que dans le cas des œuvres « ouvertes », les significations ne sont pas aussi clairement déterminées (ou déterminables). En effet, une peinture de Raphaël ou de Vermeer semble plus évidente à « lire » et à « interpréter » qu'un *ready-made* de Marcel Duchamp ou qu'une installation vidéo de Bill Viola. Néanmoins, cette évidence est forcément provisoire, car rien n'assure qu'il sera toujours évident de « lire facilement » les œuvres classiques. De plus, des travaux sur les pratiques artistiques contemporaines permettent de les encadrer par des règles et des normes favorisant la compréhension des œuvres contemporaines[147].

Ce qui nous intéresse chez Eco, c'est justement l'idée que le spectateur est sollicité pour collaborer à l'achèvement d'une « médiation ouverte » ; même si cette médiation reste toujours incomplète, elle est une réserve inépuisable de significations qui se fonde sur la subjectivité de l'artiste comme sur celle de l'interprète. Autrement, la communication en ligne est elle aussi un processus de médiation qui est toujours incomplet et qui échappe aux règles habituelles de décodage des formes sémiotiques, cela même si les systèmes informatiques se basent sur un langage informatique qui est déterminé. Pour *signifier*, ce processus nécessite toutefois un aménagement stratégique qui encadre la participation active des utilisateurs en vue de construire du sens *avec* les « unités sémiotiques d'interprétation »[148] que l'observateur reconnaît et dégage à travers son parcours interprétatif.

Dans le présent travail, nous considèrerons seulement les interfaces Web au *niveau de la tension sémio-rhétorique*[149]. Les fonctionnalités d'usage de ces interfaces révèlent les choix stratégiques des institutions patrimoniales. Nous verrons que ces institutions ont d'abord conçu leurs services comme des dispositifs où le point de vue

[147] À ce propos, voir les discussions sur les « performances » dans Genette, 2010 : 98 *sq*., ou encore les considérations sur l'« image-pensive » dans Rancière, 2008 : 115 *sq*.

[148] **Chap. 1, sect. 1.4.**

[149] Cf. **supra**, **sect. 2.3.**

construit à travers l'interface est limité à celui d'un *utilisateur-explorateur* ou d'un *utilisateur-lecteur*[150]. Le renouvellement stratégique qui se base sur la participation des utilisateurs est arrivé dans un deuxième temps, notamment à la suite de la mise en place des expositions virtuelles et de l'intégration d'applications interactives ; dans ce cas, le point de vue de l'utilisateur correspond à celui d'un *utilisateur-observateur*[151]. Les stratégies plus récentes mettent quant à elle en avant les techniques de visualisation et la *computer vision*, ce qui débouche sur un autre changement : la construction du point de vue concerne dans ce cas un *utilisateur-expérimentateur*[152].

2.4. Conclusions

Ce **deuxième chapitre** montre que plusieurs travaux s'intéressent à l'informatisation croissante des connaissances et, plus généralement, aux usages possibles du support numérique dans les recherches en sciences humaines. Nous avons insisté sur le fait que la conversion numérique des archives et des collections d'images en ressources numériques impose de mettre en place des services Web pour accéder à ces ressources. Ces services se fondent sur des stratégies de reproduction numérique dont le but est de rendre accessibles des images dématérialisées. Le rôle des sites Web est de mettre en scène un *espace virtuel* où ces ressources sont reproduites, et où ces reproductions sont manipulées par l'utilisateur. Or cet espace ne reproduit pas les archives et les collections d'images telles quelles, il les *représente* en fonction de certaines normes et conventions.

Dans cette perspective, nous avons présenté des considérations générales sur la conception des sites Web et des bases de données.

D'un côté, nous avons montré que la communication en ligne se base sur un processus complexe, mais qu'on peut répartir en trois étapes principales : (1) le service à réaliser à l'aide du réseau informatique, (2) le protocole, c'est-à-dire la conception et la programmation informatique pour réaliser le service et (3) l'interface qui permet aux utilisateurs de manipuler les ressources numériques.

---

[150] **Chap. 5**.

[151] **Chap. 6 & 7**.

[152] Cf. **quatrième partie**.

De l'autre, nous avons vu que la communication en ligne dépend d'une série de processus intermédiaires qui s'appuient sur différents niveaux qui conditionnent la lisibilité d'un contenu numérique : (1) le codage binaire du contenu, (2) le codage des métadonnées et des formats fichiers, (3) les applications et (4) les plateformes informatiques.

En nous concentrant sur le cas particulier de la communication en ligne des larges ensembles d'images, notre objectif est d'étudier plusieurs exemples de services Web qui donnent accès à ce type de ressources. Qu'il s'agisse des plateformes en ligne d'institutions patrimoniales qui proposent d'explorer des images sous la forme d'un « catalogue » de ressources numériques ou sous la forme d'une « vitrine » des activités d'une institution patrimoniale, qu'il s'agisse de « musées virtuels » ou encore d'instruments de « visualisation », dans tous les cas, il est attendu que l'utilisateur mobilise un *dispositif de lecture* pour construire son propre parcours.

Ce « dispositif de lecture » est le point de rencontre entre la raison computationnelle et les pratiques d'usages des systèmes informatiques. C'est dans ce sens que la multiplication des supports d'images et la généralisation de la conversion numérique imposent des choix stratégiques aux institutions patrimoniales. Plusieurs travaux sur le patrimoine numérique font ressortir l'avantage de prendre en compte une approche sémiotique *et* rhétorique lorsqu'il s'agit de considérer l'interaction avec des dispositifs informatiques qui manifestent des ressources numériques. Pour que cette interaction fonctionne, nous avons insisté sur le fait que la communication en ligne s'approprie les modalités d'expression d'autres régimes médiatiques, notamment celles de l'écriture et de la lecture.

La question est cependant de savoir comment l'expérience de l'utilisateur est accommodée au sein d'un environnement informatique qui soustrait la matérialité des images. D'une part, le fonctionnement de cet environnement se fonde sur des langages de programmation, sur le calcul et sur l'automatisme. D'autre part, les usages de cet environnement s'appuient sur des formes sémiotiques qui manifestent des parcours interprétatifs. À ce propos, nous avons proposé que la communication en ligne gagne à être comprise comme une « médiation ouverte » où l'utilisateur doit apporter ses actions et sa propre expérience interprétative pour attribuer du sens à cette médiation.

Nous présenterons différents exemples de communication en ligne dans la **troisième partie** et nous étudierons les possibilités analytiques du traitement informatique d'images dans la **quatrième partie**. Dans la continuité du traitement documentaire, nous constaterons que les stratégies des institutions patrimoniales ont souvent pour but d'organiser les images en fonction de métadonnées descriptives[153]. Les institutions patrimoniales emploient le traitement informatique pour fournir un ensemble cohérent de ressources textuelles en interrelation avec des substituts d'archives et de collections d'images. La classification d'images consiste alors à formater et à structurer, dans une base de données, des informations descriptives sur des archives d'images. Le problème est donc que ce ne sont pas les images en tant que telles qui sont classées par le système informatique. Dans ce cas, le système classe des descriptions verbales, des « mots-clés ».

Nous verrons dans le **chapitre suivant** que ce type de classification/recherche en fonction d'une traduction verbale du contenu des images correspond à certaines approches qui ont insisté sur l'importance de la description verbale des images, ou encore sur la réduction de l'image à un seul message. Pour ce faire, nous passerons en revue des certains travaux en iconologie et en sémiotique sur l'analyse d'images. Nous apporterons des explications sur la place de la description verbale et des caractéristiques plastiques dans les méthodes d'analyse d'images proposées par ces deux disciplines. Cela mettra en évidence un décalage méthodologique entre la conception de l'analyse dans le traitement documentaire des images et l'évolution plus récente des travaux sur l'analyse des images et sur l'étude de la culture visuelle.

[153] En effet, ces institutions ont d'abord proposé des instruments qui permettent de rechercher et de consulter rapidement des images en fonction de mots-clés (**chap. 5**) ou de catégories (**chap. 6**). Nous verrons cependant que d'autres projets se sont progressivement éloignés de ce processus, parce que l'encadrement du visuel et des images par un métalangage descriptif est problématique ; sur ce sujet, cf. **chap. 3, sect. 3.3**.

DEUXIÈME PARTIE

ANALYSES D'IMAGES ET MÉTHODES ASSISTÉES PAR UN ORDINATEUR

CHAPITRE 3
Progrès des méthodes d'analyse d'images en histoire de l'art et en sciences du langage

La **première partie** fait ressortir que la numérisation des archives et des collections d'images nécessite une réflexion critique sur le support et les stratégies de la communication en ligne. À ce stade, nous avons souligné quatre points importants :

- Les archives et les collections d'images sont converties pour être reproduites sous la forme de ressources numériques.
- Cette reproduction dématérialise et remplace les supports originaux par des copies et des métadonnées descriptives.
- Cette dématérialisation est couplée à un service Web qui se fonde à la fois sur un support perceptible (la page-écran) et un support imperceptible (la programmation informatique).
- Les dispositifs de communication en ligne dépendent des choix stratégiques des institutions patrimoniales qui encadrent cette conversion et cette reproduction.

Nous avons également suggéré que les pratiques d'archivage proposent généralement de classer ces images en fonction de mots-clés. En effet, pour organiser une grande quantité d'images, le traitement documentaire attribue un producteur, une date et un contexte de production, un contenu, etc. à chaque image, autrement dit, des informations descriptives qui les rendent intelligibles.

Le but de ce **troisième chapitre** est de montrer que ce principe de classification des images à travers une description verbale est lié à certaines réflexions théoriques et méthodologiques[154].

Nous suivrons deux voies. D'un côté, celle de travaux en histoire de l'art qui ont pour but d'interpréter des œuvres d'art visuel (**sect. 3.1.**). De l'autre, celle de la sémiologie, puis de la sémiotique visuelle (**sect. 3.2.**), deux approches qui étudient les images dans le cadre de la théorie du langage que nous avons présenté dans le **premier chapitre**. Nous constaterons que chacune de ces voies aboutit plus récemment sur des travaux qui envisagent d'utiliser l'informatique dans leurs recherches sur les images et le visuel (**sect. 3.1.4. & 3.2.4**.).

En outre, nous mettrons en évidence deux points importants. Premièrement, la question problématique de l'indexation des images, qui se base uniquement sur des descriptions verbales des images (**sect. 3.3.**). Deuxièmement, l'idée qu'il est nécessaire d'étudier l'évolution de la production d'images à travers les formes et les figures visuelles qui sont mobilisées dans les arts visuels d'une époque à une autre (**sect. 3.4.**).

3.1. De l'histoire de l'art aux sciences des images

Dans les sections suivantes, nous présenterons successivement des travaux sur l'analyse d'images en histoire de l'art, plus particulièrement, des travaux qui se rattachent à l'*iconologie*. Si nous précisons rapidement les acquis et les limites de cette discipline, c'est surtout pour montrer que d'autres travaux plus récents qui s'en inspirent ont progressivement mis en avant d'autres enjeux concernant l'interprétation des images et du visuel. Nous mettrons en évidence trois enjeux importants : (1) l'éloignement d'une conception de l'analyse des images reposant principalement sur une traduction des images en termes linguistiques ; (2) l'importance de prendre en compte les modalités de présence et la dimension plastique des images ; ou encore (3) la possibilité d'utiliser des instruments informatiques pour comprendre le visuel.

[154] **Chap. 1, sect. 1.2.** Nous ne détaillerons pas ici exhaustivement la question de l'analyse d'images en sciences humaines, question qui a déjà été commentée et discutée dans d'autres travaux didactiques ; nous pensons notamment à l'ouvrage de Laurent Gervereau (1994 : 11-35), mais aussi à celui de Jacques Aumont (2011) ou encore celui de Martine Joly et Francis Vanoye (2015).

3.1.1. Iconologie et interprétation des arts visuels

Dans un texte intitulé « Histoire de l'art et sciences humaines », André Chastel (1980 : 99-106) reconnaît évidemment la place centrale des travaux d'Erwin Panofsky (1967 [1939], 1969 [1955], 1975)[155]. Trois relations sont constantes dans les travaux de Panofsky : (1) la relation entre l'idéal et le détail de l'enquête historique ; (2) la relation entre les concepts théoriques et la particularité des œuvres ; (3) la relation entre images et concepts (cf. Barral i Altet, 2023 [1989] : 13-14). Grâce à Panofsky, l'étude scientifique des images a en effet pris « une maturité qui a changé son cours » ; selon Chastel, cette « volonté d'articuler les données de l'art, de rendre l'histoire intelligible, dans le va-et-vient incessant des motifs formels et des thèmes iconographiques, donne à ces travaux une tension particulière et féconde » (Chastel, *op. cit.* : 110).

En simplifiant, l'*iconographie* est une discipline dont l'objet est l'identification et la classification du thème des œuvres d'art (« […] branche de l'histoire de l'art qui se rapporte au sujet ou à la signification des œuvres d'art, par opposition à leur forme », cf. Panofsky, 2021 [1939] : 19). Tandis que l'*iconologie* concerne plutôt l'interprétation du sens des œuvres d'art au sein d'une culture donnée et d'une période limitée. D'après René Lindekens (1991 : 35), cette interprétation repose fondamentalement sur la distinction entre, d'une part, l'étude des formes et des motifs dans les images (« significations primaires »), et d'autre part, la composition de ces motifs en des images et les sources littéraires que le spectateur substitue à ces images (« significations secondaires »).

Dans *L'œuvre d'art et ses significations*, Panofsky explique en effet que l'historien de l'art a pour mission d'étudier les traces léguées par les cultures humaines dans une perspective qui se fonde sur une méthode scientifique. Cette méthode (l'iconologie) consiste plus précisément à combiner l'utilisation de sources littéraires (les « documents ») avec la perception directe de l'œuvre (le « monument ») pour effectuer l'interprétation des œuvres d'art visuel (Panofsky, 1969 [1955] : 38-75). C'est donc nécessairement par le moyen de cette observation et de ces sources que l'interprète peut « formuler ce que porte l'œuvre, ce ressort interne que dans un langage non ou mal contrôlé, on ne rendrait que par à-peu-près » (Chastel, 1980 : 103 & 104). L'interprétation

[155] Notons que plusieurs historiens ont œuvré, avant Panofsky, pour accorder à l'histoire de l'art une portée scientifique. À titre d'exemple, les travaux de Jacob Burckhardt, en Suisse, ceux d'Aloïs Riegl, en Autriche, ou encore ceux d'Adolfo Venturi, de Pietro Toesca et de Roberto Longhi, en Italie. Pour un point de vue (américain) sur l'histoire de l'art, voir par exemple Upjohn, Wingert & Malher (1975).

des œuvres est donc l'affaire d'un langage et surtout de discours sur l'art, puisque, selon Panofsky, une enquête iconologique vise à argumenter, à travers des discours verbaux, une interprétation du sens d'une œuvre d'art sur la base des détails visuellement perceptibles (et donc descriptibles) dans une œuvre ; ces arguments se basent sur une littérature ancienne relative aux thématiques mobilisées dans l'œuvre et aux représentations culturelles dans lesquelles l'œuvre s'inscrit.

Trois composantes principales affectent le spectateur lors de la mise en présence d'une œuvre d'art : (1) la forme concrète de l'œuvre, (2) l'idée ou le sujet de l'œuvre, (3) le contenu de l'œuvre. Ces composantes s'actualisent dans l'expérience esthétique de l'œuvre qui ne dépend pas seulement de « la sensibilité naturelle du spectateur, mais aussi de son équipement culturel » (Panofsky, 1969 [1955] : 64-65). Il est important de remarquer que, pour Panofsky, le sujet (ou la « signification ») est interprété selon trois niveaux :

- La *signification primaire* (ou *naturelle*) est saisie (1) en identifiant des *formes pures* (des représentations d'objets naturels), (2) en identifiant des *relations entre ces formes* (les événements) et (3) en observant les *qualités expressives des œuvres* ; cet « univers des formes » correspond à l'univers des *motifs*, dont le dénombrement constitue une « description pré-iconographique » de l'œuvre.
- La *signification secondaire* (ou *conventionnelle*) concerne la mise en relation des *motifs* (ou des combinaisons de ces motifs) avec des *concepts* ; les motifs auxquels l'on reconnaît une signification conventionnelle sont appelés *images* et l'identification de ces images est le domaine de l'iconographie.
- La *signification intrinsèque* (ou *contenu*) apparaît lorsqu'on prend en compte les principes sous-jacents d'une culture, d'une époque, d'une personnalité (principes qui sont condensés dans l'œuvre d'art) ; les significations primaires et secondaires sont des manipulations de ces principes sous-jacents qui, avec les procédés techniques propres à une époque et une culture, sont les symptômes d'une mentalité de base qui influence le « style de l'artiste ».

Cette distinction entre le niveau des formes et des motifs, le niveau des images et le niveau de contenu est fondamentale pour comprendre les travaux de Panofsky. Celui-ci affirme qu'il faut prendre garde aux méthodes qui relèguent à un second plan le contexte historique et le travail sur les sources littéraires, même si l'analyse des formes et des motifs est au fondement de la démarche iconologique (cf. Panofsky, 1975 : 254-255).

Dans le texte « Contribution au problème de la description d'œuvres appartenant aux arts plastiques et à celui de l'interprétation de leur contenu », l'auteur distingue en effet trois dimensions principales dans le processus d'interprétation des arts visuels : (1) l'objet de l'interprétation, (2) la source subjective de l'interprétation, (3) les correctifs objectifs de l'interprétation. L'étude iconologique d'une œuvre d'art visuel est toujours un processus interprétatif où ces trois dimensions interprétatives sont imbriquées pour former « un seul et unique processus d'ensemble parfaitement homogène se développant organiquement en une succession de tensions et de détentes » (*ibid.* : 255)[156].

Les soucis de l'historien sont donc assez évidents dans cette conception de l'analyse : l'œuvre, la perception subjective de l'œuvre, le contexte et les sources littéraires relatives à l'œuvre. Nous ne discuterons pas plus les propositions de Panofsky – celles-ci sont beaucoup plus complexes que ce que nous venons d'évoquer[157]. Cette présentation rapide indique toutefois que l'iconologie de Panofsky cherche essentiellement à faire dialoguer l'observation directe des motifs et des concepts avec une expérience érudite des images. Ce dialogue est légitimé par un cadre de réflexion historique sur les liens significatifs qui unissent, dans une culture donnée, l'expérience d'un spectateur et des phénomènes d'expressions. Dans les recherches en histoire de l'art, ces phénomènes correspondent à des formes et des styles qui sont propres aux œuvres d'art visuel[158]. Pourtant, Georges Didi-Huberman (1990) a justement souligné que cette

[156] Pour mettre en évidence ces tensions et détentes, chaque dimension est également répartie en trois sous-niveaux, qui sont proches des trois niveaux d'interprétation de la signification des œuvres : (1) le niveau du « sens-phénomène » (l'expérience subjective de l'œuvre, en tant que « sens-chose » ou « sens-expression ») ; (2) le niveau du « sens-signification » (les connaissances et ressources littéraires mobilisables) ; et (3) celui du « sens-document » (qu'il faut rapporter à des comportements humains porteurs d'une « vue métaphysique du monde » (*Weltanschauung*) ; cf. *ibid.* : 254-255). En outre, le processus subjectif d'interprétation de l'œuvre doit être confronté à des *correctifs de l'interprétation* que Panofsky fait reposer sur approche historique des œuvres. Le sens-phénomène doit être rapporté à l'*histoire des formes*, en tant que « résumé de ce qu'il est possible de représenter » ; les significations doivent être corrigées à l'aide de l'*histoire des types*, « résumé de ce dont il est possible d'avoir une notion » ; enfin, le correctif du sens-document doit être l'*histoire générale des idées*, en tant que « résumé de ce qui est possible du point de vue de la *Weltanschauung* » (Panofsky, 1975 : 254-255). Voir aussi Panofsky (1967 [1939] : 38-39) et surtout les commentaires fournis de Georges Didi-Huberman (1990 : 122 *sq.*).

[157] En amont de l'iconologie de Panofsky, voir également les travaux d'Aby Warburg (2012). Sur l'iconologie de Panofsky, cf. Panofsky (2021 [1939] : 19-39) et la présentation de Bernard Teyssèdre (*ibid.* : 11-17).

[158] Pour d'autres travaux en histoire de l'art sur la peinture, voir par exemple : Schapiro (1982 & 2000 [1973], notamment la préface d'Hubert Damisch « La peinture prise au mot », cf. *ibid.* : 5-27) ; Baxandall (1991) ; Stoichita (1993) ; ou encore Arasse (1996).

conception iconologique de l'histoire de l'art a elle-même schématisé « les limites de sa propre pratique » en favorisant les recherches autour de deux axiomes : (1) « l'art est une chose du passé » et (2) « l'art est une chose visible » qui a « son identité spécifique, son aspect discernable, son critère de démarcation, son champ clos » (*ibid.* : 54). Il fallait nécessairement dépasser ces limites. En effet, même si la visée de l'histoire de l'art est de valider des discours sur certaines périodes des arts visuels, il s'agit désormais de remettre en question l'hégémonie incontestable de l'histoire de l'art, en tant que discipline garante des « discours que portent les œuvres » (*ibid.* : 9-17, 21-64 & 107-167). À titre d'exemple, ce long passage extrait de *Devant l'image* de Didi-Huberman, où l'auteur commente l'introduction des *Essais d'iconologie* de Panofsky (2021 [1939]) :

> [La définition panofskienne du *contenu iconologique* des œuvres d'art] prétend d'abord mettre à jour ce qui, dans une image, appartient à la *sphère de la signification* – ce qui, tout bien considéré, ne va pas absolument de soi : où est, de cette sphère, le centre, où en sont l'enveloppe, les régions particulières, les limites exactes ? La signification d'autre part est-elle le seul paramètre à quoi l'on puisse référer le *contenu* d'une œuvre d'art, si cette notion a un sens ? Les œuvres d'art ne contiennent-elles pas autre chose que de la signification ? Serait-il vraiment déraisonnable d'imaginer une histoire de l'art dont l'objet soit la sphère de tous les nons-sens [sic] contenus dans l'image ? Au-delà même du *subject matter*, ou « sujet » iconographique, le *meaning* iconologique de Panofsky avait bien d'autres ambitions, en effet : il devait constituer l'instance définitive d'un lieu qui ne se contente pas d'enclore les significations portées par l'œuvre d'art mais qui, de plus, prétend les *engendrer* – donnant « leur signification même aux ordonnances formelles et aux procédés techniques mis en œuvre » dans chaque tableau, chaque sculpture et chaque édifice architecturé. Bref, le contenu iconologique « relève de l'essence » (*it is essential…*) par opposition à l'apparence, et de l'intrinsèque (*intrinsic meaning*) par opposition au conventionnel. Il répond à un concept d'où l'œuvre elle-même se pourrait déduire, comme toute superstructure se déduit de « principes sous-jacents » (*underlying principles*) ou de « principes fondamentaux qui sous-tendent le choix et la présentation » de l'œuvre elle-même, considérée en tant que phénomène expressif (Didi-Huberman, 1990 : 148-149).

Nous verrons bientôt que, à la suite de Panofsky, l'iconologie a cherché à remettre en question les fondements de ce cadre de réflexion théorique, ce qui a permis d'élargir les horizons de cette discipline incontournable. Dans la section suivante, nous

remarquerons que cette remise en question est au centre des recherches plus récentes sur les images et les médias visuels, même si ces travaux conservent à juste titre un intérêt pour cette approche iconologique.

3.1.2. Culture visuelle, *iconic turn* et réflexivité des images

Dans *Visual Culture*, Richard Howells (2003) affirme que trois approches scientifiques sont habituellement employées pour étudier et interpréter des images : l'iconologie, la sémiotique[159] et l'herméneutique (*ibid.* : 9-128). Howells précise toutefois que les spécialistes ont été amenés à accepter que ces approches ne permettent pas de comprendre toutes les significations des images, même si ces approches proposent un ensemble varié d'outils pour comprendre les images et le visuel (*ibid.* : 47-49). Selon l'auteur, étant donné que la culture visuelle est plurielle et concerne différents aspects, « il devrait en être de même pour les manières dont nous cherchons à la comprendre » (*ibid.* : 128 ; nous traduisons de l'anglais).

Dans *Voir, comprendre, analyser des images*, Laurent Gervereau (2004 : 19) affirme d'emblée que le grand mérite des travaux en iconologie est « de placer les images dans leur environnement ». L'auteur (*ibid.* : 36-91) propose notamment une grille d'analyse précise distinguant trois étapes principales : (1) la description de l'image, (2) l'étude du contexte de l'image, (3) l'interprétation de l'image[160]. Conscient des limites de

[159] Cf. **infra, sect. 3.2.**

[160] Ce type d'approche est clairement empreinte de l'iconologie (cf. **supra, sect. 3.1.1.**) ; de plus, elle convient surtout pour étudier un corpus réduit d'images. D'abord, la *description* comprend trois volets (technique ; stylistique ; thématique). La description technique concerne les producteurs de l'image et la manière de les identifier, mais encore la date et le lieu de production de l'image, le type de support et le format technique de l'image. La description stylistique met en évidence des caractéristiques sur les formes et les couleurs dans l'image. La description thématique concerne quant à elle le titre éventuel de l'image, les éléments représentés, et cherche surtout à faire ressortir la ou les thématiques d'ensemble. Ensuite, l'*étude du contexte* comprend deux volets. Le premier volet se concentre sur l'étude du contexte « en amont », c'est-à-dire qu'il s'agit de comprendre dans quel contexte technique, stylistique et thématique l'image a été produite ; il s'agit aussi de mettre en relation cette image, d'une part avec l'histoire personnelle de son producteur, d'autre part avec l'histoire et les pratiques sociales au sein desquelles l'image a été produite. Le second volet concerne l'étude du contexte « en aval ». Il s'agit d'étudier la diffusion et la réception de l'image après sa production, et de chercher s'il existe des témoignages de cette réception. Enfin, l'*interprétation* est également composée de deux volets : l'interprétation des significations et l'apport d'appréciations personnelles. D'une part, il faut chercher à savoir si le producteur de l'image a suggéré une interprétation ou si des analyses contemporaines de l'image sont disponibles (Gervereau parle alors de « significations initiales »). D'autre part, il faut prendre en compte les analyses postérieures de l'image (« significations ultérieures »). Ensuite, en fonction des éléments mis en évidence dans la

ce type de grille d'analyse, Gervereau rappelle en outre que celle-ci n'est pas forcément applicable dans son intégralité ; en effet, il semble peu probable que ce type de grilles soit appliqué sur chaque image d'une archive ou d'une collection d'images, même si ces grilles fournissent des outils pertinents pour analyser un corpus d'image[161]. En effet, lorsque la quantité et la variété des images sont importantes, un tel travail d'analyse exigerait beaucoup de temps et de main-d'œuvre.

Ces deux auteurs remarquent donc avec raison qu'il n'y a pas de clé absolue ni de discipline parfaite pour comprendre des images, notamment parce qu'il faut adapter le choix des méthodes en fonction du corpus et des objectifs scientifiques de l'analyse (cf. Howells, *op. cit.* et Gervereau, *op. cit.* : 91). Comme Howells, Gervereau a lui aussi remarqué (*ibid.* : 35) que les approches théoriques des images se situent actuellement dans une « situation d'ouverture qui permet de jeter les bases d'une *hybridation générale* des méthodes », une phase heuristique où « l'abaissement des barrières entre les différentes sciences et les doutes de chacune » rend possibles des « applications pratiques » à la croisée de ces recherches[162]. Cette ouverture s'intègre dans un contexte scientifique favorable au dialogue entre les disciplines qui étudient les images.

Dans l'introduction d'une conférence intitulée « Voir/savoir. Le partage du sensible »[163], Carl Havelange affirme en effet que l'image « s'inscrit dans une "économie générale du visuel" qui détermine un processus, passionnant, d'hybridation des objets et des approches » (Havelange, 2012 : 1). Ainsi, il faut considérer ce « désenclavement de l'image » en l'ouvrant à la question des cultures visuelles ou « cultures sensibles ». Pour l'auteur, le *sensible* peut être compris comme « une expérience que l'on fait du monde » (un « système d'ancrage, de liaison, de reconnaissance, de médiation, de relation »), et surtout comme une expérience qu'il faut penser comme une construction, et pas seulement comme une expérience donnée (cf. *ibid.* : 8). Ailleurs, le même auteur remarque aussi un autre acquis important des recherches actuelles sur le visuel : la prise

description et l'étude du contexte de l'image, il s'agit de proposer un bilan général de l'interprétation, mais aussi de s'interroger sur la manière actuelle de regarder l'image qui est étudiée.

[161] Sur la différence entre archive et corpus, cf. **chap. 1, sect. 1.3.1.**

[162] Cf. **infra, sect. 3.1.3.** Nous pouvons également constater cette inclination de la recherche vers l'interdisciplinarité dans « l'enseignement de l'histoire de l'art » (cf. Chabanne, 2016 : 59-69).

[163] Cette conférence a été présentée dans le cadre de la journée d'études « Voir/savoir : images et visualisation scientifique », organisée au Centre Alexandre Koyré (Paris).

en compte d'une autonomie de l'image par rapport au langage oral ou écrit, même s'il remarque qu'on peut parfois retrouver des caractéristiques de l'oralité et de l'écriture dans des images (cf. Havelange, 2011 : 33).

D'une manière différente des langues naturelles[164], ou encore des langages informatiques[165], le visuel impose à la fois des normes d'inscription, des règles de manipulation et un certain type de rationalité qui distribue les objets visuels dans un ensemble complexe de pratiques et d'usages[166].

Comme nous le verrons bientôt, l'histoire de l'art s'est préoccupée du problème de l'évolution des *motifs* ou des *styles* des œuvres d'art visuel, ce qui réduit beaucoup les images et les pratiques considérées. De plus, nous avons vu que certains travaux en iconologie (cf. Panofsky, 1967 [1939], 1969 [1955] & 1975) proposent sans doute des méthodes trop axées sur la description verbale et l'étude du contexte des œuvres d'art visuel. Les images sont comprises comme des œuvres (ou de monuments) auxquelles l'historien doit accorder du sens par rapport à des sources documentaires et un contexte de production et de réception.

Emmanuel Alloa souligne à ce propos que, dans le cadre de l'iconologie de Panofsky, les images « reçoivent leurs sens de ce à quoi elles renvoient », mais elles deviennent aussi des « véhicules de savoir, à condition toutefois d'avoir un "pré-texte" dont elles seront l'illustration » ; cf. Alloa, 2010 : 650). Néanmoins, l'auteur affirme aussi que l'iconologie est progressivement passée d'une conception selon laquelle l'image est une *illustration* à une conception où l'image est considérée comme un *objet spécifique* et *autonome*. Dans cette perspective, l'étude des images doit prendre en compte l'interaction avec tout le dispositif matériel qui fonde l'image et non pas favoriser uniquement une compréhension des images en les réduisant à un dispositif de représentation[167]. Si « le paradigme de la représentation était fondé sur une conception *transitive* de l'image, le paradigme de la présence renvoie à une conception *intransitive*, purement immanente de l'image qui, libérée de tout vicariat, ne tient lieu que d'elle-même » (Alloa,

[164] **Chap. 1, sect. 1.3.1.1.**

[165] Voir **chap. 1 (sect. 1.4.) & chap. 4**.

[166] Sur ce sujet, cf. les travaux sur le visuel et les images de Debray (1992), de Gervereau (2003), de Belting (2004), ou encore de Bordron (2013).

[167] Voir aussi **infra, sect. 3.2.1.**

2010 : 651). Autrement dit, il s'agit d'insister sur l'importance de la *présence matérielle*[168] des images, et surtout d'insister sur la *réflexivité* du visuel.

En effet, pour les approches qui ont mis en évidence l'expression « tournant iconique » (de l'anglais : *iconic turn*), la notion de « réflexivité du visuel » est remise au centre du débat sur l'étude des images[169]. Cette expression s'oppose à une autre, le « tournant linguistique » (*linguistic turn*), qui a pour but de mettre en évidence un changement de paradigme au sein de la philosophie : l'analyse du monde et de ses objets doit nécessairement passer par une analyse du langage, parce que c'est le langage qui donne accès au monde et à ces objets. Ce paradigme propre à la philosophie analytique implique cependant un certain « logocentrisme » – du moins une primauté du langage – que le tournant iconique vise à rectifier, car, selon cette dernière approche, le langage n'est pas le seul moyen réflexif qui nous donne accès au monde.

Comme nous l'avons déjà suggéré, il est désormais admis que le *visuel* dispose de son autonomie propre et engagerait des compositions dont le langage ne peut pas systématiquement rendre compte[170].

Dans ces travaux sur l'iconologie, Maud Hagelstein a elle aussi remarqué que cette discipline privilégie surtout des réflexions sur le *contenu des images* au détriment de questionnements sur « tout ce qui concerne les couleurs[171], les nuances, les matières, le style graphique, la densité des traits, le rythme et l'équilibre des volumes, en somme : toutes les qualités plastiques » (Hagelstein, 2017 : 493). C'est pourquoi les travaux plus

[168] La présence matérielle a d'abord été éclipsée par un intérêt pour l'effet de représentation figurative produit par le moyen des images ; cf. **infra, sect. 3.2.1**. Cette problématique de la matérialité des images, nous la retrouvons dans d'autres travaux importants en histoire de l'art sur la question de l'incarnation (cf. Didi-Huberman, 1990 & 2007).

[169] À ce propos, cf. les contributions dans Didi-Huberman & Stiegler (2008). Voir aussi les trois volumes *Penser l'image* ; cf. Alloa (2010, 2015 & 2017).

[170] Nous verrons bientôt que d'autres travaux ont proposé de distinguer deux types de lecture des images : la « lecture figurative » et la « lecture plastique » (Greimas, 1984). La première se fonde sur une grille de lecture sémantique et sur la reconnaissance des objets qui sont représentés dans une image. Tandis que la seconde se concentre sur les caractéristiques plastiques de la surface de l'image. Dans une perspective à peu près similaire, nous remarquerons également que la sémiotique visuelle a considéré la dimension rhétorique des signes visuels (notamment à travers l'opposition entre « signe iconique » et « signe plastique » ; cf. Groupe μ, 1992) ; cf. **infra, sect. 3.2.2.**

[171] Notons que les travaux sur les couleurs apparaissent plus tard en histoire de l'art ; voir notamment l'extraordinaire série d'ouvrages de Michel Pastoureau sur l'histoire du bleu (2000), du noir (2008), du vert (2013), du rouge (2016), du jaune (2019) et du blanc (2022).

récents essayent de se réapproprier le projet de l'iconologie afin de l'actualiser et d'essayer de combler ce que cette approche a d'abord laissé de côté.

> La théorie de l'image actuelle s'inscrit dans le sillage de l'iconologie en vertu de la reprise explicite de plusieurs axes problématiques centraux pour cette méthode, qui permettent son redéploiement critique : 1. la réflexion sur la représentation iconique, sur son rapport au langage et sa logique spécifique ; 2. l'enquête sur la dynamique de l'image (circulation, persistance et développement) ; 3. l'étude de l'efficacité des images et de la production de savoirs multiples sous l'effet d'opérations de montage. Ces trois axes de la recherche actuelle sur l'image gagneront certainement à être exploités encore dans de futurs travaux (Hagelstein, 2015 : §5).

Cette nouvelle approche iconologique veut comprendre les continuités et les changements dans la « mécanique des images », c'est-à-dire le jeu des forces et des dynamiques qui favorisent des continuités formelles et des contrastes visuels entre plusieurs images. Autrement dit, l'iconologie gagnerait à confronter des images qui ont été produites à différentes époques pour « déceler » des relations entre des formes visuelles que la description verbale ou l'étude du contexte peinent à circonscrire.

Pour Hagelstein, depuis la fin des années 1980, les travaux en histoire et en esthétique de l'art cherchent en effet à donner un sens actuel à l'iconologie, qui devient dès lors « l'étude transdisciplinaire des logiques iconiques » (Hagelstein, 2016 : 342). À ce propos, des travaux ont attiré l'attention sur l'approche originale de l'historien de l'art Aby Warburg et son « projet inachevé d'un vaste atlas d'images intitulé *Mnemosyne* » (Hagelstein, 2009 : 87 ; voir aussi Didi-Huberman, 2001).

L'Atlas Mnémosyne est une collection d'images en relation avec la *Kulturwissenschaftliche Bibliothek Warburg* (60 000 ouvrages et de plus de 25 000 photographies ; bibliothèque et planches d'illustrations qui proposaient un classement à partir des thèmes sur lesquels travaillait Warburg). Mis en place durant les années 1920, l'originalité du projet de Warburg est d'essayer de comprendre « le fonctionnement propre aux images, les mécanismes qui commandent le travail des formes et les grands principes de la culture », mais en examinant « les différentes *métamorphoses* que subissent les figures de l'art au cours de l'histoire » (Hagelstein, 2009 : 88)[172]. L'*Atlas* est

[172] Cf. aussi **infra, sect. 3.4.**

considéré comme un « terrain d'expérimentation visuelle » dont le but est de favoriser « les rencontres directes entre le spectateur et la masse documentaire organisée devant lui » par une confrontation visuelle des images sur plusieurs panneaux. Selon Hagelstein, cette confrontation « donne à voir l'exercice de l'inventivité des artistes eux-mêmes », tout en proposant « un espace de reprise critique et inventive des formes du passé », c'est-à-dire un espace qui contraste des images et qui, par cette mise en contraste, « fait voir que les artistes ne reprennent jamais les motifs antiques sans les investir d'un sens nouveau » (Hagelstein, 2009 : 105). Ce principe de juxtaposition et de confrontation directe à des images vise à mettre en présence et à comparer des formes visuelles dans le but d'analyser le dynamisme traversant ces images ; cette comparaison à travers la mise en présence de ces formes est considérée comme un moteur pour la compréhension de l'historien. L'idée forte de cette approche est que les formes du passé traversent les époques et servent finalement de réservoir pour les artistes, même si elles sont progressivement transformées par de nouveaux styles artistiques, car « l'histoire les "charge" de significations nouvelles » (*ibid.* : 106).

Dans l'optique de réaliser une histoire comparative de l'art sur la base des images elles-mêmes, Warburg a ainsi mis en place un outil permettant d'agencer des documents visuels pour faire ressortir des relations entre des œuvres. Le but de cet outil est de « décrire la mécanique des images » (Haglestein, 2016 : 342) en mettant en évidence des continuités formelles et des contrastes visuels entre plusieurs images qui ont été produites à différentes époques, ce qui renouvelle les conditions de lecture et d'interprétation de ces images. L'image implique toujours un contraste dont il faut « établir la logique pour décrire les effets de sens qui se développent au cœur même de la perception » (*ibid.* : 351). D'autre part, les images « ne tirent pas leur sens de similarités avec le monde réel », mais elles le tirent de « jeux internes d'oppositions visuelles qui stimulent la perception et invitent à un dépassement » (*ibid.*). Pour étudier *cette logique qui réunit ou oppose des images*, un espace théorique pour confronter des motifs et des images s'avère donc très utile. Nous verrons bientôt que ce principe de confrontation des images dans un espace théorique afin de comparer l'évolution des formes visuelles peut être rapproché d'un autre projet en histoire de l'art (celui de l'étude des œuvres d'art à travers une analyse

comparative des formes visuelles[173]) ou encore des projets en lien avec la *media visualization*[174].

3.1.3. Nouvelles conditions de visibilité des images : regarder et penser autrement

Nous avons vu que plusieurs travaux renouvellent les problématiques de l'iconologie en remarquant qu'il est préférable de prendre en compte différentes approches et méthodes pour comprendre la multiplicité actuelle des supports et des formats d'images[175]. Par conséquent, la plupart des travaux sur les images et le visuel défendent le dialogue entre différentes approches et disciplines (cf. Cohn, 2009 : 39-48).

Dans « L'œuvre d'art et les sciences humaines », Louis Marin (1973) explique déjà que le problème de l'analyse des œuvres d'art peut être abordé selon diverses approches théoriques. Pourtant, en tant qu'objet appartenant à différents champs théoriques, la spécificité et la singularité de l'œuvre sont parfois dissoutes par son « intégration particulière à un champ théorique déterminé » (*ibid.* : 479). En effet, les œuvres d'art ont longtemps été la chasse gardée de l'histoire de l'art[176]. Pour Marin, une solution possible serait de mettre en place des approches interdisciplinaires qui rassemblent les instruments méthodologiques et conceptuels en vue d'établir une « métascience de l'art »[177] (*ibid.*).

Si elle n'est pas exprimée de la même manière, cette constatation est partagée par les travaux liés à la *Bildwissenschaft* (en français : « science de l'image »). En effet, cette « science de l'image » cherche à faire dialoguer des disciplines comme l'histoire de l'art,

---

[173] Cf. **infra, sect. 3.4.**

[174] Cf. **chap. 4 & quatrième partie**.

[175] À ce propos, cf. les contributions dans Beyer, Cohn & Vladova (2010). D'autres travaux ont également proposé d'intéressantes réflexions sur les conséquences historiques et théoriques du passage progressif à un régime multimédia, c'est-à-dire aux « nouveaux médias » qui comprennent une multiplicité de techniques de production et de supports d'images (voir par exemple : Batchen, 2001 : 164-174 ; Briggs & Burke, 2009 ; Kittler, 2014).

[176] Voir **supra, sect. 3.1.2.**, en particulier les remarques de Didi-Huberman sur l'histoire de l'art.

[177] Une *métascience* a pour objectif d'étudier la science elle-même, notamment dans le but d'améliorer la qualité de la recherche scientifique ; ce terme est très peu utilisé dans les recherches francophones, où l'on rencontre davantage les termes *philosophie des sciences* ou *épistémologie*.

la sémiologie des images, ou encore la philosophie, plus précisément, l'esthétique[178]. Ces travaux en esthétique se situent souvent à la frontière de la philosophie et de l'histoire de l'art[179]. Le problème est cependant que les images étudiées par l'histoire de l'art se limitaient souvent à des œuvres artistiques (dessin, gravure, peinture, sculpture, architecture, etc.) et laissaient volontairement de côté d'autres types d'images utilisées dans des pratiques médiatiques plus récentes. C'est notamment pourquoi les recherches en *Bildwissenschaft* se sont progressivement intéressées aux modalités d'expressions visuelles à la fois dans le domaine artistique et dans certains médias de masse[180], ou encore à la place de l'expérience visuelle dans les domaines scientifiques et politiques (voir par exemple Boehm, Brandstetter & Müller, 2007).

D'un autre côté, les spécialistes des images et de l'étude de la culture visuelle intègrent davantage un point de vue archéologique et anthropologique dans leurs travaux, « parce que ce sont les hommes qui ont fabriqué et qui continuent encore de fabriquer des images » (Belting, 2004 : 8)[181]. Selon Hans Belting, pour comprendre les pratiques culturelles qui produisent des images, il s'agit de considérer ces productions dans une *configuration triangulaire* et à partir de la relation entre trois paramètres fondamentaux (« image-médium-regard » ou « image-dispositif-corps »), parce que, selon l'auteur,

---

[178] Pour un point de vue général sur l'esthétique, cf. Souriau & Souriau (1990). Pour des travaux en esthétique sur l'art contemporain, voir par exemple Heinich (1998), Jiminez (2005) ou encore Rancière (2003 & 2008). Pour un exemple de travaux en esthétique qui s'appuient sur la sémiologie, cf. Damisch (1972 & 2001), et sur les relations entre sémiotique et esthétique, cf. Château (2007).

[179] À titre d'exemple, dans *Langages de l'art*, Nelson Goodman (1990) cherche à comprendre dans quelle mesure les œuvres musicales ou picturales peuvent être considérés comme des « langages particuliers », dans le sens où ces langages reposent sur une logique autonome. L'auteur s'éloigne du point de vue iconologique au profit de celui de la philosophie analytique, notamment en considérant que les objets d'art peuvent être compris comme une forme particulière d'écriture de symboles. À travers la notion d'authenticité des œuvres, Goodman a notamment proposé de distinguer les arts *autographiques* (les œuvres sont uniques et leurs reproductions n'acquièrent dans aucun cas le même statut que les originaux ; p. ex. : une peinture) et les arts *allographiques* (les œuvres peuvent être reproduites plusieurs fois à partir d'un même modèle ; p. ex. : une partition musicale). Dans ce sens, la vérification de l'authenticité des « arts autographiques » dépend de l'authentification du support matériel original, tandis que l'authenticité des « arts allographiques » dépend d'un modèle original qui peut produire plus exemplaires d'une même œuvre. Cette distinction a notamment été reprise et largement discutée par Genette (2010 : *passim*).

[180] **Chap. 1, sect. 1.1.**

[181] Voir, entre autres : les travaux de Pierre Francastel (1956) sur les relations entre l'évolution des sciences et techniques et l'évolution des arts plastiques au XX^e^ siècle ; les travaux en paléoanthropologie d'André Leroi-Gourhan (1964 & 1965) qui remonte jusqu'aux origines des techniques de figuration ; ou encore les travaux de Howard S. Becker (2010 [1982]) sur l'aspect collectif et coopératif dans les pratiques de production des œuvres d'art.

il n'est pas possible de se « figurer une *image* sans la mettre aussitôt en corrélation étroite avec un *corps regardant* et un *médium regardé* » (*ibid.* : 9 ; nous soulignons). Belting affirme en effet que discuter « des significations que revêtent des images qui se laissent lire comme des textes est une chose ; débattre du sens que les images acquièrent et possèdent dans toute société en est une autre » (*ibid.* : 23). Il s'agit, d'une part, de prendre en compte que les images s'intègrent toujours dans un *lieu culturel*, et donc de considérer les *pratiques socioculturelles* qui encadrent le sens qu'on peut attribuer à des supports visuels[182]. Et, d'autre part, il faut considérer notre manière de regarder les images[183], plutôt que de valider une méthode scientifique parmi d'autres.

Dans un enregistrement d'une conférence sur l'interprétation des peintures, l'historien de l'art Daniel Arasse (2001) remarque justement que les œuvres picturales sont vues, aujourd'hui, très différemment qu'autrefois[184]. Les spectateurs accèdent à ces œuvres à travers les nombreuses expositions (musées, galeries) ou encore à l'aide de la reproduction photographique. Or ces œuvres sont parfois vues à travers des « conditions de visibilité » qui n'existaient pas nécessairement à l'époque de création de ces œuvres. Par exemple, dans le cas des fresques d'une église, les conditions que rencontraient les fidèles de l'époque (qui voyaient le tableau de loin) ne sont pas les mêmes que celles du public actuel d'une exposition (le public peut voir le tableau de plus près que les fidèles). Si l'historien de l'art veut comprendre le fonctionnement de l'œuvre dans l'époque et la culture visuelle où cette œuvre est apparue, il faut corriger l'impact de l'accrochage des œuvres dans les salles d'exposition des musées.

Cependant, la proximité de certaines œuvres lors d'une exposition favorise aussi un « anachronisme » (sur la question de l'anachronisme en histoire de l'art, cf. Didi-Huberman, 2000) en plaçant le spectateur dans la position rapprochée du peintre lorsqu'il peignait. Cette posture rapprochée permet, selon Arasse, de *voir autrement* ces images

---

[182] Au sujet de la sociologie de l'art et des pratiques artistiques, voir entre autres les travaux de Francastel (1956), de Duvignaud (1972), de Bourdieu & Darbel (1972).

[183] Par exemple, Didi-Huberman (1992). Voir aussi l'ouvrage de Debray (1992) qui propose une histoire du regard que l'occident porte sur les images, ou encore l'ouvrage de Havelange (1998) sur le passage, au XVI^e^ et au XVII^e^ siècles, d'un ordre ancien à un ordre moderne du regard.

[184] Voir aussi l'épisode 19 (« On y voit de moins en moins ») et l'épisode 21 (« Pour une histoire rapprochée de la peinture ») du podcast *Histoires de peintures par Daniel Arasse* (réalisé par Jean-Claude Loiseau et diffusée en juillet-août 2003) ; cf. : https://www.radiofrance.fr/franceculture/podcasts/une-histoire-de/on-y-voit-de-moins-en-moins-5991291 & https://www.radiofrance.fr/franceculture/podcasts/une-histoire-de/pour-une-histoire-rapprochee-de-la-peinture-3853419.

(d'entrer dans « l'intimité du peintre ») et surtout d'y *voir autre chose* (des détails qui ne sont pas visibles de loin).

Somme toute, les sections précédentes mettent en évidence qu'il n'y a pas de solution parfaite pour interpréter ni pour analyser des images, et à plus forte raison pour classer des images. Pour élargir le cadre de réflexion sur les théories de l'image, la plupart des approches récentes favorisent résolument un changement d'orientation dans les recherches et dans la méthodologie. Ce changement favorise un dialogue entre plusieurs disciplines et différents points de vue théoriques concernant l'interprétation des images et du visuel. Nous avons retenu trois points importants.

- Les images ne se limitent pas à des représentations figuratives qu'il faut rapporter à des sources littéraires et à un contexte de production artistique. Une image est avant tout un dispositif matériel qui s'intègre toujours dans une relation avec le regard du spectateur et dans des usages pratiques.
- Le visuel possède une réflexivité qui lui est propre. Cette réflexivité ne se fonde pas forcément sur langage verbal. C'est pourquoi des travaux en iconologie orientent davantage l'étude du visuel sur les modalités de présence des images ou encore sur l'étude contrastée d'images à partir de leurs caractéristiques formelles ou plastiques.
- L'étude de la culture visuelle ne se limite pas à des approches et des méthodes d'analyse qui se sont en grande partie intéressées aux images en tant qu'œuvres d'art du passé. Il faut aussi prendre en compte l'évolution des moyens de production des images et les changements des conditions de visibilités de ces images[185].

3.1.4. *Digital art history*

Nous avons montré que la conversion numérique des archives doit être encadrée par une stratégie globale de conservation qui prend en compte une réflexion sur la manipulation et la préservation des originaux[186]. Il s'agit de gérer le passage d'une « économie

[185] Voir aussi la discussion sur le style de visibilité des images dans le **chap. 1, sect. 1.5.**

[186] **Chap. 2, sect. 2.2.1**.

matérialisée » à une « économie dématérialisée » des archives, notamment à l'aide de la conception d'interfaces Web[187]. Dans ce contexte, des travaux en histoire de l'art envisagent plus récemment de tirer parti de cette nouvelle économie dans un but scientifique[188]. Étant donné qu'il n'y a pas une méthode parfaite pour comprendre des images, les applications des instruments informatiques dans les recherches en histoire de l'art sont sans aucun doute elles aussi limitées. Toutefois, la multiplication des travaux consacrés à la *digital art history*[189] indique que les historiens de l'art s'engagent dans cette nouvelle voie.

Dans l'article « Is There a "Digital" Art History ? », Johanna Drucker (2013) remarque que les décisions concernant la gestion de la « matérialité des images numériques » sont toutes aussi importantes que les décisions qui se rapportent à la conservation et la restauration des archives d'images matérielles.

> L'énorme corpus critique de matériaux primaires et secondaires dans le domaine de l'histoire de l'art sera mis en ligne au cours de la prochaine décennie et, ce faisant, les techniques d'exploration de données, d'analyse de réseaux et d'étude textuelle amélioreront l'histoire de l'art de manière prévisible, mais encore non reconnue et sous-utilisée. Les méthodes et les enquêtes spécifiques à la discipline de l'histoire de l'art peuvent tirer pleinement parti de la visualisation à grande échelle, de l'affichage, des présentations virtuelles et augmentées, ainsi que des modèles de connaissance rendus possibles par les technologies numériques qui découlent directement des connaissances en histoire de l'art (Drucker, 2013 : 12 ; nous traduisons de l'anglais).

Relisons certains passages de cette citation qui résume bien les attentes de l'histoire de l'art en ce qui concerne le numérique. *Les ressources concernant ce domaine seront bientôt mises en ligne afin d'être explorées*. *L'histoire de l'art bénéficiera aussi d'instruments qui amélioreront* (qui « augmenteront ») *les enquêtes des chercheurs*, notamment en visualisant des ressources et des corpus d'images en grande quantité. Nous reviendrons sur ce point crucial lorsque nous présenterons en détail ces méthodes

[187] **Chap. 2, sect. 2.3.2. & 2.3.3.**

[188] Voir aussi la discussion sur les humanités numériques dans le **chap. 2, sect. 2.1.**

[189] À ne pas confondre avec l'histoire de l'art numérique ou *history of digital art* ; cf. Paul (2015).

informatiques d'analyse d'images[190]. Or, comme nous l'avons souligné[191], même si la technologie numérique se fonde sur l'automatisme et le calcul, cette technique n'évacue pas la question des usages et de l'interprétation des ressources numériques. Si l'informatique et ses applications participent de plus en plus à la recherche scientifique par le moyen de la simulation et de l'optimisation des dispositifs qui encadrent des ressources numériques, alors il faut aussi interroger le *cadrage* qu'impose la technologie numérique dans l'interprétation critique des images.

En effet, pour Drucker, les archives d'images numériques sont autant porteuses de *valeurs interprétatives* que les archives d'images matérielles, car ce sont toujours des humains qui utilisent le potentiel analytique de l'informatique en vue de produire des interprétations.

> [Nous] savons que les dépôts vont se développer et qu'à mesure qu'ils le feront, la reconnaissance cruciale que *la numérisation n'est pas une représentation mais une interprétation* servira de tremplin critique pour la compréhension. Chaque choix effectué pour transformer une image analogique en un fichier numérique ou pour créer le format original dans le cas de ressources qui sont dès leurs créations des ressources numériques, fait partie d'une chaîne de décisions qui constituent l'artefact numérique, tout comme les décisions concernant des caractéristiques telles que la pellicule, le pigment, le substrat, la taille et/ou les techniques d'impression déterminent l'identité d'un objet analogique. La matérialité des images numériques peut être radicalement médiatisée, mais la matérialité numérique incarne des décisions et des hypothèses qui constituent l'objet en tant qu'artefact. Ces décisions sont porteuses d'une inflexion interprétative ; elles ne sont pas neutres ou sans valeur, et chacune privilégie un aspect de l'artefact numérique au détriment d'autres. Une fois prises, ces décisions ne peuvent pas toujours être inversées, et l'artefact incarne alors les hypothèses interprétatives qui ont présidé à sa formation (Drucker, 2013 : 12 ; nous traduisons de l'anglais).

D'une manière générale, la *digital art history* apporte un cadre de réflexion critique et différentes perspectives de recherches en ce qui concerne les applications de la technologie numérique dans les recherches sur les œuvres et les pratiques artistiques. Néanmoins, cette jeune discipline remarque aussi que le passage aux méthodes assistées

[190] **Chap. 4, sect. 4.2.**

[191] **Chap. 2, sect. 2.3.**

par un ordinateur ne doit pas se réaliser au détriment des approches et des concepts qui sont déjà employés dans les recherches en histoire de l'art.

Dans « Reflections on Digital Art History », Pamela Fletcher rappelle en effet que la *digital art history* « n'est pas – ou ne doit pas être – une voie distincte au sein de la discipline », et surtout pas une voie qui serait uniquement accessible aux professionnels de la technologique informatique (cf. Fletcher, 2015 ; nous traduisons de l'anglais). Il ne s'agit donc pas de remplacer les méthodes d'interprétative par des méthodes d'analyse informatique. Au contraire, l'auteur affirme qu'il faut, d'une part, contribuer à la création d'un environnement technologique « en accord avec les questions et les valeurs humanistes », et d'autre part, mettre en évidence « qu'il n'y a pas toujours de corrélation entre les techniques informatiques les plus avancées et les interprétations les plus convaincantes issues de travaux en histoire de l'art » (*ibid.*)[192]. Comme Drucker, Fletcher remarque que le traitement informatique fournit une aide précieuse en ce qui concerne la compréhension des œuvres d'art, parce que le traitement informatique des ressources numériques permet de plus en plus de « définir un contexte plus large pour leurs objets d'étude, tout en continuant à se concentrer sur l'étude d'artistes et d'objets individuels » (*ibid*.). Même s'il faut les considérer dans une perspective critique, « les techniques numériques et informatiques peuvent être des ajouts importants à la boîte à outils méthodologique de l'historien de l'art : *des machines pour penser avec*, *plutôt que des machines qui remplacent la pensée* » (*ibid.* ; nous soulignons). C'est pourquoi, selon Fletcher, il faut prendre en compte deux perspectives ouvertes par cette collaboration :

- La conversion numérique de l'histoire de l'art (*digitizing art history*).
- L'application de techniques computationnelles dans le domaine de l'histoire de l'art (*computational art history*).

Les premiers projets de conversion numérique de ressources en lien avec le patrimoine culturel et artistique présentent évidemment un intérêt pour les historiens de l'art. En effet, nous remarquerons que plusieurs projets rattachés à des institutions patrimoniales proposent des stocks d'images numériques qui ont été cataloguées et

[192] Dans ce sens, la *digital art history* partage évidemment les mêmes enjeux que les « humanités numériques » ; cf. **chap. 2, sect. 2.1.** Définir la relation de la *digital art history* « avec les domaines plus larges des humanités numériques et de l'histoire de l'art est donc un travail de collaboration en cours » (Fletcher, 2015).

indexées, puis organisées à l'aide d'un SGBD et diffusées via le réseau Internet[193]. Même si ces bases d'images permettent d'accéder à des ressources utiles pour les historiens de l'art, nous verrons que les fonctionnalités pour explorer et pour comparer ces ressources entre elles sont limitées, du moins dans un premier temps, car d'autres projets ont pour objectif de faire évoluer ces fonctionnalités, voire de mobiliser des outils informatiques pour répondre à des questions propres à la recherche en histoire de l'art[194].

Nous pouvons déjà évoquer deux raisons qui favorisent cette évolution. D'un côté, cette évolution des stratégies de communication en ligne est fondée sur la disponibilité croissante de « *big data* » (c'est-à-dire des millions images, de documents écrits et de métadonnées, qui sont disponibles dans des formats numériques). De l'autre, cette évolution s'appuie sur la disponibilité croissante d'instrument d'analyse statistique des données et sur les avancées dans les processus de traitement informatique d'images.

Cette évolution est par ailleurs discutée par une littérature scientifique et des groupes de recherche[195]. Par exemple, *Journal of Digital Art History*[196] est une revue scientifique qui vise à rassembler les progrès actuels dans le domaine du traitement informatique. L'objectif de la revue est d'interroger les exigences de l'histoire de l'art à l'égard des technologies de l'information, mais aussi de s'interroger sur la direction que va désormais prendre cette discipline (Klinke, 2016 ; Klinke & Surkemper, 2016). De plus, ce type de revue est aussi une vitrine internationale possible pour les projets de recherche en humanités numériques qui étudient actuellement cette problématique.

Les travaux qui y sont rassemblés concernent des sujets parfois très différentes. Il s'agit notamment de faire ressortir le potentiel cognitif et structurel du *cadre* en tant

---

[193] **Chap. 5**.

[194] Cf. **chap. 6, 7, 8 & 9**.

[195] À titre d'exemple, le *Digital Art History Research Group* est un groupe de recherche dont le but est d'étudier les relations entre l'histoire de l'art et les humanités numériques, ainsi que l'évolution de la remédiation numérique des connaissances et des objets propres à l'histoire de l'art germanique. Ce groupe de recherches met principalement l'accent sur des stratégies et des dispositifs innovants dont l'objectif est d'améliorer l'accès à l'information par la remédiation informatique, notamment à travers les réseaux sociaux, mais aussi à l'aide de dispositifs plus avancés de visualisation de données numériques, qui sont utilisés en tant qu'outils de traitement d'images. Voir : https://www.arthistoricum.net/en/networks/digital-art-history-research-group. Dans la même perspective, Le *Digital Art History Directory* (DAHD) est une plateforme en accès libre pour des projets en *digital art history* qui rassemblent et partagent des informations sur des projets concernant la préservation des ressources et l'archivage sur le Web ; cf. https://dahd.hcommons.org/digital-art-history-directory/.

[196] Voir : http://dah-journal.org.

que caractéristique commune dans la pratique de l'histoire de l'art et les instruments numériques (cf. Doulkaridou, 2015). Tandis que d'autres travaux ont commencé à appliquer des instruments d'analyse computationnelle en suivant le projet de Warburg que nous avons mentionné[197]. L'idée est de rapprocher les instruments informatiques de visualisations de l'*Atlas Mnemosyne* qui consistait justement à mettre en place des « montages d'images » sur différents panneaux afin de comparer les formes et les motifs composant ces images (cf. Hristova, 2016).

En suivant une autre voie, Leonardo Impett et Franco Moretti (2017) ont appliqué un traitement informatique d'images à partir de la notion de *Pathosformel* de Warburg. En simplifiant, cette notion peut être comprise comme une « force », c'est-à-dire un modèle stable qui se reproduit dans les productions visuelles à travers différentes époques, de telle sorte que nous pouvons observer que des modèles de forces survivent et persistent dans le temps. Dans cette perspective, des instruments informatiques ont été employés pour détecter des modèles de force dans un corpus réduit d'images ; cette détection ne porte pas sur les expressions faciales, elle se concentre plutôt sur les mouvements du corps (bras et jambes) dans l'image. Dans un premier temps, le traitement a pour but de segmenter et de mettre en évidence des figures humaines indépendamment du contexte. Ensuite, le traitement élimine les détails de ces figures (couleurs, vêtements, visages, mains) afin de réduire les corps à de simples *skeletons* (« squelettes »). Ces squelettes correspondent à des configurations qui font ressortir des mouvements et des postures schématiques des bras et des jambes en effaçant les particularités des corps représentés dans les images. Pour Impett et Moretti, ces configurations fonctionnent « comme un système de notation composé d'unités simples […] un alphabet pour les *Pahtosformeln* » (Impett & Moretti, 2017 : 4) que l'algorithme doit traiter comme une caractéristique morphologique.

Dans un second temps, le traitement informatique est appliqué sur images de l'*Atlas* (1 665 figures de corps à partir de 21 panneaux) en utilisant comme seule variable les articulations du corps par rapport à cet « alphabet de formes ». Même si les auteurs sont « assez confiants quant aux aspects morphologiques » mis en évidence par cette recherche, ils remarquent que l'interprétation de la signification anthropologique ou

[197] Cf. **supra, sect. 3.1.2.**

esthétique de ce type d'analyses et de résultats n'est pas évidente (Impett & Moretti, 2017 : 10).

Nous verrons bientôt que la question de l'étude des formes visuelles en fonction de « patterns » est un point de rencontre possible entre les travaux en iconologie et ceux en informatique[198]. Il ne s'agit pas de défendre ici que les historiens de l'art ont été lents à adopter les outils et les méthodes numériques ; certains projets sont entrepris par les institutions et les chercheurs du début des années 1980 au début des années 1990 (cf. Zweig, 2015). Pourtant, si les instruments d'analyse computationnelle – dont l'usage pratique était d'abord limité en histoire de l'art – sont de plus en plus accessibles et utiles pour aider la recherche, les approches traditionnelles deviennent-elles obsolètes en ce qui concerne l'analyse d'images ? La réponse est évidemment négative, puisqu'aucune approche ou aucun instrument d'analyse n'est une solution parfaite pour comprendre les œuvres et les pratiques artistiques, ou, plus précisément, les images et les pratiques qui sont liées à une culture visuelle. En étendant la méthodologie de l'histoire de l'art et de l'archéologie des médias aux instruments d'analyse informatique, il s'agit de confronter ces nouvelles méthodes aux anciennes méthodes pour étudier et comparer des images (cf. Carpo, 2018).

Nous rejoignons ici les discussions et les interrogations de la **première partie** : s'il est évident que le numérique apporte des instruments utiles pour les spécialistes qui étudient des images, comment s'agit-il d'intégrer ces instruments dans les pratiques des chercheurs ? Comment s'agit-il d'interpréter les résultats d'un traitement produit par la raison computationnelle ?

Avant de considérer plus précisément les méthodes d'analyse informatique et les solutions envisagées par les institutions patrimoniales pour communiquer en ligne leurs archives et collections d'images numériques, il reste à présenter une autre approche cruciale parmi les méthodes d'analyse d'images. Dans les sections suivantes, nous discuterons rapidement l'évolution des travaux en sciences du langage sur l'analyse d'images. Ces travaux ont d'abord tenté d'appliquer des méthodes issues de l'analyse du langage (*sémiologie*). Or nous avons vu que ce rapprochement entre la langue et les

[198] Cf. **infra, sect. 3.4.**

images est problématique (cf. **supra**, notamment les travaux sur le « tournant iconique »). Nous présenterons des travaux en sémiotique qui ont pris acte de ce problème, notamment les approches qui sont rattachées à la *sémiotique visuelle*.

3.2. Sciences du langage et interprétation d'images

En plus des travaux en iconologie, il n'est pas étonnant de remarquer que les travaux sur la gestion des archives d'images renvoient à la fois aux travaux en sémiologie et en sémiotique (cf. Kattnig, *op. cit.*). Ces deux termes se sont côtoyés au sein des théories du langage. Cependant, par rapport à la sémiologie, la sémiotique a progressivement proposé des aménagements importants concernant l'analyse des images et du visuel. D'une part, en proposant de distinguer deux manières d'appréhender l'analyse d'image (la sémiotique figurative et la sémiotique plastique). Et, d'autre part, en proposant de prendre en compte la dimension rhétorique et l'analyse énonciative des images.

3.2.1. Sémiologie et analyse d'images : message, support, représentation

Dans **la première partie**, nous avons vu que la sémiologie a d'abord proposé d'étudier la question de la signification à travers la notion de signe, même si d'autres travaux plus récents en sémiotique ont par la suite critiqué cette notion[199]. Un signe linguistique est composé d'un *signifiant* et d'un *signifié* ou encore d'un *plan de l'expression* et d'un *plan du contenu*. C'est dans cette perspective que Roland Barthes a d'abord mis en évidence que les « arts imitatifs » sont des « reproductions analogiques de la réalité » qui fonctionnent comme des « messages analogiques » (cf. Barthes, 1961). Ces messages analogiques sont composés d'un signifié et d'un signifiant, de telle sorte que le processus de signification comprend un « message dénoté » (ce à quoi le message fait référence) et d'un « message connoté » (une manière de produire, de mettre en code et d'interpréter le message). Pour Barthes, les images sont donc considérées comme des configurations de signes véhiculant un message analogique qu'il faut « décoder ».

Dans un article à propos de l'analyse sémiologique des images, Georges Mounin (1974) remarque lui aussi que *traiter une image* c'est d'abord choisir « un seul des messages spontanés qu'elle est susceptible de délivrer », puis c'est « mettre en évidence

[199] **Chap. 1, sect. 1.3.1.1. & 1.3.1.2.**

tous les points forts ou seconds […] qui assurent la communication privilégiée de ce message », enfin mettre de côté les points qui « ne contribuent pas à communiquer le message choisi » (Mounin, 1974 : 54). C'est pourquoi les premières analyses d'images en sémiologie accordent une importance centrale à la description verbale ; l'analyse revient alors seulement à rendre pertinents et à distinguer des éléments fonctionnels dans les images[200]. Cependant, cette approche sémiologique se fonde sur deux principes :

- L'image est réduite à une « reproduction analogique ».
- L'analyse des images est comprise à travers les notions de signe linguistique (signifiant/signifié) et de message (dénoté/connoté).

Sans trop entrer ici dans les détails, notons que plusieurs travaux issus des sciences du langage se sont très tôt intéressés à des images délaissées par l'histoire de l'art, notamment à travers l'analyse d'images en dehors du domaine de l'art pictural. Des travaux en sémiologie concernent par exemple les images publicitaires (Barthes, 1964) ou encore la cartographie (Bertin, 1967). En outre, plusieurs travaux en sciences du langage sont investis dans l'étude théorique de la photographie[201] et du cinéma[202]. Ces travaux s'inspirent de différentes approches : la théorie du langage de Saussure et de Hjemslev, la sémiotique de Peirce, ou encore la théorie du langage proposée par Greimas puis continuée par Fontanille.

Un point commun entre les premiers travaux en sémiologie et en sémiotique est sans doute la volonté d'affirmer la pertinence de leurs méthodologies respectives dans l'analyse des images. Certains travaux sur les images issus des sciences du langage sous-tendent pourtant un postulat problématique. En effet, il s'agit surtout de comprendre les images par rapport à la double expérience du *dicible* et du *visible*, parce que la signification des images (la peinture, mais aussi la photographie et le cinéma) sont considérées à partir de leurs ressemblances analogiques avec le monde perçu. Ce rapprochement entre la perception visuelle et les images réduit cependant ces images à une technique d'expression qui reproduit les objets tels qu'ils sont perçus dans la

[200] C'est également ce que suggère également Moles (1981). À ce propos, cf. aussi Vanmalderen (1982 : 10) et Gervereau (2004 : 26 *sq.*).

[201] À ce propos, voir Honba Honba (2018), où nous discutons plusieurs travaux importants des théories de la photographie (Barthes, 1961 & 1980 ; Dubois, 1980 ; Schaeffer, 1987 ; Basso Fossali & Dondero, 2011).

[202] Nous pensons aux travaux sur le cinéma de Christian Metz (1968, 1973, 1977, 1984 & 1991). Sur l'étude sémiotique du cinéma et de l'audiovisuel, cf. aussi Tore (2011 & 2017).

vision[203]. Comme nous l'avons vu, c'est notamment la prépondérance de cette conception « transitive » (le spectateur voit *à travers l'image*) qui a conduit les travaux plus récents en iconologie à s'intéresser davantage à l'étude des « formes visuelles » et la « présence des images »[204].

À ce propos, les travaux de Louis Marin sur l'interprétation des peintures ont souligné que le tableau est un « ensemble signifiant », mais que c'est dans le « parcours du regard » par rapport à un support que s'effectue le processus d'interprétation de l'image (cf. par exemple Marin, 1970, 1971 & 1972). Marin a en effet mis en évidence que la « fonction du regard face au tableau » institue une dissociation entre l'observation des signes et le déchiffrement du sens de l'image (Marin, 1994 : 303-312). Le tableau implique un *dispositif de représentation* où l'*invisibilité de la surface-support du tableau* « est la condition de possibilité de la visibilité du monde représenté » (*ibid.* : 305). La surface du tableau est un « dispositif réflexif-reflétant » qui s'appuie sur un *écran représentatif* qui fonctionne comme une « fenêtre à travers laquelle l'homme spectateur contemple la scène représentée sur le tableau *comme* s'il voyait la scène "réelle" du monde » (*ibid.*).

Si Marin est sans doute un des premiers à mettre en évidence l'importance du parcours du regard et du cadre dans l'analyse sémiotique des tableaux, l'auteur a aussi mis en évidence que cette réflexivité de la représentation ne s'impose pas dans l'*expérience des images* de la même manière que dans l'*expérience de la langue*[205]. Selon Marin, l'*effet-représentation* est l'un des principaux pouvoirs de l'image ; cet effet peut être compris soit comme un pouvoir de « présentification de l'absent » soit comme un pouvoir d'« autoprésentation » instituant le sujet regardant « dans l'affect et le sens » (Marin, 1993 : 14). L'efficacité de l'effet-représentation est ainsi de transformer une force

[203] Notons que ce rapprochement a été mis en évidence par des travaux en psychologie ; cf. notamment l'approche phénoménologique de Sartre (2005 [1940]) sur les différentes modalités d'apparition des images par rapport à une conscience qui les perçoit et les décrit, ou encore les travaux en psychologie d'Arnheim (1976 : 104-121 & 239-268). Ce rapprochement est également central dans l'approche bien connue de Barthes sur la photographie (Barthes, 1980 ; l'ouvrage fait directement référence au texte de Sartre).

[204] Cf. **supra, sect. 3.1.2.**

[205] C'est ce que nous avons suggéré (dans le **chap. 1, sect. 1.5.**) en présentant le point de vue de Bordron, qui estime que l'expérience d'images ouvre sur différentes scènes énonciatives, mais des scènes qu'il faut comprendre par rapport à une « économie du visuel », un « style de visibilité des images ».

potentielle en un pouvoir effectif et de valoriser cet effet « en l'instituant en état contraignant, obligatoire et légitime » (Marin, 1993 : 14).

Ce double pouvoir de l'image est déjà présenté dans des travaux antérieurs de l'auteur (cf. Marin, 1981 : 9-12). Le premier pouvoir du dispositif représentatif est de produire un effet de « présence au lieu de l'absence », tandis que le deuxième pouvoir est un « effet du sujet », c'est-à-dire un « pouvoir d'institution, d'autorisation et de légitimation comme résultante du fonctionnement réfléchi du dispositif [représentatif] sur lui-même » (*ibid.* : 10). Ce second pouvoir qui rend légitime le dispositif représentatif indique même pour Marin une certaine autonomie du visuel par rapport à la rationalité du langage, parce que ce *pouvoir de l'image* se fonde lui-même comme instrumentalisation et comme moyen d'un pouvoir, au-delà du langage et de ses règles d'expression.

Pour Marin (1994 : 343), les dispositifs de représentation comprennent donc « deux dimensions », l'une « réflexive » (« opacité du signe représentationnel ») et l'autre « transitive » (« transparence du signe représentationnel »). D'une manière très subtile, l'auteur attire l'attention sur cette « réflexivité » et cette « transitivité » dans la peinture à travers trois « dispositifs de présentation de la représentation » constituant le « cadrage général de la représentation » :

- Le fond de la représentation, la « surface d'inscription et de figuration » (*ibid.* : 344) à travers laquelle les relations entre les figures sont manifestées au regard.
- Le plan de la représentation, qui se déploie sur toute la surface du tableau et favorise donc la transparence de la représentation.
- Le cadre de la représentation, « comme frontière et limite », mais aussi comme ce qui « garnit la limite extrême de la surface géométriquement découpée de la toile », autrement dit comme un « élément structurel de construction du tableau entendu moins comme représentation ou image que comme toile » (*ibid.* : 346).

L'auteur envisage que le tableau est un « parcours narratif »[206] qui est susceptible de représenter des actions qu'il s'agit de *regarder* et de *lire*. La citation suivante est un bon exemple :

[206] Il faut sans doute comprendre ce « parcours » par rapport à un « programme narratif ». Issu de la sémiotique de Greimas, le programme narratif sert à représenter un « faire » (une action) qui est le résultat de la succession temporelle de deux états opposés ; cette succession est le produit d'un agent (« sujet de faire ») et est vécue par un patient (« sujet d'état »). Le sujet de faire peut correspondre ou ne pas correspondre au sujet d'état (« ce qui fait l'action peut être ou non ce qui en est affecté ») ; lorsque le sujet

> La représentation, dans sa dimension réflexive, se présente à quelqu'un. La présentation représentative est prise dans la structure dialogique d'un destinateur et d'un destinataire, quels qu'ils soient, auxquels le cadre fournira un des lieux privilégiés du « faire savoir », du « faire croire », du « faire sentir », des instructions et des injonctions que le pouvoir de représentation, et en représentation, adresse au spectateur-lecteur (Marin, 1994 : 348).

L'auteur ajoute ensuite dans une note :

> C'est dans cette perspective que, plus généralement, le cadre relève de l'appareil de l'énonciation et est susceptible d'être analysé selon les mêmes procédures méthodiques et techniques que celles appliquées au discours (*ibid*.).

Les réflexions de Marin sur le cadrage rejoignent les discussions sur l'énonciation que nous avons déjà présentées[207]. Dans les sections suivantes, nous verrons que plusieurs travaux en sémiotique ont également considéré l'analyse de la surface des peintures ou encore les possibilités d'analyse énonciative des images.

3.2.2. Sémiotique visuelle et rhétorique des signes visuels

Les paragraphes précédents mettent en évidence que la surface de l'image (son plan d'expression) n'est pas seulement un dispositif invisible, c'est-à-dire une « fenêtre ouverte » sur la scène qui est représentée par l'image. Cette surface de l'image est également un *ensemble signifiant* qui est inscrit sur un *support* et qui est délimité par un *cadre énonciatif*. L'expérience des peintures (et plus généralement celle des images) ne se réduit donc pas seulement à un message analogique à décoder dans le cadre restreint d'une théorie du langage. Cette expérience concerne aussi la relation ambigüe du parcours du regard par rapport à la surface de l'image[208].

---

de faire correspond au sujet d'état, il s'agit d'une action réflexive ou d'une performance, et s'il n'y a pas de correspondance, il s'agit d'une action transitive (cf. Greimas & Courtés, *op. cit.* : 297-298, ou encore Hébert, 2021, entrée « Programme narratif »). Dans tous les cas, un « état » est composé par un « sujet d'état » et un « objet d'état », entre lesquels s'établit soit une conjonction soit une disjonction (dans le cas de la conjonction, le sujet est avec l'objet ; dans le cas de la disjonction, le sujet est sans l'objet).

[207] Cf. **chap. 1**, **sect. 1.3.1.1. à 1.3.3.**

[208] Cf. **supra**, notamment lorsque nous discutions des considérations de Belting (2004).

En parallèle des travaux en sémiologie, des travaux en sémiotique ont été une étape importante concernant l'analyse des images et du visuel[209]. À propos de l'analyse des peintures, la distinction théorique entre sémiotique figurative et sémiotique plastique (Greimas, 1984) permet de considérer le plan de l'expression des images comme la surface-support de la peinture, qui est le véhicule du sens émergent de l'image (son plan du contenu)[210]. Nous avons vu que la fonction sémiotique (ou *sémiosis*) est la relation entre un plan de l'expression et un plan du contenu, relation « constituante des signes et, de ce fait, créatrice de sens » (Greimas & Courtés, *op. cit.* : 153). Le principe de départ est de considérer qu'il s'agit d'appliquer une « grille de lecture sémantique », dont le but est de distinguer des groupes de traits visuels dans la surface de l'image et à accorder des signifiés à ces traits « en transformant ainsi les figures visuelles en signes-objets » (Greimas, 1984 : 10). Ce regroupement de traits visuels constitue des « formants figuratifs », c'est-à-dire une « articulation de la surface de l'image, son découpage en unités discrètes lisibles : découpage fait en vue d'une certaine lecture de l'objet visuel » (*ibid.*). Du point de vue d'une sémiotique figurative, l'image est donc « lue », et l'interprétation est réalisée par le moyen de cette grille de lecture sémantique qui se base sur l'analogie entre l'image et les conditions de la perception naturelle.

La sémiotique figurative définit l'analyse d'images à peu près comme Marin (cf. **supra**). Cependant, le deuxième niveau de la lecture proposé par Algirdas J. Greimas vise à considérer la dimension plastique des images. Du point de vue de la sémiotique plastique, la surface de l'image est comprise comme un plan de l'expression, c'est-à-dire l'organisation du support où sont configurées les significations objectivables de l'image. Chaque élément visuel constitutif du plan de l'expression (chaque point de la surface) est défini par sa position dans le plan (catégorie topologique) et par sa couleur (catégorie chromatique) et chaque élément participe au contour d'une forme (catégorie eidétique). Dans le plan de l'expression d'une image, la configuration d'éléments manifeste des contrastes plastiques (formels et chromatiques) dans certaines zones de la surface de

[209] Concernant l'évolution des travaux sémiotiques sur l'image depuis Greimas, voir notamment les contributions dans le n°32 de la revue *La Part de l'Œil*, (Aa. Vv., 2019). Voir aussi Fontanille (2023).

[210] Suivant la théorie du langage proposée de Greimas, Felix Thürlemann (2004 [1982] : 15) avait déjà analysé la dimension plastique de *Blumen-Mythos* de Paul Klee en distinguant deux niveaux de lecture : (1) le *niveau figuratif* s'appuie sur l'application d'une lecture sémantique de l'image ; (2) le *niveau plastique* « concerne l'aspect proprement "pictural" du tableau » (Thürlemann, 2004 : 15), et non pas la fonction représentative de l'image. Cette approche sémiotique a aussi étudié les stratégies liées aux affiches et images publicitaires (Floch, 1981 & 2002 [1990]).

l'image, des « formants plastiques », auxquelles un spectateur cherchera à attribuer un sens émergent à partir de l'œuvre, et non pas à partir de la ressemblance entre l'image et le visible[211]. Par rapport aux formants figuratifs, les formants plastiques sont des « organisations particulières du signifiant qui ne se définissent que par leur capacité d'être rejointes par des signifiés et de se constituer en signes » (Greimas, 1984 : 17-18). Si les formants figuratifs signifient uniquement par le moyen d'une grille de lecture et le présupposé d'une fonction représentative, les formants plastiques favorisent d'autres significations, des composants et des relations plastiques qui concernent un « langage plastique ».

Dans un travail plus récent, Jacques Fontanille (2023 : 116) propose de « replonger la sémiotique plastique dans le bain des sciences de la culture ». Pour l'auteur, l'analyse de la dimension plastique est « exigeante et contraignante » ; dans ce travail, il rappelle notamment certaines difficultés :

> [Il] faut arrêter ou suspendre les projections symboliques, la facilité des explications contextuelles, les spéculations philosophiques ou pragmatistes ; il faut attentivement scruter l'œuvre, et s'y perdre un peu, repérer des traits répétés, ou opposés, ou singuliers, noter des corrélations, explorer des compositions, examiner avec précision des lignes et des bords, des textures et des grains, des tons et des lumières (*ibid.* : 88).

Lorsqu'il discute les propositions de Greimas, Fontanille rappelle en outre que l'enjeu d'une sémiotique plastique n'est pas d'être une « simple alternative » à la sémiotique figurative.

> Il s'agit de *faire en sorte que les formes visuelles signifient autre chose, autrement, et dans un autre langage* : la dimension plastique disposerait non seulement d'une sémiose spécifique, mais également d'une *puissance d'agir propre*. Pour que cette altérité reste perceptible, pour que le caractère « autre » de la sémiose plastique, du langage second et des signifiés profonds qu'ils portent soient sensibles et suscitent les efforts du spectateur (c'est l'*agence* du plastique), alors il faut que la signification première et la signification seconde coexistent, qu'elles soient articulables, et que les tensions, les conflits, les transpositions, les transitions de l'une vers l'autre soient toujours disponibles, dans l'œuvre, et dans l'analyse (*ibid.* : 115).

[211] Sur ce sujet, voir Bordron (2011 : 177-197) ; cf. aussi Lindekens (1986 : 45-65).

La sémiotique plastique cherche à élucider les possibilités de manifestation d'un *langage plastique spécifique à un artiste ou à une œuvre* à travers l'analyse de cette dimension plastique de l'image. Ce langage ne se fonde pas d'abord sur le postulat d'une analogie entre image et perception, mais il s'agit de l'appréhender en analysant la manière dont cette œuvre configure son plan d'expression. Lorsque nous présenterons des travaux sur l'analyse informatique d'images, nous remarquerons que ces considérations sur la sémiotique plastique peuvent être rapprochées de l'analyse des matrices de pixels[212]. Dans les deux cas, la surface de l'image est analysée au niveau de la dimension plastique. Sauf que, dans les analyses informatiques, cette dimension plastique de l'image est complètement convertie en nombres.

L'avantage de la lecture plastique est qu'elle vise à distinguer des caractéristiques plastiques propres au plan d'expression d'une image, tout en considérant le positionnement d'un observateur[213]. Même si une image ne repose pas sur un effet d'analogie entre l'image et la vision de cet observateur, celui-ci peut rapporter les formants plastiques constituant le plan de l'expression à un parcours d'interprétation (dans ce cas, l'ensemble signifiant est « non-figuratif »). Ce point est important. Au-delà de la relation avec le visible instaurée et forcée par la peinture figurative, l'observateur peut en effet attribuer du sens à une peinture non figurative en considérant plutôt l'image comme un assemblage complexe de formes et de couleurs, dont la configuration lui propose une signification cohérente dans le contexte de la pratique artistique qui en est à l'origine et qui élabore un langage spécifique[214].

D'autres approches ont proposé de considérer davantage la dimension rhétorique de l'image, tout en maintenant l'opposition entre sémiotique figurative et sémiotique plastique. Dans le *Traité du signe visuel* (1992), le Groupe μ a en effet proposé une approche de la structure sémiotique des signes visuels dans le cadre d'une rhétorique générale qui distingue le « signe plastique » et le « signe iconique ». Jean-Marie Klinkenberg a précisé ailleurs que l'objectif de cette rhétorique générale « est de décrire le fonctionnement rhétorique de toutes les sémiotiques par des opérations puissantes,

[212] **Chap. 4**.

[213] Pour une approche sur la sémiotique de l'observateur et sur la sémiotique du visible, voir notamment Fontanille (1989 & 1995).

[214] Voir aussi Marin (1994 : 177-376) ou encore Didi-Huberman (1992 : *passim*).

restant identiques dans tous les cas » (Klinkenberg, 1996 : 423)[215]. Selon l'auteur, dans les *énoncés iconiques*, il est possible d'identifier des opérations rhétoriques qu'on peut décrire. Cette description peut reposer sur des concepts déjà présents dans les travaux sur la rhétorique, mais aussi sur d'autres concepts, notamment le couple *in absentia* (dans le cas où le degré conçu de la figure rhétorique n'est pas manifesté dans le contexte) vs *in præsentia* (le degré conçu est manifesté dans le contexte ; cf. *ibid.* : 423 *sq*.).

D'un côté, le Groupe μ a proposé de comprendre le *signe iconique*[216] en fonction de quatre pôles qui sont en relation et agissent simultanément :

- Le stimulus, qui est le support matériel du signe visuel.
- Le signifiant iconique, en tant qu'ensemble de stimuli visuels correspondant à un type stable.
- Le type, la représentation mentale constituée par processus subjectif et qui permet d'établir une équivalence entre le référent et le stimulus.
- Le référent, c'est-à-dire l'objet dont le signe visuel rend compte.

De l'autre, les travaux du Groupe μ ont fait ressortir trois éléments principaux concernant la *dimension plastique des signes visuels* : (1) la forme, (2) la couleur et (3) la texture. La « position » (équivalent de la caractéristique « topologique » de Greimas) détermine l'emplacement de la *forme* par rapport à un *fond* (cf. Groupe μ, 1992 : 213), tandis que la *texture* d'une image repose sur la « microtopographie » de l'image, c'est-à-dire au niveau microscopique de l'image, où la surface est uniquement constituée par la répétition d'éléments (un élément textural est un fragment de petite dimension de telle sorte « qu'on ne puisse pas en faire une forme, car la perception individuelle de ces éléments cesse à partir d'une certaine distance, et est remplacée par une appréhension globale grâce à une opération d'intégration » ; cf. *ibid.* : 197-199). Enfin, la *couleur* correspond soit à la « couleur physique » (spectre des couleurs en longueurs d'onde, qui se base sur le rapport entre la quantité de lumière reçue et la quantité de lumière réfléchie)

[215] **Chap. 1, sect. 1.5.**

[216] Notons que (même s'ils utilisant la notion de signe) les travaux du Groupe μ intègrent à la fois l'approche post-structuraliste et des considérations sur les sciences cognitives (voir Groupe μ, 2015). Voir notamment le chapitre « Problèmes d'une sémiotique des icônes visuelles » dans Klinkenberg, 1996 : 377-432 ; pour Jean-Marie Klinkenberg, l'un des membres du Groupe μ : « L'originalité de ce système tétradique est de faire éclater la relation binaire entre /signifiant/ et "signifié", relation binaire que l'on a souvent utilisée pour décrire le signe iconique, et qui a posé des problèmes non surmontés à ce jour » (Klinkenberg, 1996 : 382).

soit à la « couleur phénoménologique » (couleur perçue en fonction de la lumière naturelle)[217].

Nous constatons donc que, d'une manière très différente, le renouvellement de l'iconologie[218] et les travaux en sémiotique visuelle se rejoignent sur un même point. Par rapport à la sémiologie qui, comme nous l'avons vu, cherche à comprendre l'image comme un message à décoder, d'autres travaux sur la peinture ont mis en évidence le rôle essentiel du parcours du regard par rapport à la surface du tableau, ou encore la distinction entre deux niveaux de lecture de l'image (sémiotique figurative et sémiotique plastique). L'analyse de la dimension plastique des images se base sur quatre catégories : catégorie topologique (position), catégorie chromatique (couleurs), catégorie eidétique (formes) et catégorie microtopographique (texture).

Ce que nous voulons mettre en évidence, c'est que les méthodes d'analyse d'images se sont éloignées de la seule question de la transitivité des images en faveur de l'analyse de l'opacité des images (dimension réflexive).

Dans la continuité de la sémiotique visuelle, nous avons suggéré que l'étude rhétorique des signes visuels indique aussi qu'il est possible de considérer les mécanismes de l'énonciation visuelle. Même si le principe d'énonciation concernait à l'origine des travaux sur le langage verbal[219], nous verrons dans la section suivante que des sémioticiens appliquent la théorie de l'énonciation à des langages non verbaux.

---

[217] La forme, la couleur et la texture d'un signe visuel sont des éléments autonomes qui participent ensemble à la constitution d'un ensemble signifiant. S'il est éventuellement possible de dégager des effets de sens (des signifiés plastiques) à partir de ces « signifiants », cependant, l'étude de ces effets de sens est complexe (les recherches du Groupe µ se sont surtout concentrées sur des énoncés visuels simples et où la relation icono-plastique peut facilement être mise en évidence). Concernant ces « effets de sens » de cette communication visuelle, d'autres travaux en sémiotique visuelle sur la peinture se basent sur le postulat que la *signification* est à la fois le résultat de l'ensemble des articulations entre expression et contenu qui délimite « la forme abstraite de l'objet de sens soumis à l'analyse », mais elle est aussi « le mode d'accroche que peut adopter, dans telle œuvre donnée, la mise en relation de sa forme de l'expression et de sa forme du contenu » (Hénault & Beyaert, 2004 : 235-236). La « signification picturale » n'est pas seulement une donnée abstraite, elle est aussi une « manière d'adhérer » (une union entre le sens et la matière). Le *sens* est quant à lui compris comme une « pure émergence, à partir de composantes dont aucune, prise isolément, ne contient un atome des qualités de ce sens » (*ibid.*).

[218] Cf. **supra, sect. 3.1.2. & 3.1.3.**

[219] **Chap. 1, sect. 1.3.1.2.**

3.2.3. Analyse énonciative des images

Comme nous l'avons suggéré, l'analyse sémiotique doit prendre en compte différents niveaux dans le parcours génératif de l'expression (Fontanille, 2008)[220]. Dans le même sens, d'autres travaux proposent de comprendre l'image en tant que résultat d'un acte d'énonciation, c'est-à-dire comme un énoncé visuel[221]. Parmi ces travaux, Maria Giulia Dondero (2015) affirme que, même si ces énoncés visuels sont observés et saisis au niveau des pratiques qui les mobilisent, il s'agit aussi de prendre acte de la dimension pratique dans l'expérience des images. En effet, l'énonciation visuelle est un domaine où l'expression prend forme par le moyen de supports qui sont saisis dans des *usages pratiques*. En mettant l'accent sur les pratiques de production et de réception des images, les sémioticiens cherchent à comprendre le visuel en fonction du *statut* (artistique, documentaire, scientifique, publicitaire, etc.) de ces images et des *contextes* où elles ont été produites, reçues ou utilisées, ou encore en fonction de la relation entre les images et leurs supports, ainsi que les formes et les genres de discours auxquels les images participent[222].

Concernant les formes de discours, nous avons remarqué que la *discursivisation* est considérée comme un mécanisme créateur de la personne, de l'espace et du temps de l'énonciation, en même temps qu'elle rend compte de la représentation actantielle, spatiale et temporelle de l'énoncé[223]. Lorsqu'ils se sont intéressés aux possibilités de discursivisation des catégories de la personne, de l'espace et du temps dans l'énonciation, les sémioticiens ont proposé la notion de *débrayage* pour définir l'opération « par laquelle l'instance d'énonciation se disjoint et projette hors d'elle, lors de l'acte de langage, et en

[220] **Chap. 1, sect. 1.3.2.**

[221] La sémiotique cherche à s'institutionnaliser autour d'axes de recherches et de paradigmes en commun où l'énonciation occupe une place importante. Au sujet des rapports actuels entre sémiotique et énonciation, voir notamment les contributions dans Colas-Blaise, Perrin & Tore (2016). Cet ouvrage propose un large éventail des applications de l'énonciation dans l'étude de langue, des discours, des textes et des images, en rassemblant divers spécialistes des sciences du langage (analyse du discours, linguistique, sémantique, sémiotique) qui proposent plusieurs points de vue sur l'énonciation.

[222] La prise en compte de certains usages peu étudiés des images a amené des travaux en sémiotique de l'énonciation à s'intéresser au rôle de l'image dans le discours scientifique (cf. Basso Fossali *et al.*, 2011 ou Dondero & Fontanille, 2012). De plus, comme pour l'histoire de l'art, notons que les recherches en sémiotique s'ouvrent aussi au dialogue interdisciplinaire avec la linguistique, la philosophie, l'esthétique, les sciences cognitives, la communication et les sciences sociales (cf. les contributions dans Basso Fossali *et al.*, 2010, ou plus récemment, Dondero, 2018a).

[223] **Chap. 1, sect. 1.3.3.**

vue de la manifestation, certains termes liés à sa structure de base pour constituer ainsi les éléments fondateurs de l'énoncé-discours » (Greimas & Courtés, *op. cit.* : 79)[224].

Cette réflexion sur le débrayage a notamment été reprise dans des travaux en sémiotique concernant des images (Fontanille, 1989 ; Dondero, 2016). Du point de vue sémiotique, étant donné que l'énoncé est compris comme une totalité signifiante et accomplie, cet énoncé peut être verbal, pictural, photographique, filmique, etc., mais le point commun est que ces différents types d'énoncés sont le résultat d'un acte énonciatif qui est observable à travers des marques énonciatives. En effet, un texte, une peinture, une photographie, un film, etc. sont caractérisés par le fait que l'énoncé a été produit « à partir d'un acte de projection qui lui a assuré des limites et des frontières » (par exemple : la couverture et la quatrième de couverture d'un roman, le cadre d'un tableau, les génériques d'un film ; cf. Dondero, 2016 : 245-246).

De plus, comme l'évolution des travaux en iconologie l'a montré, il s'agit aussi de prendre en compte la *matérialité de l'image*, puisque celle-ci est le support de l'énoncé visuel. En s'appuyant sur les travaux de Fontanille (2005 & 2008), Maria Giulia Dondero et Everardo Reyes (2016) ont montré que le niveau de l'objet-support de l'image correspond à celui d'une « structure d'accueil » qui permet, d'une part, d'accommoder l'objet en vue du processus interprétatif de l'actant opérateur, et d'autre part, cette structure permet d'attribuer un statut à des supports d'images et de les rattacher à un genre de pratiques énonciatives.

Dans *Image et vérité*, Jean-François Bordron remarque en effet que le propre de l'expérience d'images est d'ouvrir sur une scène énonciative par rapport à laquelle sont situés à la fois le producteur et l'observateur de l'image. Elle peut aussi être comprise *en tant que support (ou plan d'expression)*, mais un support qu'il sera toujours nécessaire de

[224] Le débrayage est par là même considéré comme un acte constitutif de l'énoncé, parce que cet acte vise à disjoindre, d'une part, les simulacres actantiels, spatiaux et temporels au sein de l'énoncé (l'énonciation énoncée), et d'autre part le sujet, l'espace et le temps qui sont propres à l'acte d'énonciation. Cette disjonction se réalise de deux manières, selon l'effet de sens recherché : soit par le moyen d'un « débrayage énonciatif » soit par le moyen d'un « débrayage énoncif ». Le débrayage énonciatif vise à installer un *je* (débrayage actantiel), un *ici* (débrayage spatial) ou un *maintenant* (débrayage temporel) dans l'énoncé et produit un « effet de subjectivité » (formes de l'énonciation énoncée). Tandis que le but du débrayage énoncif est d'instaurer dans l'énoncé un *il* (débrayage actantiel), un *ailleurs* (débrayage spatial) ou un *alors* (débrayage temporel), ce qui produit plutôt un « effet d'objectivité » du discours (formes de l'énoncé énoncé). Cet « effet d'objectivité » est souvent employé dans le discours scientifique, puisque celui-ci tend à éliminer les marques énonciatives dans l'énonciation et de faire uniquement reposer l'argumentation sur l'enchainement des énoncés.

lire, d'organiser en éléments signifiants. Enfin, l'expérience d'image ouvre sur l'*horizon de l'image*, la scène qui « disparaît au moment même où l'on découvre son existence » (cf. Bordron, 2013 : 54-55). L'auteur distingue trois « types idéaux » d'images à partir de ces trois scènes caractéristiques de l'expérience d'images, même si cette distinction n'est pas complètement fermée (car il n'est pas toujours facile de savoir « ce qu'il faut voir » par le moyen d'une image, cf. *ibid.* : 131-132). Chaque type met en évidence une manière de comprendre l'image[225] : (1) l'image événement met l'accent sur l'énonciation mise en œuvre par le moyen de l'image ; (2) l'image écriture met l'accent sur l'image en tant que plan d'expression, en tant que support à interpréter ; (3) l'image horizon se caractérise quant à elle « par l'importance accordée à ce dont il y a image, à ce qui est visé par elle » (*ibid.* : 55-64 & 131).

Les travaux de Dondero proposent également des considérations intéressantes sur la place de l'objet-support dans la relation énonciative qui s'établit entre l'image et l'observateur. Si, dans le discours verbal, les marques énonciatives concernent par exemple les pronoms personnels, dans les images, l'acte de projection de la personne est plutôt observable à travers le système des regards. Dans un travail de synthèse sur l'énonciation visuelle (cf. Dondero, 2016), l'auteure affirme que l'analyse énonciative des images a mis en évidence que le « point de vue » d'où l'on produit un énoncé visuel est toujours un « lieu de savoir », c'est-à-dire un lieu d'où des choses sont vues et à partir duquel le spectateur construit sa propre vision.

Denis Bertrand[226] a apporté des remarques éclairantes sur ce sujet. Le *point de vue* est défini comme un « ensemble de procédés utilisés par l'énonciateur pour sélectionner les objets de son discours et en orienter l'éclairage » ; dans les différentes formes de discours (descriptif, narratif, argumentatif), le point de vue concerne « le jeu des positions énonciatives (du débrayage à l'embrayage) » ou encore « les stratégies de structuration que déterminent les contraintes de la textualisation (ce qui est avant/ce qui est après, les relations entre les parties et le tout, le passage du particulier au général ou inversement, etc.) ». Bertrand renvoie en outre à deux autres notions, plus précises : (1) la focalisation,

[225] Nous avons en effet vu que, selon Bordron, les images s'intègrent dans un « style de visibilité », une « économie visuelle » qui permet de comprendre « ce qui est représenté » ou « ce qui est visé » par le moyen d'images ; cf. **chap. 1, sect. 1.5.**

[226] Voir les entrées « Point de vue », « Observateur » et « Focalisation » dans le glossaire, cf. http://denisbertrand.unblog.fr/glossaire-de-semiotique/m-p/ & http://denisbertrand.unblog.fr/glossaire-de-semiotique/f-l/ ; ce glossaire renvoie à Bertrand (2000).

à savoir la « procédure de débrayage cognitif qui détermine la position et le mode de présence du narrateur (ou de l'observateur) » ; (2) l'observateur, c'est-à-dire le « sujet cognitif installé par l'énonciateur, grâce au débrayage, chargé de recevoir l'information et de la transmettre ».

Dans cette perspective, la sémiotique a montré que l'analyse énonciative permet d'étudier « la manière dont le spectateur est censé se positionner face à l'image », puisque c'est justement l'image qui détermine ce positionnement ; comme dans la production des discours verbaux, la production d'une image prépare la « place » que l'observateur doit « occuper s'il veut reconstruire le parcours de la signification » (Dondero, 2016 : 249). Chaque « énoncé visuel » est porteur d'un questionnement sur le « savoir partagé » qui circule et qui est parfois disputé entre l'énonciateur et l'énonciataire. Le rapport entre énonciateur et énonciataire est le lieu où se joue ce savoir. L'énonciateur est caractérisé par une activité de production, tandis que l'énonciataire est caractérisé par une volonté de connaissance.

D'un autre côté, ces travaux en sémiotique de l'énonciation considèrent aussi la question de la réflexivité du visuel en prenant acte des innovations technologiques et médiatiques qui sont actuellement à l'origine de nouvelles formes d'images[227]. Si l'analyse énonciative met en valeur « le fait qu'il existe dans toute image des conflits de points de vue plus ou moins affichés », Dondero affirme plus récemment que cette analyse permet aussi « de relier l'image, en tant que produit fini, avec son parcours d'effectuation » et « avec les parcours perceptifs et interprétatifs offerts à l'observateur » (Dondero, 2020 : 196). Selon l'auteur, ce type d'analyse nous dirige vers une réflexion sur « la capacité de l'image à s'auto-analyser et à analyser d'autres images » (*ibid.* : 196). Comme le suggère l'auteure, l'image engendre pourtant elle-même un *langage explicatif* fait de formes, de structures et de schèmes qui émergent par le moyen d'un support (Dondero, 2020 : 112).

Cette prise de position est proche de celle d'autres travaux en histoire de l'art que nous avons présentés et qui défendent le visuel met en jeu une réflexivité qui lui est propre

[227] Voir par exemple l'ouvrage collectif *Les plis du visuel* (Dondero, Beyaert-Geslin & Moutat, 2017), qui rassemble plusieurs contributions concernant le problème de l'énonciation visuelle, problème auquel les auteurs apportent des solutions en suivant trois thématiques : (1) les pratiques médiatiques audiovisuelles et leurs relations avec l'énonciation, d'un point de vue théorique et formel, (2) la relation entre l'énonciation et la perception, (3) les caractéristiques de l'énonciation, en tant que métalangage.

(à savoir la logique, les mécanismes et les relations entre support et perception qui sont en jeu dans les images)[228]. À ce propos, Dondero affirme que la notion de « métalangage » est sans doute trop attachée au « langage » et à la « grammaire ». L'idée importante défendue par l'auteure est qu'il est possible d'identifier d'autres voies pour rendre compte de la dimension réflexive du visuel en dehors d'une conception uniquement grammaticale ou linguistique de la réflexivité[229]. Une image peut notamment participer à la construction d'une « activité autoréflexive » à partir de la « stratégie de composition de l'image », de la « spécificité médiatique de l'image », des « habitudes perceptives » ou encore du « cadre statutaire de l'image » (art, science, publicité, etc.)[230].

Puisque les stratégies de communication en ligne ont pour but de produire un support visible (la page-écran), et puisque cet espace est compris comme une scène d'énonciation, nous considèrerons que ce support et cette scène participent à la *construction du point de vue d'un observateur*. La question n'est pas de savoir si ce point de vue est réflexif, car nous avons suggéré que le support numérique peut être utilisé dans une perspective réflexive à condition de ne pas confondre la réflexivité informatique (le calcul et l'automatisme aveugle de l'ordinateur) avec la réflexivité et les compétences humaines. Il s'agit plutôt d'éclairer les modalités d'énonciation qui permettent de construire ce point de vue à l'aide d'un objet-support et d'une stratégie de communication.

---

[228] En effet, nous avons vu que la question de la réflexivité du visuel et des images est l'une des problématiques centrales dans plusieurs travaux récents ; cf. **supra, sect. 3.1.2.**

[229] Notons que la question du sens est indissociable de la réflexivité, qui est « une manière de définir et d'analyser ce qu'on appelle le "sens" » (Tore, 2013 : §5). Dans un travail sur ce sujet, Gian Maria Tore (*ibid.*) remarque notamment une tendance dans les sciences humaines et sociales à distinguer une dimension référentielle du langage (ce que la langue permet de figurer) de la dimension sui-référentielle du langage (ce que le langage produit), mais en soulignant qu'aucune de ces deux dimensions ne peut exister sans l'autre. Selon Tore (*ibid.* : §16), nous retrouvons également cette dynamique dans les pratiques qui utilisent des images, puisque toute image, en tant que fait sémiotique, est à la fois « référentielle » et « sui-référentielle », représentation et réflexive, mais elle n'est jamais exclusivement l'un ou l'autre. Il convient mieux de dire que représentation (transparence) et présentation (opacité) sont deux points de vue théoriques différents sur un même phénomène de sens ; cf. aussi **supra, sect. 3.2.1.**

[230] Dondero (cf. *ibid.* : 143-145) distingue quatre acceptions du « préfixe *méta-* en sémiotique selon les grandeurs discursives qu'il accompagne » : le *métatexte* (ou *méta-énoncé*) concerne la réflexion d'une image sur lui-même ; le *métalangage* concerne la réflexion d'une image sur sa « spécificité médiatique » et sur « la relation entre support et apport » ; la *métalangue* est la réflexion de l'image sur « les forces et les formes qui régissent le visuel selon le fonctionnement de la perception » ; le *métastatut* concerne la réflexion des images sur leurs statuts, voire sur les pratiques énonciatives qui les organisent.

3.2.4. *Computational semiotics*

Nous avons vu que les travaux en histoire de l'art font ressortir des usages possibles du traitement informatique d'images. Dans le même sens, nous présenterons rapidement quelques travaux en sémiotique qui envisagent aussi d'utiliser le traitement informatique pour analyser des images.

Au contraire d'autres approches en sciences humaines et sociales qui ont mis en évidence l'apport des *big data* (notamment à travers le contexte des humanités numériques) la sémiotique ne s'est pas de suite intéressée aux instruments informatiques d'analyse. Deux raisons ont favorisé cet investissement tardif :

- La sémiotique a longtemps été influencée par une tendance structuraliste qui proposait surtout des analyses d'objet culturel particulier, du moins des corpus réduits.
- Nous l'avons vu, cette tradition structuraliste a d'abord voulu appliquer aux autres pratiques médiatiques un cadre de réflexion et d'analyse qui correspond à l'étude des langues, des signes linguistiques, voire des discours et des textes.

Le cas des *big data* n'a pas été considéré par la sémiotique structuraliste, car il ne s'agit pas d'une configuration de signes ou d'un texte. L'analyse des *big data* correspond plutôt à un dispositif ouvert et dynamique (un point de rencontre) entre l'utilisateur et le traitement automatisé de données numériques[231].

Plus récemment, la sémiotique commence à considérer le potentiel analytique de l'informatique. La *computational semiotics* est définie comme « l'étude des phénomènes sémiotiques dans un contenu créé par l'homme – tous les types de contenu, du texte à l'œuvre d'art – à l'aide de modèles et de méthodes informatiques » (Chartier *et al*., 2018 : 2-3 ; nous traduisons de l'anglais). Pour ce faire, il s'agit de développer des modèles informatiques qui sont applicables à des ensembles de données dans le cadre de l'étude de la production du sens. Comme nous venons de le suggérer, ce type d'approche est encore rare dans les recherches en sémiotique, parce que certains sémioticiens rejettent

[231] Nous avons toutefois montré que plusieurs travaux s'inspirent de la sémiotique pour définir un cadre théorique et épistémologique concernant le support numérique et le patrimoine numérique ; cf. **première partie**. Nous avons remarqué que les réflexions de Tanaka-Ishii suggèrent que la théorie sémiotique peut être rapprochée de la programmation et de l'ingénierie des systèmes informatiques.

l'hypothèse selon laquelle les processus sémiotiques pourraient être modélisés de manière computationnelle (cf. Meunier 2014).

En suivant différentes voies, des travaux en sciences du langage ont pourtant montré que le traitement informatique est une étape importante dans l'évolution des méthodes d'analyses des langues (Rastier, Cavazza & Abeillé, 1994 ; Mehler, 2003 ; Léon, 2018). Or, ces travaux ont surtout mis en évidence les méthodes qui se basent sur *Semantic Vector Space* (SVS ; cf. notamment : Leopold 2005) ; méthodes qui permettent de découvrir des modèles abstraits à partir desquels il est possible d'inférer des structures de sens complexes comme la relation paradigmatique (Sahlgren 2006) ou l'étude des connotations (Neuman, Cohen & Assaf, 2015).

Les modèles SVS désignent soit une modélisation mathématique de comportements combinatoires au sein d'un large ensemble de données soit un ensemble de techniques basées sur des algorithmes qu'il s'agit d'appliquer sur ces données (cf. Schmidt, 2015). Ces techniques sont généralement utilisées dans l'analyse de langues ou de corpus de textes, mais il est également possible de les utiliser dans l'analyse d'un corpus d'images. Dans cette perspective, des chercheurs du *Laboratory of Cognitive Analysis of Information* (Université du Québec à Montréal) ont réalisés plusieurs expériences dont le but est d'explorer les possibilités d'une analyse sémiotique et computationnelle sur un corpus de peintures de Magritte (cf. Chartier *et al*., 2018). Le but de ces expériences est de considérer les apports des méthodes relatives au SVS dans les recherches en *computational semiotics* afin de montrer comment ces techniques pourraient être utilisées dans des analyses sémiotiques d'images.

Sans entrer trop dans le détail, le traitement informatique de descripteurs sémantiques relatifs aux images du corpus et l'emploi de techniques de *data visualization* permet de faire ressortir des thématiques récurrentes au sein d'un corpus d'images[232]. Pour les auteurs (cf. Chartier *et al*., 2018), ces expériences aboutissent sur des résultats prometteurs en ce qui concerne une sémiotique computationnelle à partir de modèles SVS. La question est toutefois de savoir si les techniques de visualisation sont des outils d'analyse efficace pour les recherches sur les images et le visuel, du moins un dispositif technique qui aide l'utilisateur à interpréter des ressources numériques.

[232] Nous expliquerons davantage ces techniques dans le **chap. 4**, **sect. 4.2.**

L'avantage du traitement informatique et des techniques de visualisation est sans aucun doute qu'ils sont applicables à de nombreuses images. Les analyses concernent donc de très larges corpus, ce qui n'est pas possible pour un historien ou un sémioticien. Cependant, nous avons souligné à plusieurs reprises qu'il n'existe pas de méthode d'analyse parfaite. En outre, nous verrons bientôt que ce type de traitement informatique accorde encore une fois une importance capitale à la description linguistique dans l'analyse d'images, même s'il faut remarquer que les outils informatiques d'images sont capables de mettre en évidence des relations entre les formes visuelles dans des images en fonction d'une quantification de la surface des images.

À ce propos, nous avons constaté que des travaux en histoire de l'art contribuent à utiliser davantage l'informatique comme une aide pour l'analyse ; il s'agit par exemple d'utiliser le traitement informatique d'images pour mettre en évidence des similitudes visuelles au sein d'un corpus d'images à l'aide d'un « dispositif de visualisation »[233]. Dans le même sens, d'autres travaux plus récents en sémiotique proposent des réflexions sur les méthodes d'analyse quantitative (Compagno, 2018 ; Reyes & Sonesson, 2019) ou encore sur l'intelligence artificielle appliquée aux images (Leone, 2021 ; Dondero, 2021). C'est dans la continuité de ces travaux que se situe le présent travail.

Avant de considérer les applications possibles de l'informatique dans l'analyse d'images, nous devons insister sur un point très important. Le traitement informatique d'images ne résout pas tous les problèmes qui concernent l'analyse d'images. Néanmoins, ce traitement augmente la « boîte à outils » que les historiens de l'art et les sémioticiens mobilisent dans leurs recherches. Le principal obstacle allant à l'encontre de l'utilisation du traitement informatique dans les recherches scientifiques n'est pas un obstacle technique. En effet, ces instruments fonctionnent, mais il faut encore comprendre les résultats qu'ils produisent et intégrer ces résultats dans une recherche. En outre, même si des dispositifs peuvent certainement aider les chercheurs, nous constaterons que (1) le cadre de réflexion sur le domaine du traitement informatique d'images est encore sommaire ; (2) ce domaine de recherche est surtout alimenté par des travaux récents en

[233] Pour Dondero (2020), ces dispositifs de visualisation constituent des processus d'analyse, mais une analyse qui doit être comprise comme le résultat d'opérations méréologiques (relations entre une totalité et ses parties), plus précisément des opérations de division, de regroupement, de superposition, etc.

informatique qui veulent montrer les performances de leurs algorithmes, parfois sans prendre en compte les besoins spécifiques des spécialistes des images[234].

3.3. Indexation d'images et limites du principe de métalangage de description

Afin de compléter les discussions précédentes sur le rôle de la description verbale dans la classification et l'analyse des images[235], la section suivante présentera quelques considérations sur l'indexation des images et de l'audiovisuel. Nous verrons que des travaux remarquent que cette indexation ne concerne pas les mêmes principes de *grammatisation* que dans le cadre de la documentation textuelle. Cependant, si l'indexation d'images qui se base sur des descriptions verbales est certainement limitée, une indexation uniquement à partir des caractéristiques visuelles est elle aussi problématique. Pour Bruno Bachimont, il faut rapporter la classification des archives d'images à un principe de « grammatisation des contenus » (cf. Bachimont, 2005 : 348-353)[236]. À ce propos, l'auteur souligne deux processus complémentaires qui sont à la base de l'indexation :

- Un processus de structuration qui isole certaines parties du contenu et qui décrit l'articulation entre ces parties.
- Un processus de qualification qui vise à qualifier ces parties du contenu par le biais de la langue naturelle.

L'indexation aboutit par conséquent sur un ensemble de représentations structurées en fonction d'une langue.

La structuration et la qualification des documents font l'objet de travaux sur des ontologies. Les « ontologies » sont au cœur des outils d'indexation en tant que « systèmes de référence », c'est-à-dire qu'il s'agit de « structures de concepts » sous la forme de mots-clés et qui concernent un domaine en lien avec les tâches et les activités de ce

[234] Nous nuancerons ces deux points dans la **quatrième partie**, notamment en présentant des projets dont le but est de proposer des instruments de visualisation pour les spécialistes des images et de la culture visuelle.

[235] Cf. **chap. 1** (**sect. 1.2.**) et **supra** (**sect. 3.1.1.** & **3.2.1.**).

[236] Voir aussi **chap. 1 (sect. 1.4.) & chap. 2 (sect. 2.3.)**.

domaine[237]. Si l'indexation des documents textuels se rapporte commodément à ces principes d'ontologies et de grammatisation des contenus, l'indexation des archives d'images et audiovisuels est plus difficile à mettre en place. D'un côté, les textes ont une structuration et une qualification qui sont étroitement liées aux modalités de l'expression linguistique (grammaire, syntaxe, vocabulaire, etc.) et de l'écriture (par exemple, alphabétique, mais pas nécessairement). De l'autre, les images et l'audiovisuel concernent des modalités d'inscription et de communication qui dépendent d'une structuration qu'on ne peut pas uniquement rattacher à des règles linguistiques.

Ces remarques de Bachimont sur l'indexation mettent en avant que « l'indexation multimédia est avant tout un problème conceptuel avant de renvoyer à des difficultés techniques » (Bachimont, 2007a : 170)[238]. Même si des travaux soulignent que l'indexation d'images peut aussi se baser sur les caractéristiques visuelles du document, dans la plupart des cas, les archives d'images sont donc organisées et classées à l'aide de métadonnées descriptives (voir par exemple Boulogne *et al*., 2006 ou Bachimont, 2007b). D'autres travaux plus récents affirment en outre que des ontologies, des systèmes de référence et des structures de concepts appropriés aux images et au visuel ne sont pas encore disponibles (voir notamment les contributions dans Dahlgren, Hansson, Reichert & Wasielewski, 2020). En attendant que ces ontologies soient disponibles, l'objectif de l'indexation multimédia est de rapporter un ensemble d'images à un cadre de notions linguistiques. Pour structurer un ensemble de ressources multimédias, il faut alors traduire ces ressources en concepts exprimés à l'aide d'une langue.

Cette problématique de la pertinence des ontologies dans l'indexation des archives d'images est amplifiée par un autre problème qui est mis en évidence par les recherches en *traductologie* (ou *translation studies*). Dans l'introduction d'un numéro de la revue *Hermès* sur ce sujet, Joanna Nowicki et Michaël Oustinoff remarquent en effet que les langues ne sont pas interchangeables, et donc qu'« une question clé pour la communication est celle qui consiste à déterminer dans quelle mesure les concepts sont tributaires des langues dans lesquelles ils se formulent et se traduisent » (Nowicki &

[237] Dans cette acception, une ontologie est considérée comme un ensemble structuré des termes et concepts propres à un champ d'informations ou d'un domaine de connaissances. On peut dire qu'une ontologie constitue un « modèle de données » qui est représentatif d'un ensemble de concepts et des relations entre ces concepts.

[238] Voir aussi Bachimont (2007b : 185-187).

Oustinoff, 2007 : 11). Ainsi, fonder l'indexation d'images sur des concepts propres à une langue contraint à comprendre ces images à travers une communauté discursive et des compétences sémiolinguistiques et socioculturelles qui donnent du sens à ces concepts ; or, d'une langue à l'autre, il ne s'agit plus de la même communauté. Concernant les activités de republication d'archives audiovisuelles dans un environnement multimédia, Peter Stockinger (2007 : 72) remarque que la traduction est une « activité spécifique de republication d'un document audiovisuel ou d'un corpus de documents audiovisuels » ; activité dont le but peut être d'aider à la compréhension linguistique de contenus audiovisuels, mais aussi de traduire complètement un contenu (d'une « langue source » à une « langue cible »). À la contrainte de réduire l'image à une description s'ajoute le problème de l'éventuelle traduction de cette description d'une langue à une autre[239]. Pour contourner cette problématique linguistique, l'usage de plus en plus généralisé de la langue anglaise vise à pallier ce problème en employant cette « langue internationale »[240]. En effet, nous verrons que certains projets institutionnels proposent une version en anglais de leurs sites Web, cela même si l'anglais n'est pas une langue officielle du pays[241]. Nous verrons également que deux projets proposent même de nombreuses versions linguistiques de la même interface Web[242].

L'utilisation d'un processus standardisé de description est pourtant largement employée pour encadrer et garantir l'intelligibilité des archives multimédias. À ce propos, d'autres travaux de Stockinger (2011 & 2012) ont montré que la sémiotique peut aider à conceptualiser des modèles de description pour analyser des archives audiovisuelles. En

[239] La traduction linguistique et l'adaptation cognitive font partie de la même problématique : « celle de la valeur ajoutée (de la pertinence, de l'intérêt, de l'utilité,…) d'un document ou d'un fonds de documents au sens d'une ressource (cognitive) potentielle pour un acteur et un contexte d'usage donné » (Stockinger, 2007 : 73). La plupart des projets de republication d'archives audiovisuelles sont confrontés à un même défi conceptuel et technique, mais aussi socioéconomique : « comment faire circuler des documents monolingues à profils fortement marqués (généralement des documents scientifiques) dans un "marché" intrinsèquement multilingue et multiculturel tout en devant supporter la concurrence de documents réalisés en langue anglaise » (Stockinger, 2007 : 73-74). C'est sans doute pourquoi les projets institutionnels (surtout ceux qui proviennent de pays plurilingues) apportent un soin particulier à traduire les informations en plusieurs langues, même si d'autres projets ne semblent pas considérer cet obstacle.

[240] Voir notamment d'Angelo (2019) : « Le développement de l'usage de l'anglais correspond au bénéfice d'une langue véhiculaire globale, y compris pour le petit groupe des langues utilisées à l'international (allemand, arabe, espagnol, français, portugais, russe). On peut toutefois s'interroger de savoir jusqu'à quel point la croissance de l'usage de la langue centrale/globale sera possible » (d'Angelo, 2019 : 75).

[241] **Chap. 5**.

[242] **Chap. 6**.

effet, cette discipline semble tout indiquée pour spécifier le « métalangage approprié », c'est-à-dire l'*ontologie* à utiliser « pendant le traitement, la description, l'indexation, le commentaire, la version/traduction, l'interconnexion, des données des médias composant les fonds d'une archive » (cf. Stockinger, 2015 : 80-81)[243]. Un métalangage description se base sur modèle de description « pour identifier et décrire les parties pertinentes de son corpus audiovisuel », c'est-à-dire « les segments ou passages d'un document audiovisuel qui apportent des informations concrètes à un modèle de description donnée » (Stockinger, 2011 : 110-111). Pour l'auteur, un métalangage de description permet de distinguer différents « types de modèles de description » (p. ex. : métadescription, éléments paratextuels, analyse du contenu, mise en scène, adaptation linguistique du contenu ; cf. Stockinger, 2012 : 351) qu'il s'agit d'appliquer à des archives audiovisuelles, afin de faciliter l'interaction de l'usager avec les ressources numériques.

Ces dernières considérations sur le rôle du métalangage de description dans l'indexation des images font ressortir un point important. Les archives d'images sont décrites et organisées sur la base du langage verbal, ce qui indique notamment une volonté d'encadrement des archives d'images dans une communauté discursive[244]. Mais, en retour, cet encadrement favorise uniquement une classification des images à partir d'un *métalangage descriptif*, autrement dit, un modèle formalisé qui vise à rendre compte d'une langue ou d'un autre langage. Pourtant, cette notion est problématique parce qu'elle s'inscrit dans un cadre limité, celui d'une compréhension logico-grammaticale du sens[245]. Toutefois, si les outils d'indexation d'images favorisent une classification en fonction de métadonnées descriptives et non pas en fonction de formes visuelles, c'est sans doute parce que les images ont d'abord été analysées dans un cadre théorique qui s'appuie sur une *conception transitive* (paradigme de la représentation).

---

[243] Il existe, en outre, plusieurs applications de la sémiotique dans le domaine des archives et bibliothèques numériques (audiovisuelles), notamment comme cadre méthodologique pour un projet d'archives numériques ; plus précisément : « pour la création de données multimédias numériques à l'aide de modèles de publication/réédition (spécifiques à l'archiviste) » ou pour définir « la portée d'un projet d'archivage (son domaine et son contexte d'utilisation) » (*ibid.*).

[244] **Chap. 1 (sect. 1.3.) & chap. 2 (sect. 2.3.)**.

[245] Le métalangage « ne concerne que les langues naturelles, considérées comme des langues-objets, et le métalangage dont il s'agit est plus au moins coextensif à la grammaire (ou à la théorie grammaticale) » (Greimas & Courtés, *op. cit.* : 255 ; entrée « Métalangage »). Pour un point de vue plus récent sur le métalangage et la sémiotique, voir aussi les contributions dans Basso Fossali *et al.* (2013).

C'est seulement dans un deuxième temps que des travaux en iconologie et en sémiotique visuelle se sont éloignés de cette conception. En effet, nous avons vu que le visuel s'appuie sur sa propre logique, à savoir sur une dynamique des formes visuelles et des forces qui sont en jeu dans l'image. Plusieurs recherches plus récentes en iconologie et en sémiotique que nous avons présentées plaident en faveur d'une approche plurielle du visuel qui s'éloigne d'une interprétation axée sur cette compréhension logico-grammaticale des images[246]. Depuis les années 1980, des travaux en sémiotique visuelle ont proposé une approche pour prendre en compte les propriétés formelles et plastiques dans le processus d'analyse et d'interprétation des images. D'un côté, la sémiotique figurative propose d'interpréter l'image à travers une grille de lecture sémantique qui permet de dégager des formants figuratifs. De l'autre, la sémiotique plastique concerne l'analyse de la surface de l'image, en tant que plan d'expression où la dimension plastique est définie en fonction de plusieurs catégories (topologique, eidétique, chromatique, microtopographique). Nous avons également vu que d'autres travaux en sémiotique ont par la suite proposé de considérer l'énonciation visuelle dans le cadre d'une approche rhétorique des signes visuels.

Somme toute, l'application stricte d'une grille de lecture sémantique dans l'indexation d'images met de côté la complexité de la dimension plastique, c'est-à-dire la surface de l'image et la composition des formes et des couleurs sur cette surface. C'est pourquoi Dondero remarque qu'il est nécessaire d'apporter quelques aménagements à la notion de métalangage lorsqu'il s'agit d'images[247], car il faut mettre en évidence l'autonomie réflexive du visuel par rapport à la réflexivité limitée du langage. Dans cette perspective, l'analyse énonciative montre comment les images, en tant qu'énoncés visuels, préparent une place pour le spectateur, qui doit de se frayer un parcours de lecture qui s'appuie sur des agencements de formes et de couleurs, sur un support et les limites de son cadre, sur un point de vue, sur le statut de l'image, etc. Les images sont alors comprises (1) comme le résultat d'une pratique énonciative, où l'image est considérée comme le lieu de production d'un savoir, tandis que l'utilisateur est caractérisé par une volonté de savoir,

[246] Cf. **supra, sect. 3.1.2. & 3.2.3.**

[247] Cf. **supra**, **sect. 3.2.3.**

et (2) dans une dynamique réflexive qui les rattache les unes aux autres, mais aussi une dynamique qui les rattache à des formes perceptives, à des supports, à des statuts.

Dans les sections précédentes, nous avons vu que les historiens de l'art et les sémioticiens envisagent notamment d'utiliser l'informatique pour étudier et pour comparer des formes visuelles présentes dans des corpus d'images[248]. En effet, des progrès récents en informatique indiquent que des instruments permettent d'analyser une grande quantité de données relatives à la culture visuelle en fonction des caractéristiques plastiques des images, ce qui offre des perspectives intéressantes aux chercheurs en ce qui concerne la classification automatique d'un corpus d'images[249]. Cependant, selon les spécialistes de l'indexation multimédia, le remplacement des outils d'indexation d'images par des instruments classant automatiquement des images doit être considéré avec prudence et rigueur (cf. Gros, 2005 : 383-391).

Pour terminer ce **troisième chapitre**, nous remarquerons que le projet d'une étude du visuel à travers les similitudes plastiques des images les unes par rapport aux autres est également une question centrale dans d'autres travaux en sciences humaines et sociales. Dans la section suivante, nous verrons que ce projet trouve son origine dans d'anciens travaux en histoire de l'art partageant la conviction qu'il faut étudier les formes et de figures visuelles dans une perspective comparative. En parallèle, nous présenterons aussi des approches en sociologie et en psychologie de l'art qui partagent cette conviction.

3.4. Revenir à l'analyse des formes visuelles

En plus des approches en sémiotique visuelle, nous avons constaté un regain d'intérêt pour la dimension plastique et pour le principe de comparaison des formes visuelles dans des travaux en histoire et esthétique de l'art. En effet, si nous avons souligné que l'iconologie accorde un rôle important aux sources, aux contextes de production, et surtout à la description verbale des œuvres d'art visuel, des travaux plus récents en iconologie insistent aussi sur le projet de l'*Atlas Mnémosyne* d'Aby Warburg. Ce dernier aborde le problème de l'interprétation des arts visuels en insistant sur l'idée de résurgence de certaines *formes* (ou *forces*)[250]. Pourtant comment s'agit-il d'identifier les relations

[248] Cf. **supra**, **sect. 3.1.4. & 3.2.4.**

[249] **Chap. 4.**

[250] Cf. **supra**, **sect. 3.1.2.**

entre ces formes ? Cette question rejoint sans doute d'autres considérations du début du XX^e siècle en histoire de l'art sur l'évolution des formes. En effet, certains travaux affirmaient déjà à cette époque que les formes et les styles artistiques s'influencent et dynamisent à long terme les productions culturelles. En France, nous pouvons mentionner les travaux d'Élie Faure (1992a & 1992b [1927] et ceux d'Henri Focillon (2010 [1934]). En Allemagne, ceux de Warburg sur la Renaissance (Warburg, 2003) et ceux d'Heinrich Wölfflin (1997 [1943] & 2017 [1915]) sur l'art classique et sur l'art baroque.

Dans *La vie des formes* (Focillon, *op. cit.*), Focillon propose qu'un *style* se constitue à travers des formes qui ont pour l'historien une « valeur d'indice » (il faut comprendre que ces formes sont des indicateurs qui fournissent des informations pour rattacher l'utilisation de certaines formes à un style). Pour ce faire, l'étude des représentations figurées d'un thème ou sujet dans les arts visuels (ou « iconographie ») est définie par l'auteur « soit comme *la variation des formes sur le même sens*, soit comme *la variation des sens sur la même forme* » (Focillon, *ibid.* : 10). Tandis qu'un style défini est, pour Focillon, un « milieu formel homogène, cohérent, à l'intérieur duquel l'homme agit et respire, milieu qui est capable de se déplacer en bloc » ; le style est un ensemble de relations entre des éléments formels, même si ces « formes ne sont pas leur propre schéma, leur représentation dépouillée » (*ibid.* : 24-25). Ces formes sont définies dans leurs rapports dynamiques à l'espace, à la matière, à l'esprit et au temps, tandis que ces styles doivent être mis en relation avec une culture et des activités humaines : la *vie des formes* « s'exerce dans un espace qui n'est pas le cadre abstrait de la géométrie ; elle prend corps dans la matière, par les outils, aux mains des hommes » (*ibid.* : 25). Les formes (qui sont les indices d'un style) ne sont donc pas seulement des relations de structures entre des formes géométriques, mais aussi des significations qui sont déjà investies d'un sens anthropologique.

Focillon souligne en outre que l'étude des œuvres d'art visuel montre que les *formes plastiques* sont « soumises au principe des métamorphoses, qui les renouvelle perpétuellement, et au principe des styles qui, par une progression inégale, tend successivement à éprouver, à fixer et à défaire leurs rapports » (*ibid.* : 10). En tant que répertoire ou vocabulaire d'un style artistique, l'étude de ces formes doit par conséquent être rapportée à d'autres exemples de formes du passé, même si souvent les nouvelles figures paraissent « sans rapport avec leur origine » (*ibid.* : 8).

Ailleurs, dans *L'Esprit des formes*, Élie Faure affirme quant à lui que la « forme universelle est bâtie sur un plan unique » et que l'évolution des styles progresse dans un cadre anthropologique, car la « tradition, s'il le faut par la révolution, a pour objet de maintenir ou de rétablir le style dans les formes expressives d'une époque, d'un peuple, d'une fonction, d'un homme » (Faure, 1992 [1927] : 114 & 163).

Dans *Style, artiste et société*, Meyer Schapiro remarque justement que « les spécialistes ont rapidement remarqué que les traits qui constituent un style ont une qualité en commun. Ils ont tous l'air d'être marqués par l'expression de l'ensemble, ou bien il existe un trait dominant auquel les éléments ont été adaptés » (Schapiro, 1982 : 45). Pour cet auteur, la description d'un style se rapporte surtout à trois « aspects de l'art » : (1) éléments formels (motifs), (2) les relations formelles et (3) les qualités (*ibid.* : 39). Schapiro précise :

> Quand on relie des œuvres d'art à un individu ou à une culture, ce sont ces trois aspects qui donnent le critère de jugement le plus large, le plus stable et, donc, le plus sûr, qui s'accorde le mieux à la théorie moderne de l'art, bien qu'à des degrés divers selon les points de vue. La technique, le sujet et le matériau peuvent caractériser un certain groupe d'œuvres et on les inclut parfois dans les définitions ; mais, en général, ces éléments sont moins propres à l'art d'une certaine période que les traits formels ou qualitatifs. On peut facilement imaginer un changement marqué dans le matériau, la technique ou le sujet qui ne soit pas accompagné d'un changement notable dans la forme fondamentale. […] Quand une technique coïncide effectivement avec le développement d'un style, plus que le travail lui-même, ce sont les traces laissées par la technique sur la forme qui comptent pour décrire le style. Les matériaux ont une signification surtout en ce qui concerne la qualité de la texture et la couleur, bien qu'ils puissent affecter la conception des formes. En ce qui concerne le sujet, on constate que des thèmes assez différents – portraits, natures mortes et paysages – peuvent apparaître à l'intérieur d'un même style (Schapiro, 1982 : 39).

Pourtant, Schapiro souligne que « les éléments formels ou motifs, bien qu'ils soient très frappants, essentiels même, en ce qui concerne l'expression, ne suffisent pas à caractériser un style » (*ibid.* : 40). Nous l'avons vu, Panofsky a lui aussi montré que l'histoire de l'art cherche à faire ressortir cette évolution des *motifs* et des *éléments*

formels. L'analyse des formes est au cœur du travail de l'interprète des arts visuels[251]. Dans l'un des textes que nous avons déjà évoqués, Panofsky (1975 : 195) affirme à peu près comme Schapiro que la « forme » est toujours complémentaire d'un *style*. D'un côté, l'artiste est à la recherche de son « mode d'expression » par le moyen d'un « modelé qui n'appartient qu'à lui et à une conception et à une définition de l'objet qui lui sont propres » (*ibid.*). Cette expressivité est réalisée à l'aide de « formes générales » qui sont liées à une époque. De l'autre, l'histoire de l'art doit fournir une explication à propos de ce « mode d'expression » qui s'appuie sur des formes plus générales.

Nous retrouvons ce rapprochement entre forme et étude des arts visuels en dehors de l'histoire de l'art. Dans *La Figure et le Lieu*, le sociologue de l'art Pierre Francastel (1967) propose que le *domaine des figures* est en fait « un fil conducteur imaginaire » pour se guider à travers l'évolution des arts figuratifs, parce que l'ordre figuratif suggère toujours au spectateur « l'existence d'un lieu, qui n'évoque pas nécessairement une configuration réalisée, ni même réalisable, mais une simple relation entre des signes fragmentaires » (Francastel, 1967 : 350-351 ; à ce sujet, voir aussi les propositions de Belting, 2004 : 77 *sq.*)[252]. Francastel (1967 : 342-343) distingue plus précisément trois niveaux de compréhension de la *figure* :

- Au niveau des signes[253] « qui matérialisent la sélection intentionnelle des éléments simples de reconstitution du visible ».
- Au niveau des structures « qui déterminent la disposition de ces éléments à l'intérieur du champ figuratif ».
- Au niveau des « chaines d'idées » ou des « ordres » qui déterminent « des principes généraux suivant lesquels se lient et se combinent les signes sélectifs et les structures associatives et stabilisatrice de l'expérience » (*ibid.* : 354).

Des dynamiques – donc, des mouvements de forces et de structures – sont comme dissimulées derrière cette notion de figure, qu'il ne faut pas seulement réduire à des signes

[251] Cf. **supra**, **sect. 3.1.1.** Pour le point de vue de Panofsky sur la notion de forme en théorie de l'art, on consultera aussi Panofsky (1975 : 37-182 & 235-255 ; 1989 : 17-23).

[252] Le but de l'auteur est de défendre, dans une perspective sociologique, l'idée que les arts figuratifs constituent « un système de signification irréductible à tout autre », mais aussi l'idée que ce système se plie pourtant « aux mêmes lois d'insertion dans la réalité et dans la pensée que tous les autres codes constitués d'une manière permanente, parce qu'ils matérialisent une fonction, c'est-à-dire une conduite, un mode d'être de l'homme » (Francastel, 1967 : 355).

[253] Nous pouvons remplacer « signes » par « ensembles signifiants », cf. **supra, sect. 3.2.2.**

linguistiques. Il faut rapporter la compréhension de ces figures à des pratiques socioculturelles d'où elles émergent, en prenant en compte les relations complexes qui se tissent entre ces « structures dynamiques », mais aussi l'équipement culturel du spectateur des arts figuratifs. De ce point de vue sociologique, l'interprétation des figures visuelles est considérée selon trois niveaux (subjectif, significatif, anthropologique) qui rappellent sans doute l'approche iconologique. Or il faut remarquer que ces considérations sur les figures conviennent surtout à un contexte historique et à un certain type d'images, notamment la peinture européenne du XV^e^ siècle et l'ordre visuel qui en découle[254].

D'un autre côté, des approches en psychologie se sont intéressées à cette relation entre forme, image et perception visuelle. L'étude des images et du visuel a notamment été rapprochée d'une conception de la forme qui est inspirée des travaux en *Gestaltpsychologie* (« psychologie de la forme »)[255]. Selon cette approche, « la réalité est toujours déjà composée, agencée suivant des *formes*, qui sont ce par quoi nous saisissons toujours les apparences en premier lieu », Hansen-Løve (2000 : 183 ; entrée « Gestalt ») ; cette conception se fonde vraisemblablement sur un point de vue théorique qui associe la notion de forme à des « structures apparentes » ou au principe d'« arrangement logique »[256]. Sans trop entrer dans le détail, la psychologie de la forme a permis de mettre en évidence que la perception visuelle complète automatiquement certaines compositions formelles ; par exemple : un carré constitué de points et non de quatre lignes sera tout de même perçu comme un carré. Ainsi, il y aurait des structures visuelles que la perception reconnaît comme des modèles d'images génériques.

Cependant, en s'appuyant sur la psychologie de la forme, Rudolph Arnheim (1976) affirme dans *La pensée visuelle* qu'il ne faut pas systématiquement confondre la « signification des formes visuelles dans les œuvres d'art visuel » avec la « signification

[254] Sur l'art figuratif qui est issu de la peinture du *Quattrocento*, voir aussi les travaux de Daniel Arasse (2008 [2019] & 2009).

[255] Sur la psychologie de la forme, cf. entre autres Kölher (2000 [1929]) et Guillaume (1992 [1937]).

[256] En logique, la « forme » désigne la « structure d'une proposition ou d'un raisonnement, indépendante de son contenu de réalité […] » ; cf. le dictionnaire de Laurence Hansen-Løve (2000 : 170, entrée « Forme ») : (1) « Caractères apparents des choses » ; (2) « Arrangement, façon dont une pensée est présente » ; (3) « Caractéristiques d'une institution politique ou culturelle ». Bien entendu, la forme n'est pas uniquement considérée par des approches concernant les images et le visuel, car cette notion est également discutée dans de nombreuses disciplines scientifiques ; voir par exemple les contributions dans Aa. Vv. (1994).

des formes visuelles perçues ». Pour l'auteur, les « *modèles formels* que suggèrent un paysage ou une nature morte ne se rapportent que très indirectement, quand ils sont pris dans leur unicité, aux *formes* et aux *significations standards* d'arbres, de fermes, d'artichauts ou de poissons » (*ibid.* : 312 ; nous soulignons). Ces formes et ces significations standards des objets perçus sont établies sur la base d'une expérience subjective, tandis que la valeur des modèles formels qui ont été mis en évidence par la psychologie concerne « des modèles de force beaucoup plus génériques » (*ibid.* : 312).

Nous avons déjà présenté le point de vue de la sémiotique à ce sujet : la question de la *figuration picturale* est le domaine de l'*effet-représentation* et de la sémiotique figurative qui, rappelons-le, consiste à appliquer une « grille de lecture sémantique » aux images afin de distinguer des formants figuratifs dans le plan de l'expression de l'image. De plus, les travaux en sémiotique visuelle nous ont permis de faire ressortir que la *dimension plastique* des peintures est caractérisée par la *disposition* des *formes* et des *couleurs* sur la surface-support de l'image. Dans un cas comme dans l'autre, tout se passe comme si l'expérience des images se rattachait nécessairement à un questionnement sur la « perception du monde »[257] et au principe de « composition de formes visuelles »[258] qui constituent cette surface-support.

Certes, il ne s'agit pas de trancher ici entre les travaux sur les formes visuelles en histoire de l'art, en sociologie, en sémiotique ou encore en psychologie, mais plutôt de remarquer que cet intérêt pour la dimension plastique de l'image met en avant la logique des « structures », des « formants plastiques » ou des « modèles de forces » qui sont éventuellement susceptibles de constituer un « champ figuratif », des « formants figuratifs », des « effets de représentation » et des « effets de sens ».

Nous constatons donc qu'il est courant, au sein de différentes disciplines des sciences humaines et sociales, de fonder l'étude des œuvres d'art sur l'analyse des formes

[257] À ce propos, cf. les réflexions de Maurice Merleau-Ponty (1960) sur certains tableaux de Paul Cézanne ; pour Merleau-Ponty, la peinture a été l'occasion de redécouvrir l'opacité de l'image et l'illusion de sa profondeur, notamment à travers l'ambigüité de la relation de l'image avec le monde perçu.

[258] Cette idée est déjà présente dans l'un des plus anciens traités sur la peinture, au XV^e^ siècle : « La *composition* est dans la peinture le procédé par lequel les parties sont disposées dans l'œuvre de peinture. Le plus grand mérite du peintre n'est pas de faire un colosse mais une histoire. Et son talent aura plus de mérite dans une histoire que dans un colosse. Les parties de l'histoire sont les corps, la partie du corps est le membre, et la partie du membre est la surface. Les premières parties d'un ouvrage sont donc les surfaces, parce que d'elles sont faits les membres, des membres les corps et des corps l'histoire qui constitue le dernier degré d'achèvement de l'œuvre du peintre », cf. Alberti (2014 [1435] : 159).

visuelles qui les composent. Un passage de *Fables, formes, figures* (Chastel, 1978) résume les différents points de vue que nous avons présentés dans les sections et paragraphes précédents. Dans cet ouvrage, André Chastel affirme comme d'autres historiens de l'art que, s'il est vrai que l'activité artistique et les approches théoriques se métamorphosent, que les styles des œuvres se transforment au fur et à mesure, ce dynamise de l'activité artistique est toutefois indissociable de ce que Chastel nomme la *Fable*, à savoir l'au-delà imaginatif de l'œuvre d'art. L'auteur ajoute que l'activité artistique s'exerce « dans l'espace qu'elle enrichit d'objets matériels à travers des procédures et en vue d'effets, que ramasse la notion de *Forme* », et que cette activité est « productrice de *Figures* aux multiples résonances psychologiques et morales » (Chastel, 1978 : 28). Ainsi, pour passer de l'état potentiel à un état réalisé, la *forme* est l'actualisation de structures dans l'espace et la matière. En effet, pour l'auteur, la forme est avant tout un moyen de coordination, d'ajustement et d'articulation, mais cette forme n'est « signifiante que dans une composition » où des *figures* deviennent une « présence qui efface [la Fable] et [la Forme], au moment où elle bénéficie de notre regard » (*ibid.* : 38). Dans une certaine mesure, le propos de Chastel préfigure des considérations récentes sur l'évolution de l'iconologie en une science des images qui, nous l'avons vu, défend un paradigme de la présence des images[259]. Cependant, nous ne mettons pas seulement en évidence ce propos pour illustrer cette correspondance, mais plutôt pour indiquer une remarque de l'auteur que nous approfondirons dans la suite de ce travail.

Au regard des approches que nous avons présentées dans ce **troisième chapitre**, la forme est comprise comme une *composition*, ou selon Chastel, comme « un "schème-directeur" (mais "*pattern*" est décidément mieux) qui a le pouvoir d'appeler et de coordonner les motifs » (Chastel, 1978 : 37). Pour dégager un contexte historique qui permet d'expliquer une œuvre, nous l'avons remarqué, la solution proposée par l'iconologie ou la sociologie de l'art est d'étudier le contexte de production de cette œuvre pour mettre au jour le sens des « *patterns* » qui traverse cette œuvre. L'obstacle que rencontrent toutefois l'histoire et la sociologie de l'art est d'essayer « de rendre compte des propriétés que chaque époque et chaque milieu accordent aux formes et aux figures » (*ibid.* : 46). Certes, durant les années 1960-1980, il était difficile de détecter des

[259] Cf. **supra, sect. 3.1.2.**

modèles formels dans un ensemble d'images, parce que ces images étaient dispersées dans différents musées de par le monde.

Plus récemment, la conversion numérique et la diffusion sur Internet des images et l'évolution des instruments pour traiter informatiquement des images indiquent que cette difficulté de la dispersion des images pourra à l'avenir être palliée. En effet, nous verrons que cette conception de la forme proposée par Chastel (c'est-à-dire la forme comprise comme « un pattern qui coordonne les motifs ») n'est pas incompatible avec le point de vue de l'informatique sur l'analyse d'images[260].

Pour finir, remarquons que lorsque Chastel explique comment il faut entendre les expressions « Fable », « Forme » et « Figures » par rapport aux orientations et paradigmes scientifiques de son époque : « la Fable est l'arrière-plan du signifié, le référent par excellence ; la Forme, manifestation de la structure ; la Figure, instance métapsychologique de l'image »[261] (*ibid.* : 45). Pour Chastel, à la suite de l'iconologie, « l'ère de la sémiologie » impose de considérer les œuvres d'art et les images « en termes de signes et de sens, c'est-à-dire interprété selon le modèle linguistique » (*ibid.*).

Nous l'avons vu, des travaux en sciences du langage sur l'image occupent une place importante dans les théories des images et du visuel qui sont contemporaines des écrits de Chastel. Dans la suite du présent travail, nous montrerons que tout indique que cette « ère de la sémiologie » est actuellement supplantée par « l'ère du numérique ».

3.5. Conclusions

Ce **troisième chapitre** nous a permis de mettre en avant une problématique concernant les outils d'indexation employés par les institutions patrimoniales. En effet, l'archivage d'image s'appuie sur un « paradigme » d'analyse qui est influencé par les travaux en histoire de l'art et en sciences de langage. Ce paradigme est influencé par les conceptions défendues par les premiers travaux en iconologie et en sémiologie des images, travaux que nous avons rapidement présentés à travers plusieurs auteurs reconnus dans ces disciplines. L'évolution de ces approches affirme toutefois une ouverture à d'autres

[260] **Chap. 4**.

[261] Le référent et les structures de l'œuvre s'effacent derrière cette notion ambiguë de figure en tant qu'instance métapsychologique (le rapprochement de l'histoire de l'art et de la théorie psychanalytique est considéré dans Marin, 1971, ou encore dans Didi-Huberman, 1990).

disciplines scientifiques en ce qui concerne l'interprétation des images et du visuel. Néanmoins, nous avons constaté que ces approches ont d'abord accordé une place centrale à l'analyse du contexte de production et surtout à la description verbale des images.

Nous avons également remarqué que cette problématique est mise en évidence par des travaux sur le métalangage de description des images. L'indexation des images a d'abord favorisé la réduction des archives d'images à des termes linguistiques, ce qui a permis de développer des instruments de recherche par mots-clés dans un ensemble d'images. Toutefois, ce modèle d'indexation contraint à ne pas prendre en compte la dimension plastique et la réflexivité propre au visuel (réflexivité qui est placée au centre des recherches plus récentes sur le visuel et les images). Cette problématique risque de s'intensifier, car nous remarquerons que ces métadonnées descriptives sont aussi fondamentales dans les processus d'entraînement des ordinateurs pour reconnaître des objets dans les images ou pour annoter automatiquement des images[262].

Pourtant, ce paradigme (qui propose de « contrôler » les images et le visuel par le biais d'un langage de description) a été remise en question par plusieurs travaux sur la culture visuelle et sur l'analyse d'images. Même s'ils sont issus de différents courants disciplinaires, nous avons vu que ces travaux défendent pour la plupart un décloisonnement des théories et des méthodes d'analyse d'images, notamment : (1) en affirmant que le visuel met en jeu une réflexivité qui lui est propre et (2) en envisageant l'étude contrastée de corpus d'images pour faire ressortir des mécanismes inhérents au visuel à travers différentes productions visuelles.

Nous avons suggéré que des travaux en histoire de l'art et en sémiotique envisagent d'utiliser des instruments informatiques pour analyser des corpus d'images, même si ces instruments se fondent sur un cadre théorique et méthodologique assez différent des approches en sciences humaines et sociales. En effet, les progrès dans des domaines spécifiques de l'informatique ont permis de développer des programmes informatiques capables d'analyser rapidement une image, mais surtout de reproduire le processus sur de nombreuses autres images. Ces systèmes fonctionnent à l'aide de langages de programmation et de logiciels qui nécessitent des considérations techniques pour bien comprendre les opérations d'analyse de ces dispositifs informatiques.

[262] Voir **chap. 4, sect. 4.2.3.**

Dans le **chapitre suivant**, nous verrons que l'un des avantages de ces progrès en informatique est qu'ils offrent des instruments pour traiter des images en grand nombre, à partir de description verbale, mais aussi en fonction des éléments plastiques (formes, couleurs) des images. Avant d'être analysables et analysées, les images sont définies sous la forme de nombres, parce que les ordinateurs ne traitent pas les images telles que l'humain les perçoit. Même si le numérique influence les pratiques de manipulation et d'interprétation des images parce que le support imperceptible réalise des opérations computationnelles sur des données qui ne sont pas réalisables par un humain, il faut garder à l'esprit que la visée du traitement informatique d'images est d'être utilisé par un humain. Nous constaterons surtout un intérêt envers des opérations analytiques qui visent à classer automatiquement des images et à les présenter au sein d'un espace de visualisation (une interface) qui fait ressortir des relations sémantiques, ou encore des relations de similarités formelles et plastiques entre ces images.

CHAPITRE 4
Images numériques, nouveaux médias et analyse informatique d'images

Dans les **chapitres précédents**, nous avons suggéré à plusieurs reprises que des travaux proposent d'interroger les exigences des sciences humaines à l'égard de la technologie numérique et des outils informatiques. Dans ce **quatrième chapitre**, nous présenterons des travaux sur l'image numérique et sur l'analyse informatique d'images. Ces « méthodes d'analyse assistées par un ordinateur » (*computer-assisted methods*) s'appuient principalement sur la quantification complète de la surface des images (**sect. 4.1.1.**). Nous verrons que les performances analytiques des ordinateurs sont de plus en plus efficaces, à tel point que les informaticiens envisagent d'utiliser la programmation informatique pour mettre en œuvre des instruments automatisés pour classer des images en fonction de certaines caractéristiques plastiques (**sect. 4.1.3.**). Lorsque nous expliquerons brièvement le fonctionnement de ce type d'analyse quantitative, nous verrons que trois domaines ont été décisifs : la *computer vision*, la *data vizualisation* et le *deep learning* (**sect. 4.2.**). Avant de présenter ces instruments et ces domaines informatiques, nous présenterons rapidement quelques travaux théoriques dans le but de mieux comprendre les points communs et les différences entre les images qui sont habituellement étudiées par les spécialistes des arts visuels et de la culture visuelle (peintures, photographies, cinéma, etc.) et les images numériques (**sect. 4.1.2.**).

4.1. De l'analyse d'images à l'analyse d'images numériques : quelle différence ?

Dans un entretien avec Émile Noël sur la « théorie de la morphogénèse »[263], le mathématicien René Thom suggère une remarque intéressante sur la conversion numérique des images :

> On découpe la surface du tableau contenant la forme en petits carrés que l'on appelle des pixels et ensuite on note tous les petits carrés qui rencontrent la forme. De plus, on code éventuellement la couleur, la nuance, la brillance, etc. Avec ces notations, on peut reconstituer une forme qui n'est plus la forme initiale, mais qui en est une approximation assez fine pour que l'impression subjective soit pratiquement la même (Thom & Noël, 1994 : 180).

Convertir un tableau en image numérique, c'est découper la surface du tableau (ou de l'image) en unités élémentaires et transformer ces unités selon un « code », des « notations ». En effet, un ordinateur se base sur des procédés et des opérations qui traitent uniquement des valeurs numériques correspondant à un ensemble de « petits carrés », les « pixels » (*picture elements*). Ce « découpage » et ce « codage » se fondent sur une *quantification* qui remplace la matérialité et la surface de l'image par des milliers de pixels. Ainsi, la définition d'une image numérique concerne seulement des valeurs numériques (les « notations »). Tandis que l'affichage de cette image fonctionne comme une « carte de points lumineux et colorés » qui reconstitue visuellement chaque point de l'image à l'aide d'un dispositif d'affichage.

Face au tableau ou à ce dispositif d'affichage, l'œil du spectateur (« l'impression subjective ») voit donc à peu près la même image. Pourtant, dans un cas, le regard du spectateur parcourt la surface matérielle de l'image, et dans l'autre, le parcours du regard s'appuie sur une simulation de cette surface à travers la matérialité de l'écran.

Dans les deux sections suivantes, nous expliquerons pourquoi cette question de la ressemblance entre des images matérielles et des reproductions numériques n'est pas tant

[263] Dans ses travaux, René Thom a proposé une approche mathématique des formes ; cette approche considère les formes comme des entités géométriques et cherche à mettre en évidence la dynamique des formes, car « une succession de formes spatiales ne peut être prédite que si l'on peut lui appliquer une théorie mathématique » (Thom & Noël, 1994 : 187).

un problème. Comme nous l'avons suggéré, la conversion numérique offre plutôt de nouvelles conditions de visibilité et de manipulation des images.

4.1.1. Quantification et affichage des images numériques

En parallèle de ses activités artistiques et de ses travaux sur l'*art numérique* (cf. Couchot & Hillaire, 2003), Edmond Couchot a publié *Images, de l'optique au numérique* (Couchot, 1988). Dans cet ouvrage important, l'auteur défend l'idée que l'image numérique simule les objets qu'elle représente. Comme nous l'avons expliqué[264], cette simulation se base sur deux dispositifs complémentaires. D'un côté, l'image numérique doit être affichée par le moyen d'un *écran*, le « dispositif périphérique de sortie » connecté à l'ordinateur. De l'autre, cette image est générée par la *carte graphique* de l'ordinateur, c'est-à-dire le système de traitement de la machine qui transmet à l'écran les données numériques stockées dans la mémoire de l'ordinateur.

Pour Couchot, l'*écran informatique* est une « *mosaïque totalement ordonnée* où les valeurs affectées à chaque point élémentaire de l'image définissant sa position et sa couleur sont discrètes et quantifiées » (Couchot, 1988 : 145). Du point de vue informatique, l'image numérique est réduite à un « plan » qui est complètement fractionné. Ce plan est découpé en un « damier de points » où l'intensité lumineuse et la couleur sont définies pour chaque point (ou pixel) donné dans un plan de largeur *w* (en anglais : *width*) et de hauteur *h* (*height*). À chaque point correspond une intensité quantifiable et les valeurs numériques de ces intensités sont stockées dans la mémoire de l'ordinateur (*ibid.* : 140-146). Dans cette mémoire (le support électronique de stockage de données), l'image numérique est donc caractérisée par :

- Sa *définition* (sa hauteur multipliée par sa largeur, exprimée en « pixels »).
- Sa *résolution* (le nombre de pixels par unité de longueur de l'image ; plus la résolution est élevée, meilleure est la qualité de l'image).
- Les *informations sur la distribution des pixels* (informations qui sont mémorisées sous une forme binaire).

Ainsi, la spécificité de l'*image numérique* est directement liée à la manière de stocker et d'afficher ce type d'image. L'image numérique est une configuration visible

[264] Cf. les discussions sur le support perceptible et le support imperceptible des textes informatiques dans le **chap. 1, sect. 1.3.4.**

grâce à un *ensemble de pixels qui simule la présence d'une image.* Chaque pixel correspond à un triplet de couleurs (rouge, vert, bleu) et à un vecteur de valeurs en rapport avec la taille de l'image (hauteur et largeur)[265]. Le revers de cette configuration prend toujours la forme abstraite d'une *matrice*, « un tableau de nombres composé de colonnes et de rangées qui a l'avantage de permettre de nombreuses opérations mathématiques particulièrement intéressantes dans la synthèse et dans l'analyse d'image » (*ibid.* : 191). C'est pourquoi, selon Couchot, l'image numérique rend possible « une image totalement pénétrable », une image qui est à la fois « réelle » et « potentielle », même si l'auteur ajoute qu'il ne faut pas considérer cette image comme un « simple moule » (*ibid.* : 193). Cette « image-matrice » ne concerne pas seulement « l'image affichée sur l'écran, ni la mémoire image dont elle est la transduction complète ou partielle » ; cette notion désigne plus largement l'ensemble des *processus informatiques invisibles* qui permettent de générer ou de manipuler une image (cf. *ibid.* : 145 & 189-198).

Ceci est un point fondamental dont nous avons déjà suggéré l'ampleur[266] : à l'*image visible sur l'écran* (celle qui est composée de pixels) correspond une *image invisible* qui est uniquement *composée de nombres* qui sont rassemblés sous une forme binaire. Par conséquent, c'est la « nature du support » et « l'organisation du support » qui différencient l'image traditionnelle (picturale ou optique) de l'image numérique, car le support spécifique de l'image numérique est une « collection de nombres » (*ibid.* : 192). Comme l'explique Couchot, les procédés pour générer une image numérique sont différents des procédés picturaux ou optiques. Cependant, « générer » ou « manipuler une image numérique » nécessite toujours de « donner des instructions à l'ordinateur », d'« écrire un programme », et donc de « *définir la forme avec du langage* », un « langage de programmation dictant ses instructions logiques et mathématiques à une machine qui les exécute fidèlement » (*ibid.* : 191). En effet, un ordinateur traite uniquement une représentation quantifiée de la surface des images, indépendamment de la description verbale de ces images.

À ce propos, l'auteur souligne un paradoxe : « au moment même où nous croyons contrôler totalement le plus petit élément constituant physique de l'image, sa forme s'évanouit – et son sens – dans l'incertitude paradoxale de la précision mathématique »

[265] Dans une image numérique en noir et blanc, l'intensité correspond à un niveau de gris (plus un point est sombre, plus son niveau de gris est faible).

[266] Cf. **introduction & chap. 1, sect. 1.4.**

(cf. *ibid.* : 150). En effet, les matrices de pixels sont concrètement une collection impénétrable de données chiffrées. Ces matrices offrent une représentation mathématique des caractéristiques (*feature representations*) plastiques d'une image sous la forme d'une fonction vectorielle[267]. L'avantage est que cette représentation est de moindre dimension que le nombre effectif de pixels dans une image, tout en gardant les principales caractéristiques qui déterminent l'information visuelle contenue dans cette image. Cette information est circonscrite dans des représentations statistiques auxquelles l'utilisateur de l'ordinateur n'a pas directement accès. Même si elles sont parfaitement adéquates pour le traitement informatique, ces informations sont ininterprétables pour l'humain, car la machine n'analyse pas quelque chose de visible ou de présent, elle analyse des nombres correspondant à une sélection programmée de caractéristiques représentatives issues d'une matrice de pixels.

Ces dernières remarques suggèrent une question simple : l'étude du visuel et des images s'est-elle dégagée d'un paradigme d'interprétation fondé sur le langage verbal pour finalement engager ce domaine d'étude dans un paradigme mathématique, celui de la quantification et des langages de programmation ?

À ce stade, il faut remarquer une limitation cruciale : les méthodes d'analyse que nous présentons dans **ce chapitre** ne sont pas applicables directement sur un corpus d'images parce qu'il faut maitriser un *langage* pour réaliser des analyses. En effet, toutes ces possibilités d'analyse doivent être programmées, au préalable, en s'appuyant sur un langage informatique[268]. Il ne faut donc pas se méprendre : le traitement informatique

[267] Cette représentation correspondant à un *tenseur*, à savoir un objet dont la valeur s'exprime dans un espace vectoriel (structure d'ensemble permettant de réaliser des combinaisons linéaires, qui sont construites sur cet ensemble) ; il s'agit d'un « objet abstrait », qui se modifie lorsqu'on passe d'une représentation d'un ensemble à une autre.

[268] Dans certains cas, ces programmes sont archivés et parfois accessibles sur Internet, de telle sorte que nombreuses ressources et codes informatiques pour mettre en place des analyses computationnelles sont disponibles en libre accès. À titre d'exemple, à l'aide du langage de programmation *Python*, il est possible d'utiliser un ensemble de fonctions (*OpenCV*) relatives à la *computer vision*. Ces fonctionnalités fournissent des outils pour manipuler les espaces de couleur dans une image numérique ; cf. : https://www.python.org/about/. Les outils d'*OpenCV* peuvent être utilisés pour mettre en place des systèmes de traitement automatique d'images. Pour ce faire, *OpenCV* est combiné à deux autres *packages Python* : *NumPy*, le principal *package* de calcul scientifique utilisé en *Python* ; *Matplotlib*, une bibliothèque de représentation graphique. En s'appuyant sur le principe de libre partage (*open source*), le langage de programmation *Python* offre des possibilités intéressantes pour réaliser « soi-même » des traitements

d'images doit au préalable être codé (c'est-à-dire « écrit ») par un humain dans un environnement adéquat. Cet environnement est l'interface système (*shell*) où les opérations d'analyse sont programmées (le programmeur code, tandis que la machine exécute le programme). S'il n'est pas nécessaire de comprendre toutes les fonctions des logiciels utilisés dans ces protocoles, il faut néanmoins une compréhension de base du codage informatique. Même s'il ne s'agit pas d'entrer ici dans le détail de ce sujet (notre propos ne concerne pas le codage informatique), nous pouvons néanmoins remarquer de quelle manière ce type programmation est mis en place.

Nous avons vu que des travaux en sémiotique considèrent que, comme les langues naturelles, les langages informatiques peuvent être considérés comme des systèmes réflexifs[269]. Dans ce sens, les langages de programmation sont compris comme des systèmes sémiotiques comprenant des processus d'énonciation/interprétation, même s'il ne faut pas penser par-là que les systèmes informatiques concernent les mêmes types de signes que dans les langues naturelles. Du reste, cette dimension énonciative est observable dans la programmation informatique.

À titre d'exemple, le *script* comprend plusieurs étapes qui sont écrites dans un langage informatique qui permet de démarrer et de coordonner l'exécution de programmes. Ce « script » est exécuté à partir du code source[270] à l'aide d'un interprète (fig. 1) qui (1) *lit* et *analyse une instruction*, (2) *l'exécute si la syntaxe de l'instruction est correcte*, puis (3) *traite l'instruction suivante*. Ainsi, en tant qu'outil informatique, l'interprète évalue et réalise les instructions au fur et à mesure de la lecture des lignes du script.

---

computationnels sur des images numériques ; cf. : https://realpython.com/python-opencv-color-spaces/#simple-segmentation-using-color-spaces.

[269] **Chap. 1, sect. 1.4.** Dans une certaine mesure, le codage informatique partage cette caractéristique avec le visuel : chacun dispose de ses propres mécanismes de réflexivité.

[270] Ce code source est un texte (fichier texte) qui comprend toutes les instructions qui ont été écrites dans un langage de programmation. D'un autre côté, notons qu'un « programme compilé » est exécuté à partir de la traduction du fichier source dans un langage machine. Un *assembleur* ou un *compilateur* permet justement de transformer un *code source* en un *code objet* qui comprend des instructions exécutables par un système informatique (le code objet est une suite d'instruction sous la forme de données binaires et non plus d'un texte).

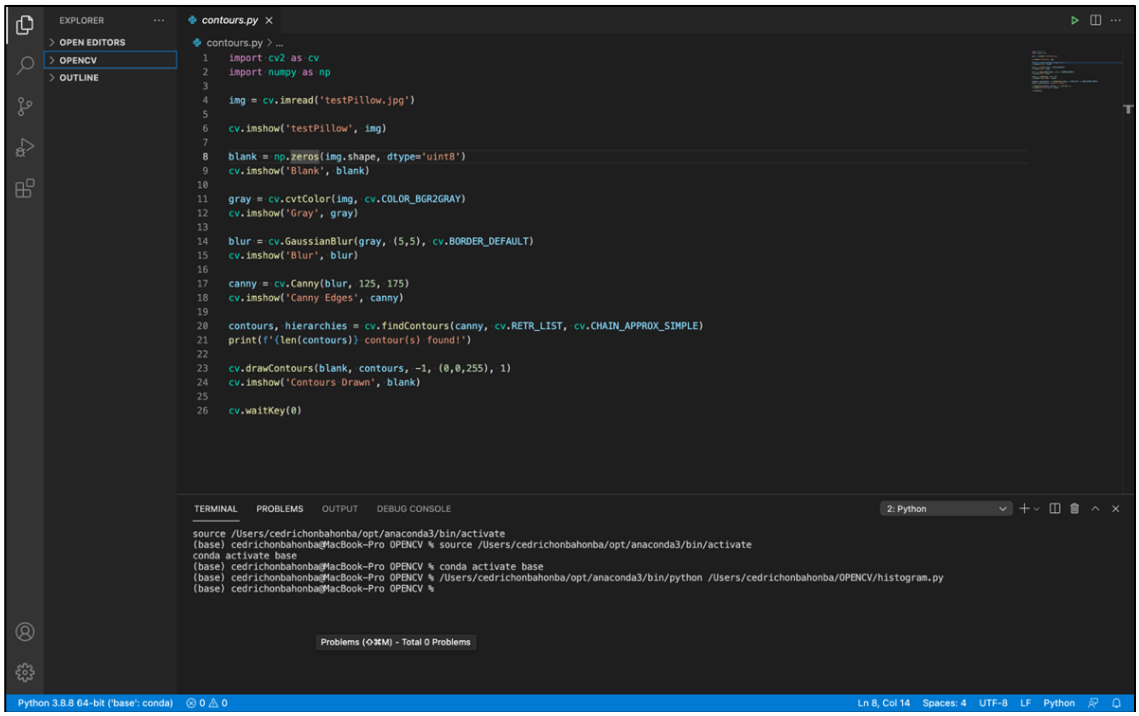


Figure 1 : Exemple d'un interprète permettant d'exécuter un programme écrit dans un langage informatique (dans cet exemple, il s'agit de Python)

Ce changement d'environnement et de techniques pour traiter et analyser des images entraîne aussi de nouveaux défis, notamment la manière d'interpréter les résultats des analyses computationnelles qui sont produites à l'aide d'une programmation informatique. Il faut toutefois remarquer que le but n'est pas d'exposer cette programmation et les opérations de traitement informatique, mais plutôt de proposer des interfaces ergonomiques, c'est-à-dire des dispositifs de lecture simples à utiliser.

Dans la **première partie**, nous avons montré qu'il faut intégrer le numérique dans des usages pratiques qui ne se réduisent pas à un système informatique ou aux langages de programmation. En effet, même si l'informatique offre de nouveaux instruments pour analyser des images sous la forme de nombres, le traitement informatique d'images n'est pas une « solution miraculeuse » ; au contraire, les informaticiens sont les premiers à souligner les limites que rencontrent encore leurs instruments, car ceux-ci permettent d'analyser de grandes quantités d'images *seulement si* ces images sont converties en données numériques (alors que la plupart des archives ou collections d'images ne sont pas encore numérisées) et *seulement si* ces images sont décrites au préalable par des humains[271].

[271] Cf. nos commentaires sur la classification d'images et les difficultés concernant la structuration particulière des images et du visuel dans le **chap. 3, sect. 3.3.**

4.1.2. Le langage des nouveaux médias

Dans la section précédente, nous avons présenté deux caractéristiques principales de l'image numérique. Celle-ci est caractérisée, d'une part, par une quantification complète de la surface de l'image transformée en valeurs numériques et, d'autre part, par des modalités d'affichage des images par le moyen d'une matrice de pixels et d'un écran d'ordinateur.

Parmi les travaux que nous avons évoqués[272], un ouvrage important a considéré les conséquences de cette transformation en données numériques des médias. En effet, *Le langage des nouveaux médias* de Lev Manovich (2015a [2001]) est sans doute un ouvrage qui a bouleversé la théorie des médias. L'auteur propose notamment cinq principes fondamentaux pour caractériser ce qu'il nomme les « nouveaux médias » par rapport aux technologies médiatiques non numériques (cf. *ibid.* : 98-131). Ces principes complètent les considérations sur le support numérique que nous avons déjà présentées[273].

Le premier principe est le *codage numérique*. Le numérique est un code qui permet de transformer un média en une représentation numérique par le moyen d'un échantillonnage et d'une quantification de ce média.

Le deuxième principe est la *structure modulaire des médias et nouveaux médias* : les éléments constitutifs d'un média (niveau 1) sont assemblés en objets médiatiques (niveau 2), qui peuvent également être combinés en d'autres objets plus complexes (niveau 3), et ainsi de suite, sans que chaque niveau perde son indépendance.

Les deux premiers principes rendent possible un troisième principe : l'*automatisation* des opérations pour créer et manipuler des nouveaux médias à l'aide d'algorithmes.

Le quatrième principe est la *variabilité des nouveaux médias*, parce que ceux-ci peuvent exister dans de nombreux formats techniques et différentes versions.

---

[272] **Chap. 1, sect. 1.1.** Voir aussi les travaux sur le « musée virtuel » de Bernard Deloche et ceux de Matteo Treleani sur le « patrimoine numérique » (cf. **chap. 2, sect. 2.2. & 2.2.2.**) ou encore l'approche théorique de Bruno Bachimont qui distingue quatre niveaux qui conditionne la lisibilité des ressources numériques : (1) le codage du contenu, (2) le codage des métadonnées et des formats fichiers, (3) les applications, et (4) les plateformes informatiques (cf. **chap. 2, sect. 2.3.3**).

[273] Cf. **chap. 1 (sect. 1.3. & 1.4.) & supra (sect. 4.1.1.)**.

Ces quatre principes ont pour conséquence le cinquième principe, que l'auteur nomme *transcodage culturel* : « transcoder » consiste à transposer un contenu dans un autre format technique, et dans ce sens, on peut envisager que l'impact de l'informatique dans les pratiques culturelles impose une transformation des contenus, et surtout un changement dans notre relation aux contenus qui sont transformés[274].

Dans le même ouvrage, Manovich a également mis en évidence que la popularité croissante des ordinateurs et d'Internet influence profondément nos interactions avec les objets culturels. En effet, l'ordinateur est devenu le « filtre de la culture dans son ensemble ; une forme de médiation à laquelle toutes sortes de productions culturelles et artistiques ont recours » (*ibid.* : 157). Pour mettre en œuvre cette « forme de médiation »[275], l'*interface de l'ordinateur* « agit comme un code qui véhicule des messages culturels dans divers médias », car « tout ce à quoi l'on a accès » sur le réseau Internet *transite* soit par « l'interface du logiciel de navigation » soit l'interface du « système d'exploitation » (*ibid.*). L'auteur ajoute même que ces interfaces ont tendance à déterminer « la manière dont l'utilisateur conçoit l'ordinateur » et surtout « la façon dont il se représente tout autre objet médiatique auquel l'ordinateur lui donne accès » (*ibid.* : 203-204)[276].

Pourtant, s'il est admis que la société s'est progressivement transformée en une *société de l'écran*[277], il reste à savoir comment « les affichages informatiques prolongent ou remettent en cause la tradition de l'écran » (*ibid.* : 204). Ne faudrait-il pas interroger cette « forme de médiation » qui se situe entre l'espace physique dans lequel se trouve le corpus du spectateur et l'espace de l'écran ?

---

[274] Ce principe rejoint le point de vue de Bachimont, lorsqu'il remarque que le numérique « modifie profondément l'économie de nos supports, dispositifs de pensée et d'action », notamment parce que ce support implique des opérations de calcul et des processus automatisés dans la manipulation des connaissances.

[275] **Chap. 2, sect. 2.2.**

[276] Sur les interfaces, cf. **chap. 2, sect. 2.3.**

[277] À ce propos, voir par exemple le chapitre « L'Écran-monde » dans Lipovetsky & Serroy (2007 : 281-330) où sont distingués trois moments du pouvoir de l'écran : (1) la toile de l'écran de cinéma, (2) la télévision, et enfin (3) l'ordinateur : « L'ordinateur et son écran son chaque jour mobilisés davantage dans les domaines les plus variés de la création, de la communication et de l'information. Pouvant simuler la peinture, la photographie, le cinéma, la vidéo, l'architecture, la voix, la musique, la danse, et s'appliquant à la quasi-totalité des activités humaines, le numérique est bien la technique englobante, universelle, du temps de l'écran global » (Lipovetsky & Serroy, 2007 : 320).

À propos de cette relation entre « espace du spectateur » et « espace de la représentation », Manovich (*ibid.* : 214) propose de définir « l'écran comme une fenêtre ouvrant sur l'espace de la représentation »[278]. L'auteur distingue aussi trois types d'écrans à partir de la notion de temporalité : (1) l'*écran classique*, lorsque le dispositif affiche une image statique et permanente ; (2) l'*écran dynamique*, lorsque le dispositif affiche une image (en mouvement) du passé ; et (3) l'*écran en temps réel*, lorsque le dispositif « montre le présent ». Un point important est que, dans chaque cas, l'écran nécessite un « spectateur immobile » qui regarde un dispositif fixe. Peu importe le type d'écran, l'observateur « regarde toujours une surface plane, rectangulaire, existant dans l'espace occupé par notre corps » ; cette « surface plane » sert de « fenêtre ouvrant sur un autre espace » (*ibid.* : 229).

Cette relation « surface-regard » est évidemment apparue avant l'écran ordinateur, puisqu'elle concerne aussi les écrans de télévision et de cinéma. Dans un passage de *L'écran global*, Gilles Lipovetsky et Jean Serroy (2007) suggèrent par exemple que l'écran d'ordinateur favorise un « nouveau rapport au monde »[279].

> [Un] nouveau rapport à l'espace-temps qui s'est mis en place, une espèce d'hyper espace-temps dans lequel tout s'effectue tout de suite, en flux tendu, dans l'instantanéité du temps réel. Et cela touche tous les domaines de l'activité humaine, de la vie économique à la vie quotidienne : partout, l'écran travaille à faire sauter les limites du temps et de l'espace (Lipovetsky & Serroy, 2007 : 329).

Comme le suggère ailleurs Vilém Flusser – dans « Une révolution dans le domaine des images » – si l'*imagination* est « la démarche complexe de la création d'images », les progrès techniques et scientifiques « ont transféré l'imagination de l'homme dans l'appareillage pour la perfectionner » (cf. Flusser, 2006 : 59). Selon l'auteur, ce transfert de l'homme à la machine a permis à la fois une « imagination plus efficiente » et un renouvellement de l'imagination, car le pouvoir d'imagination de l'image technique est tellement puissant que « nous prenons les images pour la réalité, mais encore nous vivons en fonction d'elles » (*ibid.* : 67). Pour Flusser, les nouvelles possibilités d'imagination (notamment celles qui sont le fruit de la technologie numérique) permettent même de *voir les connaissances en image* (en tant que « faculté de mettre en image au moyen

[278] Voir nos commentaires des travaux de Louis Marin sur la représentation, dans le **chap. 3, sect. 3.2.1.**

[279] Voir **note 277**.

d'appareils les concepts les plus abstraits », cette nouvelle imagination « commence à nous permettre à présent de voir les connaissances en images, donc d'en faire une expérience d'ordre esthétique » ; cf. *ibid.*).

Ces considérations sur l'écran et l'imagination technique rappellent sans doute l'engouement envers le potentiel de la reproduction photographique que nous avons constaté dans les pratiques documentaires et d'archivage[280]. Dans la continuité de la reproduction photographique, le *transcodage* (cf. **supra**) suscite des questionnements sur la relation du spectateur aux connaissances lorsque celles-ci sont reproduites et représentées dans un environnement informatique ; notons au passage que Flusser met lui aussi en évidence le rapprochement entre la reproduction numérique et la reproduction photographique, car le dispositif photographique a profondément transformé notre rapport au visuel et au visible en poursuivant un projet scientifique général qui visait à prendre du recul vis-à-vis de la « réalité objective », notamment à l'aide d'appareils dont le but est de « fixer » et de « codifier ce qui était vu » (Flusser, 2006 : 55-68).

Tom Gunning (2006) a quant à lui considéré ce passage à l'image numérique dans le but de montrer qu'il n'y a pas de différence entre la photographie chimique (ou « traditionnelle ») et la photographie numérique. La photographie chimique ou la photographie numérique comportent évidemment plusieurs différences pertinentes (technologie du support, manipulation par l'intermédiaire de procédés chimiques ou de programmes informatiques, perception culturelle), comme elles comportent aussi des points communs (règles optiques, dimension figurative ou plastique). Mais, pour Gunning, même si des données numériques ne ressemblent pas à une photographie, cette différence de format technique ne modifie pas forcément la *qualité indicielle de l'image* qui est finalement observée par le moyen d'un écran. En effet, qu'il s'agisse de photographie analogique ou numérique, l'inscription sur un support est « inhérente aux effets de la lumière » sur des produits chimiques ou sur des capteurs électroniques. Dans chaque cas l'image produite est déterminée « de façon indicielle »[281] par rapport à des

[280] **Chap. 1, sect. 1.1.**

[281] Il faut être prudent lorsqu'on définit l'image photographique par rapport à la notion d'indice. Un indice est un signe résultant d'un contact physique. Dans le cas de la photographie, les rayons lumineux agissent sur le support d'inscription pour réaliser ce contact. Cependant, si cette conception de l'image photographique se focalise évidemment sur le processus d'inscription sémiotique, elle met de côté le fait que la photographie signifie surtout à travers des pratiques qui l'emploient ; à ce sujet, voir les travaux dans Basso Fossali & Dondero (2011).

« objets extérieurs à l'appareil » (*ibid.* : §6), et elle n'est observable que par le moyen d'un dispositif qui rend matériellement visible l'inscription (tirage ou écran).

Selon Gunning, il faut nuancer « cette idée déjà bien ancrée que l'on peut soumettre les images numériques à des manipulations infaisables avec des photographies argentiques », car cette idée supposant « que seul le numérique transformerait ses données en une forme intermédiaire encourage le mythe que la photographie impliquerait un processus transparent, un transfert direct de l'objet à l'image photographique » (*ibid.* : §6). Le propre de la photographie chimique n'est pas le caractère « non manipulable » de l'image, bien qu'on ait tendance à le croire, notamment parce que « la photographie traditionnelle dispose aussi de procédés pour amoindrir, ignorer, voire anéantir l'indicialité », même si « les procédés numériques peuvent accomplir ces altérations plus vite et de façon moins criante » (Gunning, 2006, §8). D'un côté, même si les techniques de production et de présentation sont différentes, il n'y a aucun problème à admettre que la photographie, chimique ou numérique, manifeste des images de la même manière. De l'autre, « la commodité, la rapidité et la qualité de la manipulation numérique constituent une nouvelle étape importante dans la technologie de l'image » (*ibid.*). Il faut toutefois prendre en compte :

- Les contextes et les dispositifs d'énonciation visuelle parmi lesquels ces possibilités de manipulation acquièrent une valeur dans des pratiques sociales.
- La « dette considérable que ces *transformations* doivent à l'*histoire de la photographie* (même si elles nous en détournent notre attention) » (*ibid.* : §7 ; nous soulignons).

Ainsi, le passage de la photographie traditionnelle au support numérique ne modifie pas l'image observée par le moyen des écrans (ni son iconicité ni son indicialité). Ce changement de support concerne plutôt les modalités d'affichage et de transformation des images, ou encore les possibilités d'interaction avec une image[282].

[282] Plus récemment, Maxime Fabre (2019) a lui aussi remarqué que l'empreinte photographique est généralement considérée comme le résultat d'une technique « qui fait appel à l'expérimentation et à la manipulation pour produire des *effets de sens* » (*ibid.* : §30). Fabre propose de faire ressortir la complémentarité et la continuité des dispositifs chimique et numérique, plutôt que de les opposer, ou de se focaliser seulement sur l'exactitude analogique de leurs rendus : « Si rupture il y a entre le chimique et le numérique, elle ne se fait pas dans la captation de l'image mais bien dans la configuration, le conditionnement et la circulation de celle-ci » (*ibid.* : §33).

La question n'est donc pas de savoir si l'image numérique représente mieux le réel que les autres supports d'images, mais de faire ressortir que les *nouveaux médias* s'appuient, comme nous l'avons vu, sur certains principes : (le codage numérique, la modularité, l'automatisation, la variabilité, le transcodage ; cf. Manovich, *op. cit.*). Pour Manovich, ces nouveaux médias sont composés d'une *strate culturelle* et d'une *strate informatique*, et chaque strate s'influence mutuellement, car « la logique de l'ordinateur exerce une influence importante sur la logique culturelle traditionnelle des médias », à tel point que la seconde strate modifie radicalement la première (*ibid.* : 126). C'est pourquoi, selon l'auteur, il faut s'aider des approches informatiques pour comprendre la logique des nouveaux médias : « De l'étude des médias, nous passons à quelque chose que l'on pourrait appeler l'"étude des logiciels" (*software studies*) ; ainsi de la théorie des médias nous passons à la théorie des logiciels » (*ibid.* : 127).

Les *software studies* correspondent à un champ de recherche interdisciplinaire qui étudie les logiciels et leurs effets dans la société et la culture (Manovich, 2013a). Ces recherches suivent trois perspectives qui sont complémentaires :

- Étudier la programmation informatique des algorithmes et des logiciels en prenant acte de l'importance des infrastructures numériques et de l'environnement technologique qui résulte de ces infrastructures.
- Considérer la relation entre nouveaux médias et méthodes d'analyse statistique et computationnelle des données à travers le prisme des sciences humaines.
- Mobiliser l'avancée des recherches scientifiques sur l'analyse automatique de données pour explorer de larges bases d'images.

Pour Manovich, il ne s'agit pas seulement d'emprunter des théories et des méthodes d'analyse statistique et computationnelle. L'auteur propose de penser une nouvelle théorie des médias plus adéquate à l'ère du numérique, et surtout de comprendre le fonctionnement et les concepts qui gouvernent le support numérique afin d'estimer comment ce support peut être utilisé dans l'étude des pratiques culturelles. Dans cette perspective, Manovich (2013b, 2015b, 2017a & 2017b) a proposé d'utiliser des ordinateurs pour effectuer des analyses informatiques sur un grand nombre de médias visuels en fonction des paramètres visuels associés à des caractéristiques plastiques (couleurs, luminance) ou à des informations relatives aux métadonnées (auteur, année, format, etc.). Pour ce faire, l'auteur combine différentes techniques et domaines qui sont propres à l'ingénierie informatique – notamment, le « webdesign », les « sciences des

données » (*data science*), le « design numérique » (*interaction design*), la « visualisation de données » (*data visualization*), l'« apprentissage automatique » (*machine learning*), ou encore des techniques de l'analyse statistique – avec des approches en sciences humaines comme la « sociologie des médias » (*media studies*) ou l'« étude des cultures visuelles » (*visual studies*).

Comme nous l'avons suggéré[283], d'une part, de nombreuses collections et archives d'images numériques sont actuellement accessibles sur Internet, et d'autre part, les sciences humaines s'interrogent sur les usages du numérique, notamment à travers le domaine des humanités numériques. Les travaux de Manovich font cependant ressortir que les *software studies* ne disposent pas de beaucoup d'échantillons concernant la diversité médiatique actuelle dans les pratiques visuelles. En effet, les premières entreprises de conversion numérique d'archives d'images s'intéressent surtout à des archives photographiques et à des collections artistiques. Par conséquent, les larges ensembles d'images que nous avons évoqués concernent surtout des artefacts culturels du passé. Pour réagir à ce problème, nous verrons que des entreprises plus récentes proposent également d'étudier les images issues de pratiques amateurs et des réseaux sociaux, ou encore d'étudier l'influence des dispositifs médiatiques contemporains sur la culture visuelle (p. ex. : smartphone, ordinateurs, caméra numérique, etc.)[284].

C'est du reste la perspective choisie par le *Cultural Analytics Lab*[285]. Ce laboratoire (qui a inspiré de nombreuses initiatives semblables) rassemble des chercheurs de plusieurs disciplines pour interroger la problématique du traitement automatique et de l'exploration de la grande quantité de données numériques issues de pratiques culturelles, qu'il s'agisse de données qui sont produites et diffusées par des professionnels des médias ou par les utilisateurs du réseau Internet.

---

[283] **Chap. 2, sect. 2.2. & 2.3.**

[284] **Chap. 6, 7 & 8**.

[285] Voir http://lab.culturalanalytics.info. Ce laboratoire de recherche affirme qu'il est nécessaire de développer une procédure systématique d'échantillonnage et de représentation des données culturelles telles que les grands ensembles d'images ou de vidéos. Depuis que les progrès en informatique ont permis d'élaborer des programmes qui sont capables de réaliser des tâches analytiques complexes sur de larges quantités de données, le but est d'utiliser la programmation informatique pour analyser les données culturelles.

4.1.3. *Image Processing*, détection des formes et classification d'images

Dans les sections précédentes, nous avons vu que l'image numérique peut être considérée comme un ensemble de pixels, correspondant à une *quantification complète* de l'image et sa transformation en valeurs numériques. Comme nous l'avons suggéré, l'échantillonnage (*sampling*) sert à découper une image en suivant une grille régulière, puis à associer une intensité à chaque cellule de cette grille ; la détection de la position de chaque point de ce plan se base sur un système des coordonnées, comparable à une sorte de « repère cartésien » de coordonnées *x*,*y* (*x* étant son abscisse, *y* son ordonnée). La principale différence entre les approches analytiques développées dans le cadre des sciences humaines et le traitement informatique d'images est que l'ordinateur traite uniquement ces représentations numériques des caractéristiques principales de la surface de l'image. Le « traitement d'images » (*image processing* ou *image science*) est la branche des sciences informatiques qui traite les images comme des « entrées » (*inputs*) et qui transforme les propriétés plastiques des images en données numériques pour faciliter sa manipulation technique, son étude et, dans la mesure du possible, sa compréhension. Plus précisément, cette branche désigne un ensemble de techniques computationnelles pour manipuler des images sous la forme de *matrices de pixels*, puisque nous avons vu que ces matrices « représentent numériquement » l'information visuelle contenue dans une image.

Dans *Fundamentals of Digital Image Processing*, Chris Solomon et Toby Breckon (2011) affirment que la conversion des images en données numériques rend possibles diverses techniques, notamment l'*amélioration* (*ibid.* : 87 *sq.*) et la *restauration* (*ibid.* : 141 *sq.*) des images à l'aide d'opérations mathématiques de filtrage (souvent afin de rendre plus clairement visible les informations dans l'image). Parmi ces techniques, nous voulons surtout faire ressortir celles qui concernent la *distribution de pixels dans une image*. À titre d'exemple, en informatique, un *histogramme d'image*[286] est un graphique de la fréquence relative d'occurrence de chacune des valeurs de pixel de l'image par rapport aux valeurs elles-mêmes. Ce graphique permet surtout de « révéler le contraste de base dans l'image », ainsi que les différences dans la distribution des couleurs qui compose l'image » (*ibid.* : 63 ; nous traduisons de l'anglais). En outre, Solomon et Breckon mentionnent également que des « opérations morphologiques » peuvent être

[286] Cf. **infra, sect. 4.2.**

appliquées sur les pixels dans le but de faire ressortir des *formes* et des *structures* dans des images.

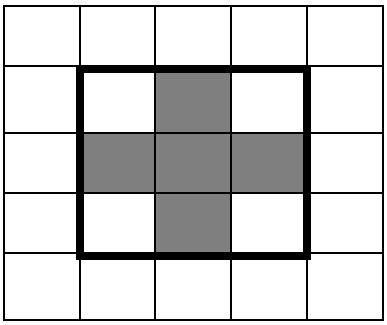
> Dans le traitement des images, la morphologie mathématique est utilisée comme moyen pour identifier et extraire des descripteurs d'images significatifs basés sur les propriétés de la forme ou de la structure de l'image. Les principaux domaines d'application sont la segmentation ainsi que le comptage et l'inspection automatisés (*ibid.* : 197).

Quelles sont ces opérations que l'*image processing* met à disposition pour analyser ces « structures » et ces « formes de l'image » ? Pour Solomon et Breckon, les *shape descriptors* (« descripteurs de forme ») tiennent en rôle important dans les opérations du traitement informatique d'images comme la *feature extraction* (« extraction de caractéristiques ») ou encore la *pattern recognition* (« reconnaissance de formes »). En simplifiant, ces « descripteurs » ont pour but de réduire les « paires de coordonnées (x, y) » dans la représentation numérique d'une forme ; cette réduction produit une mesure approximative, plus compacte (voire une mesure à paramètre unique) de la forme qui est représentée (*ibid.* : 169 *sq.*). En effet, ces opérations sont appliquées sur des « images *binaires* » (*ibid.* : 197).

Dans une image binaire, chaque pixel peut prendre *deux valeurs discrètes* : *1* (la valeur est décrite comme activée) ou *0* (la valeur est considérée comme désactivée). Les seules propriétés retenues dans les images binaires sont la *forme*, la *taille* et l'*emplacement des objets dans l'image numérique*, de telle sorte que les pixels d'une image binaire qui ont une valeur de 1 sont appelés « pixels *de premier plan* », tandis que les pixels ayant une valeur de 0 sont appelés « pixels *d'arrière-plan* »[287]. En fonction des relations entre les différents pixels d'une image, le but de l'analyse d'images binaires est de déterminer si un pixel appartient à l'avant-plan ou à l'arrière-plan de l'image. Le changement de valeur d'un pixel dépend du type d'image et du type de traitement morphologique appliqué, mais ce changement de valeur dépend aussi d'un « élément structurant » (*structuring element* ; cf. *ibid.* : 197-199). Un *élément structurant* détermine un pixel central dans une image binaire et il détermine aussi comment les pixels entourant

[287] Cette conception est proche du principe de la perception figure-fond, principe qui est issu de la psychologie de la forme ; cf. la discussion sur les formes visuelles dans le **chap. 3, sect. 3.4.** La perception visuelle se diviserait en un avant-plan (la figure qui se détache et possède un contour défini) et un arrière-plan (le fond, moins distinct, sur lequel des figures « se détachent »).

ce point central doivent être considérés. Ainsi, à chaque section d'une image matricielle correspond un élément structurant : la distinction entre *forme* (pixels de premier plan) et *fond* (pixels d'arrière-plan) repose sur une opposition binaire (présence/absence) par rapport à un pixel central donné dans l'image (cf. fig. 2).



0	1	0
1	1	1
0	1	0

Figure 2 : Représentation schématique d'un « élément structurant » ; à gauche un « damier de pixels » contenant une forme visuelle, à droite la représentation numérique de cette forme.

Sur ce point, l'*image processing* rejoint certaines prises de position qui ont été mise en évidence par les approches en sciences humaines[288].

D'un côté, nous pouvons rapprocher la notion d'*image-matrice* des propositions de la sémiotique. Nous avons vu que les images peuvent être interprétées selon deux niveaux de lecture : au niveau figuratif ou au niveau plastique. Dans l'*image processing*, la surface de l'image est à peu près définie comme le plan d'expression des images en sémiotique plastique. L'image est dans chaque cas considérée comme une surface plane déterminée par la position (catégorie topologique) des formes (catégorie eidétique) et des couleurs (catégorie chromatique) sur cette surface. Les matrices de pixels se basent sur une *discrétisation extrême* d'au moins deux catégories de la sémiotique plastique ; seul un nombre limité d'informations sur les *couleurs* et leurs *positions* dans la surface de l'image est conservé par le système informatique. De plus, si nous reprenons la distinction proposée par le Groupe μ entre couleur, forme et texture des signes visuels, nous constatons que les éléments structurants de l'*image processing* concernent la texture de l'image, sa « microtopographie ».

D'autre part, plusieurs travaux insistent sur l'importance d'une approche pluridisciplinaire pour étudier le visuel, une approche qui prend en compte différentes méthodes pour analyser et comprendre des images, notamment les outils informatiques.

[288] **Chap. 3, sect. 3.1.2., 3.2.2. & 3.4.**

En outre, l'évolution des approches en iconologie a montré que, lorsqu'il s'agit d'étudier des images, il faut prendre en compte les continuités et les contrastes entre les formes et les structures dans les productions visuelles au fur et à mesure des époques. Sans même le savoir, l'*image processing* renoue avec l'analyse des formes visuelles, mais à un niveau plastique encore plus « profond » que les « formes » et les « motifs » des historiens de l'art ou que les « formants plastiques » des sémioticiens. Comme nous venons de le suggérer, le traitement informatique opère au niveau de la *microtopographie de l'image*.

Il ne faut pas minimiser ce point de rencontre, même s'il reste beaucoup de travail pour que cette rencontre se réalise. Le problème est que ces éléments structurants et les valeurs numériques des pixels n'ont pas directement de significations pour un humain. L'*image processing* fonde l'analyse d'images sur des *structures formelles* qui ne sont pas des *formes* et des *structures visibles*. Les approches analytiques de ce domaine reposent entièrement sur cette possibilité de passer de l'image au nombre afin de fournir très rapidement des résultats statistiques globaux à propos des caractéristiques plastiques des images transformées en données numériques (par exemple : des lignes qui composent des formes, des formes qui composent un objet, etc.). L'*image processing* peut participer, en tant qu'instrument, à une interprétation en classant les images en fonction d'une représentation numérique de la dimension plastique de l'image.

Par conséquent, des approches en sciences humaines s'intéressent à l'*image processing*, puisque ces techniques de détection des formes sont utilisées pour faire ressortir des similitudes visuelles dans un corpus d'images. Cependant, « ce qu'il est possible d'analyser » dans cette quantification de l'image dépend toujours, d'une part d'un langage de programmation, et d'autre part d'objectifs et de décisions humaines. Depuis les années 2000, plusieurs recherches ont mis au point des dispositifs informatiques pour aider les travaux en histoire de l'art, notamment la classification de corpus de peintures. Certains travaux ont montré que les performances des ordinateurs permettent de classer des peintures en fonction de caractéristiques formelles générales. Étant donné la grande quantité d'informations visuelles (autrement dit : la variation de valeurs des pixels) au sein de ce type d'images, il est nécessaire de diminuer la quantité

d'information traitée par l'ordinateur en limitant le traitement à des représentations de caractéristiques (cf. Herik & Postma, 2000)[289].

Néanmoins, si l'*image processing* offre évidemment des avantages en proposant des méthodes d'analyse assistées par un ordinateur pour reconnaître des formes dans des images sous forme de nombres, ces méthodes ne remplaceront pas de sitôt les méthodes d'interprétation l'histoire de l'art (cf. Stork, 2007 & 2009). En effet, la détection de formes n'aboutit pas sur une explication de ces formes, de telle sorte que cette détection doit nécessairement être complétée par les connaissances des historiens et spécialistes des arts visuels qui utilisent ces méthodes informatiques d'analyse.

D'une manière générale, l'objectif de l'*image processing* est de développer « des solutions pour comprendre avec précision et rigueur les significations sémantiques des données visuelles, qui comprennent à la fois des images en deux dimensions et des vidéos » (Chen, Ren & Kuo, 2016 : 1)[290]. Ce domaine de recherches a beaucoup évolué depuis l'émergence et la disponibilité de larges ensembles de données (*big data*) qui ont permis d'éprouver des techniques en relation avec un autre domaine des sciences informatiques. Les *big data analytics* est domaine spécialisé dans le traitement automatique de larges ensembles données numériques (cf. Simon, 2014, ou encore Suh & Anthony, 2017). Cette automatisation dans les processus d'analyse des données se rapporte à un projet plus général, celui de *simuler certaines fonctions cognitives humaines par le moyen de la programmation d'ordinateurs*, afin de *leur apprendre à réaliser des tâches de manière beaucoup plus rapide que ne pourraient le faire des humains*.

[289] La contribution de Jaap van den Herik et d'Eric O. Postma (2000) passe en revue les techniques actuelles de « prétraitement » (l'objectif principal du « prétraitement » est la réduction de la dimension d'une image en transférant les données originales vers une représentation numérique ; cette représentation retient les informations pertinentes pour une classification d'images). Les auteurs remarquent que la classification de peintures est une tâche difficile parce que les « caractéristiques spécifiques au peintre » doivent être conservées dans une représentation numérique qui est une réduction de l'image. Les résultats de cette étude montrent que les techniques de réseaux neuronaux (cf. **infra, sect. 4.2.2.**) sont les plus appropriées pour exécuter et affiner le prétraitement et la classification des images. Les auteurs concluent que le prétraitement doit être adapté à la tâche que l'ordinateur doit accomplir, à condition de combiner les connaissances du domaine relatif à cette tâche avec les techniques de réseaux neuronaux.

[290] Voir aussi : Morris (2004) ; Forsyth & Ponce (2003) ; Davies (2005).

L'hypothèse est que les phénomènes sociaux et les productions culturelles sont devenus accessibles aux compétences analytiques des ordinateurs[291]. En effet, en associant les outils de diverses disciplines (mathématique, statistique, informatique), la science des données (*data science*) vise actuellement à développer des instruments pour analyser un grand nombre de données numériques et en tirer des connaissances scientifiques. Si ces méthodes, statistiques ou informatiques, sont parfois utilisées dans les travaux en sciences humaines et sociales pour analyser et comprendre un phénomène social, la science des données cherche à systématiser les processus de traitement des données dans à peu près tous les domaines de recherches scientifiques. La difficulté rencontrée est pourtant double : d'une part, des compétences techniques sont nécessaires pour gérer les importantes quantités de données numériques qui résultent de ces analyses computationnelles ; d'autre part, ces instruments d'analyses doivent être optimisés parce que leurs applications ne sont pas toujours généralisables à divers types et formats de données.

Notons que l'analyse automatique d'images n'est pas la seule tâche que les informaticiens tentent d'apprendre à des machines. Le traitement automatique des langues (TAL) est lui aussi en plein essor[292]. Robert Martin remarque que « l'informatisation a changé complètement la donne », parce que les linguistes peuvent désormais facilement manipuler « des masses de données précédemment impossibles à réunir » (Martin, 2018 [2002] : 30). Pour l'auteur, le traitement est surtout utilisé pour automatiser certaines tâches, principalement des fonctions de reconnaissance ou de génération (*ibid.* : 120-122). Cependant, comme le constate Jacqueline Léon, il faut aussi remarquer les problèmes que ce type de traitement amène, dans une discipline telle que la linguistique, et à laquelle les sciences informatiques feraient dès lors « concurrence »[293]. Lorsqu'elle propose un historique des rapports entre linguistique et

[291] Voir Fielding & Lee (1998). Des travaux étudient aussi les possibilités informatiques d'analyse des comportements humains (Salah & Gevers, 2011).

[292] Voir notamment : Rastier, Cavazza & Abeillé (1994) ; Jurafsky & Martin (2008) ; Mitkov (2005) ; Kurdi, (2017 & 2018).

[293] « Les approches et les enjeux deviennent en effet radicalement distincts. Pour le linguiste, il s'agit d'accroître ses connaissances sur les langues et le langage contrairement aux acteurs du TAL pour lesquels il s'agit d'améliorer la performance d'un système technologique. Le rapport de force s'est inversé et les informaticiens deviennent dominants. Les projets de TAL sont développés au sein de laboratoires d'informatique et les recrutements se font parmi les informaticiens. Les travaux sont publiés dans des revues d'informatique : les mots-clés employés en TAL ressortissent essentiellement à l'informatique et le contenu des publications est le plus souvent très technique. Le public des grandes conférences vient

TAL, l'auteure distingue dans les grandes lignes quatre périodes décisives, mais qui ne se succèdent pas « de façon linéaire » : la première (de 1950 à 1960) est celle des « balbutiements et cohabitation des méthodes » ; la seconde (de 1960 à 1990) concerne « la linguistique computationnelle » ; la troisième (de 1990 à 2000) durant laquelle se forment les « grands corpus et outillage de la linguistique » ; la quatrième (à partir de 2010) celle des *big data* (Léon, 2018 : 195).

À cet égard, la situation de la linguistique et celles des disciplines qui étudient des images sont à peu près similaires. Suivant la remarque de Léon, la généralisation du traitement informatique concurrence et impose progressivement leurs conceptions à l'étude des images. C'est évidemment une crainte qu'il faut prendre en compte. Nous l'avons vu, l'énorme quantité d'archives d'images numériques semble imposer ce type de traitement. Il ne faut toutefois pas penser que le traitement automatique des langues ou des images résolve tous les problèmes des disciplines qui étudient les langues ou les images. Le numérique n'est pas une solution miraculeuse, apte à réaliser toutes les tâches et opérations que les scientifiques voudraient qu'un ordinateur soit capable d'exécuter.

En effet, des spécialistes remarquent que le traitement informatique ne résout pas tous les « problèmes classiques » qui sont liés à la classification des images, de sorte qu'il est encore difficile d'envisager une indexation automatique d'images complètement fonctionnelle[294]. Nous pouvons souligner au moins quatre enjeux principaux (nous reprenons ici les constatations de Picard, Gosselin & Gaspard, 2015).

(1) De *nombreuses archives d'images ne sont pas encore numérisées* et sont seulement accessibles dans les institutions qui conservent les supports matériels. L'activité de numérisation de certains fonds et collection est intense, mais la quantité de support à intégrer dans des environnements informatiques est vaste et augmente constamment. En amont de ce problème, l'inégalité quantitative entre les supports d'inscription fait qu'il y a des catégories d'images dont on dispose de peu d'exemples, tandis que d'autres formats d'images sont beaucoup plus nombreux.

(2) Les *institutions n'ont pas forcément indexé leurs archives ou collections d'images à partir des mêmes protocoles*. Il en résulte une importante fluctuation des supports et des

aujourd'hui quasi exclusivement des laboratoires d'informatique (ou d'ingénierie notamment pour le traitement de la parole) et quasiment plus de laboratoires de linguistique » (cf. Léon, 2018 : 201).

[294] Voir aussi **chap. 3, sect. 3.3.**

informations sous la forme de métadonnées. Dans la plupart des cas, ces informations sont systématiquement des descriptions verbales des images. En outre, nous avons remarqué que le traitement informatique est lui aussi soumis à un langage. Or, comme nous l'avons souligné, plusieurs travaux insistent sur la nécessité de s'éloigner du paradigme langagier dans l'étude des images et du visuel, puisque ceux-ci disposent d'une propre réflexivité qui leur est propre.

(3) L'*application d'une programmation informatique pour analyser et classer des images est considérée par les spécialistes de ce domaine comme une tâche complexe*. Nous l'avons remarqué, un ordinateur n'analyse pas des images, mais des représentations numériques qui sont la quantification complète des caractéristiques plastiques de la surface de l'image. Le problème est que les images renvoient souvent à des structurations et des relations particulières, difficiles à généraliser et qui requièrent de vastes bases d'images numériques et des architectures informatiques complexes pour entraîner les machines.

(4) Les *choix stratégiques pour visualiser les résultats du traitement informatique d'images aboutissent sur la conception d'interfaces qui permettent à l'utilisateur d'observer et de manipuler des ressources numériques*[295]. L'utilisation pratique de ces résultats exige de les présenter à travers un dispositif de lecture qui est compréhensible pour un humain. Nous avons montré que le risque est de réduire la compréhension des images à une compréhension uniquement verbale, ce qui influence la visée rhétorique du dispositif.

De plus, l'une des principales difficultés que rencontre le domaine du traitement informatique des données est le *semantic gap*, c'est-à-dire l'écart entre la manière dont un être humain peut comprendre des données, et la manière dont une machine considère et traite les mêmes données. En effet, l'analyse automatisée se base fondamentalement sur des techniques d'apprentissage de tâches (prise de décision, prédiction, production de connaissances) qui utilisent des algorithmes. Qu'il soit complexe ou simple, l'algorithme se définit comme un ensemble de règles opératoires propres à la résolution d'un certain type de calcul ou de raisonnement logique. Il s'agit d'une suite finie de règles à appliquer

[295] Dans la **quatrième partie**, nous présenterons différents projets dont le but est de concevoir des dispositifs de visualisation d'images qui permettent à l'utilisateur d'explorer et de manipuler une grande quantité d'images numériques en favorisant un renouvellement de la dimension rhétorique du traitement informatique d'images.

dans un ordre déterminé à un nombre fini de données pour arriver, en un nombre fini d'étapes, à un résultat, indépendamment des données qui sont traitées.

Plusieurs travaux sur la description, la classification et l'exploration de grands ensembles de données suggèrent pourtant que l'informatique offre de nouvelles possibilités pour comprendre ces données. Dans la section suivante, nous verrons que ces possibilités s'appuient également sur des travaux en *computer vision* et sur des recherches concernant les techniques d'apprentissage des ordinateurs. Étant donné que des institutions et des laboratoires de recherche envisagent d'employer ces techniques pour analyser une grande quantité d'images, nous présenterons quelques opérations qui sont utilisées dans la description et à la classification d'images numériques.

4.2. *Computer vision*, *data visualization* et *deep learning*

En parallèle de l'*image processing*, la *computer vision*[296] (ou « vision par ordinateur ») est un domaine scientifique en relation avec les recherches sur l'intelligence artificielle. Le but est d'automatiser certaines tâches effectuées par des ordinateurs dans le traitement et la compréhension d'images ou de vidéos numériques, et surtout de produire des informations quantitatives sur les images traitées.

Dans ce domaine, l'appellation *pattern recognition* (« reconnaissance de formes ») renvoie à un ensemble de protocoles informatiques dont le but est d'identifier des structures dans des images pour attribuer une *catégorie* à ces structures.

Comme nous l'avons suggéré, pour extraire de ces données des informations qui sont significatives du point de vue humain, il faut transformer ces images en données numériques. À ce propos, plusieurs travaux ont cherché à mettre en évidence les bénéfices des *neural networks* (Colagrossi, Sciarrone & Seccaroni, 2003) ou encore des « processus stochastiques »[297] (Li, 2004) dans les techniques de classification d'œuvres d'art. Pour ce faire, la cohérence interne d'une peinture est évaluée du point de vue géométrique,

[296] Voir notamment : *International Journal of Computer Vision* ; *IEEE Transactions on Pattern Analysis and Machine Intelligence* ; *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

[297] Un « processus stochastique » est un processus ou une fonction aléatoire qui vise à utiliser le calcul des probabilités afin de représenter l'évolution et la réalisation d'une variable aléatoire. Ce processus est notamment employé dans différents domaines de recherche pour interpréter des observations d'un phénomène qui comprend une dimension aléatoire, mais en fonction d'un facteur temporel.

notamment afin de mesurer des objets dans l'image, ou encore pour détecter des « centres d'intérêt » (des « motifs » dans les peintures), voire pour reconstruire les parties manquantes d'une peinture (Criminisi, Kemp & Zisserman, 2005). Pour « extraire des caractéristiques visuelles » (*visual features extraction*), ces méthodes consistent principalement en des opérations mathématiques calculées sur des images binaires substituées aux propriétés plastiques de l'image. Il s'agit, par exemple, de mettre en place des descripteurs globaux et des descripteurs locaux pour décrire des « zones d'intérêt » : un point d'intérêt est un point qui ressemble le moins (ou le plus) possible aux points voisins, et ainsi de suite, jusqu'à l'obtention d'une bonne représentation des caractéristiques formelles de l'image (cf. Lowe, 2004)[298].

Par rapport aux méthodes des sciences humaines qui cherchent à analyser la dimension plastique des images ou les mécanismes formels qui survivent d'une époque à une autre, la *computer vision* se concentre uniquement sur l'extraction de motifs formels.

L'objectif est alors d'évaluer si un pixel est plus ou moins entouré de pixels similaires ou différents, à partir d'une représentation numérique de la surface de l'image. La programmation informatique repère les zones dans une matrice de pixel où des propriétés changent brusquement, mais cette représentation numérique est une réduction de l'image qui est traitée par la machine, (la *computer vision* conserve seulement les informations les plus pertinentes pour déterminer des éléments formels dans une image, et non pas tous les détails).

Comme nous l'avons déjà suggéré, l'utilisation du traitement informatique ne se limite plus seulement à la transmission d'images, il est aussi appliqué dans des processus d'analyse d'images[299]. L'*image processing* et la *computer vision* trouvent actuellement de l'écho dans des disciplines qui n'utilisaient que très peu des instruments computationnels ; par exemple, des travaux en histoire de l'art et en sémiotique cherchent actuellement à mettre en évidence les possibilités analytiques qu'offre l'informatique dans leurs recherches.

---

[298] Dans le cadre de la détection et la reconnaissance d'objets, on dénombre aussi plusieurs méthodes de détection : « mise en correspondance de graphes », « méthode bayésienne », « méthode des hypercubes », « réseau de neurones convolutif », « *local feature focus* », « *support vector machine* » (SVM) ; etc. Nous présenterons la méthode des réseaux convolutif, qui est désormais beaucoup utilisée dans le traitement automatique d'images.

[299] **Chap. 3, sect. 3.1.4. & 3.2.4.**

Dans le même sens, un autre changement important est directement lié aux progrès en informatique concernant les *big data* : il s'agit de combiner le traitement informatique à un *dispositif* qui présente les données (ou les images) qui sont traitées au sein d'une « vue d'ensemble ». Cette présentation des résultats d'un traitement informatique de données numériques est manifestée par le moyen de différentes *techniques de visualisation* (listes, diagrammes, séquences vidéo, animations en 3D, etc.).

4.2.1. Visualisation de données

La visualisation des données (*data visualization*) est un autre domaine où l'évolution des techniques informatiques offre de nouveaux moyens pour visualiser des informations et pour manipuler ces informations à des fins scientifiques (Yau, 2013 ; Reyes-Garcia, 2017). L'un des avantages des techniques de visualisation de données est qu'elles rendent possible un point de vue abstrait et schématique sur ces données. Couplée à un système informatique, la visualisation de données est actuellement utilisée pour organiser une grande quantité de données afin d'aider les humains durant le processus d'analyse de ces données. Comme les autres types d'interfaces, la visualisation de données comprend une dimension énonciative. Dans ce cas, l'énonciation vise à *représenter* et surtout à *mettre en contraste* des informations à travers des opérations de traitement qui révèlent des relations entre des données numériques. Ces techniques sont en effet appliquées dans différents domaines scientifiques, ou encore par des entreprises et par des institutions patrimoniales, afin de produire des techniques médiatrices qui tiennent un rôle dans l'argumentation et les discours scientifiques. Ces techniques participent à une rhétorique qui s'appuie sur l'idée qu'une représentation visuelle et abstraite permet de voir d'un seul coup d'œil des relations quantitatives au sein d'une grande quantité de données.

Même si des techniques de visualisation de données existe en effet depuis deux siècles[300], depuis la généralisation de l'informatique, la *data vizualization* a plus récemment popularisé différents modèles de visualisation pour présenter les résultats d'un traitement informatique d'un ensemble de données numériques. Ces « modèles » prennent souvent la forme de « systèmes graphiques » (cf. Bertin, 1967) qui sont le

[300] La visualisation de données apparaît au XVIII^e^ siècle et se développe durant le XIX^e^ siècle notamment à la suite des travaux de William Playfair (1786) et de Charles Joseph Minard (1862). Pour un point de vue plus récent qui prend en compte les avancées en informatique, cf. notamment Tufte (2001[1983]), Chen *et al.* (2008), Young *et al.* (2011) ou encore, en langue française, Lagnel (2021).

résultat d'un traitement statistique des données[301]. Dans ces systèmes, les données numériques sont présentées à l'aide de points, de lignes ou de barres afin de communiquer visuellement et efficacement une représentation quantitative des données, mais aussi pour faciliter la comparaison de ces données entre elles. Nous présenterons rapidement trois exemples diagrammes quantitatifs :

- Les *nuages de points*.
- Les *histogrammes*.
- Les *visualisations en réseaux*.

Les « nuages de points » (*scatter plot*) sont des diagrammes qui utilisent un système des coordonnées cartésiennes qui représentent différentes valeurs d'un ensemble de données en fonction de deux variables. À travers des processus de réduction, le but est de mettre en évidence des relations entre ces variables dans les dimensions d'un plan ; les nuages de points fonctionnent souvent sur l'algorithme t-SNE (*t-distributed stochastic neighbor embedding*) qui est une technique de réduction de dimension pour la visualisation de données développée par Geoffrey Hinton et Laurens van der Maaten (2008). En simplifiant, cet algorithme se base sur un processus stochastique qui aboutit

[301] Pour Jacques Bertin (1967 : 9-10), la « représentation graphique » est la transcription d'une information par le moyen d'un système graphique de signes. Il faut distinguer l'analyse préalable de l'*information* (le « contenu traductible d'une pensée » constitué par une ou plusieurs « correspondances originales entre un ensemble fini de concepts de variation et un invariant ») et les moyens (de représentation) du système graphique. Trois points sont nécessaires pour déterminer l'analyse d'une information : (1) le *choix du nombre de composantes* (les éléments (ou catégories ») sont les « différentes parties identifiables d'une composante ou d'une variable ») ; (2) la *longueur des composantes ou d'une variable* est « le nombre d'éléments ou catégories [que cette variable] permet d'identifier » ; (3) les *niveaux d'organisation* (niveau qualitatif ou combinatoire ; niveau de l'ordre ; niveau quantitatif) qui « forment le champ des significations universelles, des analogies fondamentales auxquelles peut prétendre la transcription graphique ». Dans cette perspective, l'ouvrage de Bertin fournit des notions et des explications précises sur le fonctionnement de certaines représentations graphiques (diagrammes, cartes, réseaux). Dans ces représentations, l'*invariant* est « la définition commune à toutes les correspondances originales », tandis que les *composantes* désignent « les concepts de variation mis en œuvre ». Les *variables* (visuelles) sont les « composantes du système graphique de signes ». Les *dimensions du plan* sont les deux variables produites par le plan. À travers une *implémentation* dans un plan, une tache peut soit représenter un « point », une « ligne » ou une « zone ». Cette « tache » peut ainsi recevoir trois significations par rapport aux dimensions du plan : (1) l'*implémentation ponctuelle* (x est représenté par un point) ; (2) l'*implémentation linéaire* (x est représenté par une ligne) ; (3) l'*implémentation zonale* (x est représenté par une zone). D'un autre côté, l'*imposition* est « l'utilisation des deux dimensions du plan [du système graphique] » ; cette utilisation est séparée en trois groupes en fonction de la manière dont les correspondances s'établissent dans le plan : (1) le *diagramme* (correspondances « entre tous les éléments d'une composante et tous les éléments d'une autre composante ») ; (2) le *réseau* (correspondances « entre tous les éléments d'une même composante ») ; (3) la *carte géographique* (correspondances « entre tous les éléments d'une même composante géographique inscrits dans le plan suivant la distribution géographique observée »).

sur une distribution de points dans un espace qui rapproche certains points et en éloigne d'autres. Cette technique représente la configuration optimale d'un ensemble de données dans un espace à deux ou trois dimensions. Chaque point est associé à un terme x et y qui détermine son emplacement dans le plan (fig. 3). Dans le système, les données sont représentées éventuellement par des points de couleur, de forme ou de taille différentes pour exprimer des variables supplémentaires. Dans le même sens, les « nuages de points tridimensionnels » (*scatter plot 3D*) représentent des relations entre variables dans un espace à trois dimensions (fig. 4).

Les histogrammes[302] (*histograms*) sont une représentation approximative de la distribution des données numériques. À partir d'un ensemble de données numériques, l'ensemble de valeurs est divisé en une série d'intervalles déterminés, de telle sorte que le nombre de valeurs comprises dans chaque intervalle est rassemblé dans un même « groupe » (fig. 5).

La « visualisation en réseaux » (*network* ; cf. fig. 6) met en évidence des clusters de données et des relations entre ces clusters en fonction de quelques variables (couleurs, tailles des points et des liaisons). La visualisation est donc envisagée comme une *aide* informatique qui offre des dispositifs d'analyse pour défendre des hypothèses sur la base d'un traitement statistique des données numériques. L'avantage de ces techniques est qu'elle offre une analyse computationnelle sans qu'il soit nécessaire de maitriser des méthodes statistiques et informatiques complexes pour pouvoir interpréter les visualisations des données.

Les approches théoriques de la visualisation des données cherchent encore à expliquer la visualisation et ses usages potentiels (voir par exemple les contributions dans Chen *et al.*, 2020). En effet, dans l'introduction de *Foundations of Data Visualization*, nous pouvons lire :

> Alors qu'un nombre croissant de recherches sur la visualisation se déploie – la visualisation de données est encore une discipline relativement jeune, comparée à d'autres comme la physique, la chimie, etc. – ce domaine cherche actuellement des explications sur les principes de la visualisation de données, ainsi que des fondements théoriques (*ibid.*, 2020 : 1).

[302] Voir aussi **supra, sect. 4.1.3.**

La première partie de cet ouvrage présente une discussion approfondie sur des questions théoriques centrales dans la recherche sur la visualisation : (1) l'interaction complexe des différents aspects de la visualisation de données ; (2) l'abstraction, plus précisément l'abstraction visuelle ; (3) comprendre ce que nous pouvons mesurer dans la visualisation et comment le mesurer ; (4) le rôle des connaissances préalables dans la visualisation ; (5) les fondements mathématiques importants de la visualisation ; (6) les concepts essentiels des *mappings* et des transformations dans la visualisation des données. Ces questions théoriques aboutissent sur des considérations concernant l'utilisateur et les actions que celui-ci peut réaliser à l'aide de la visualisation, ou encore sur des réflexions théoriques concernant les concepts mathématiques et technologiques qui permettent de visualiser des données.

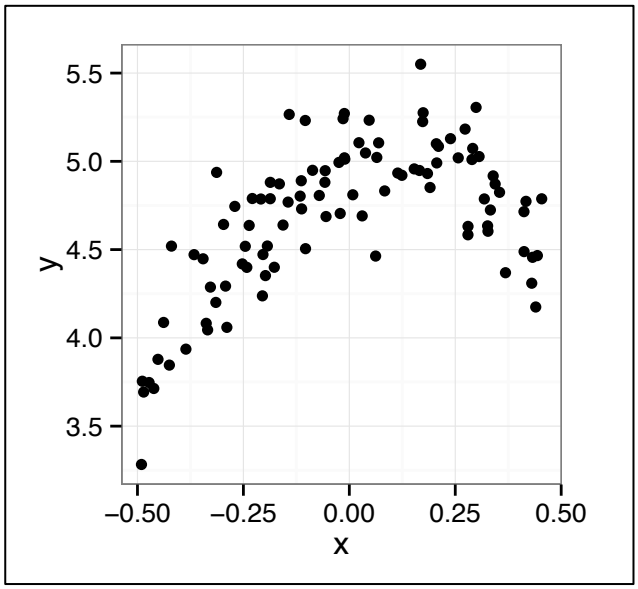


Figure 3 : Exemple de *scatter plot* (« nuage de points ») à deux variables.

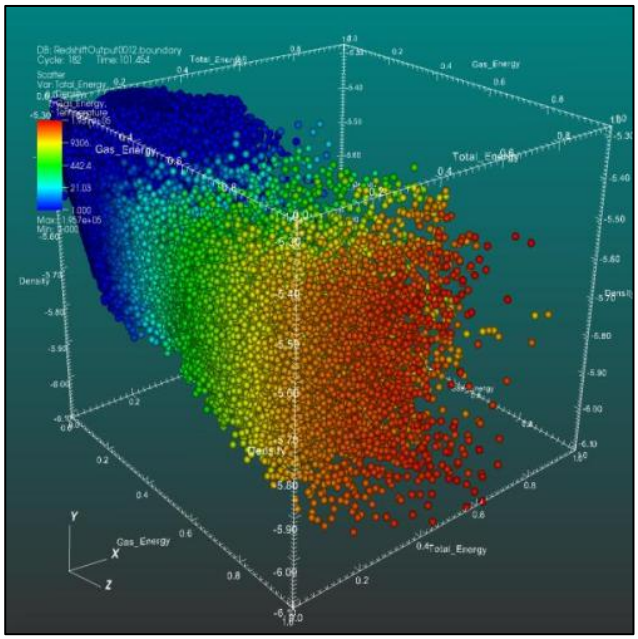


Figure 4 : Exemple de *scatter plot (3D)* ; cf. https://commons.wikimedia.org/wiki/File:Scatter_plot.jpg

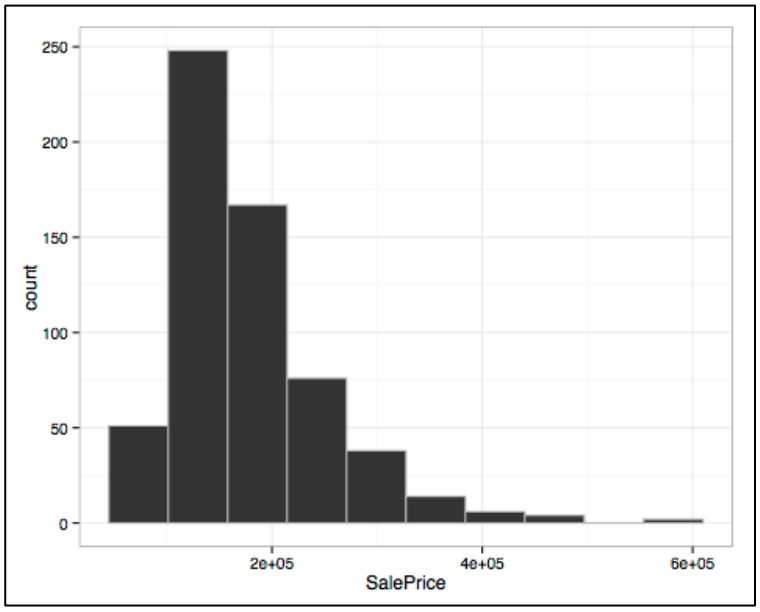


Figure 5 : Exemple d'histogramme (*histogram*) à deux variables ; cf. https://commons.wikimedia.org/wiki/File:Housingprice.png

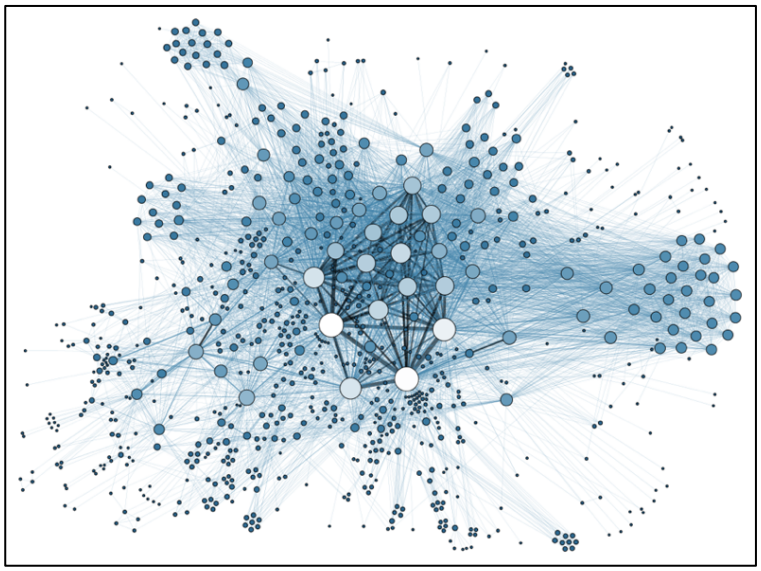


Figure 6 : Exemple de *network* (« visualisation en réseaux ») composé de différents points et liaisons ; cf. https://commons.wikimedia.org/wiki/File:Social_Network_Analysis_Visualization.png

Ces dispositifs présentent visuellement les résultats d'une analyse informatique. Ce point est très important, puisque des travaux en sémiotique proposent que les stratégies de composition d'une énonciation visuelle peuvent participer à la construction d'une activité autoréflexive, c'est-à-dire une réflexivité hors du cadre de la réflexion métalinguistique[303]. Nous pouvons remarquer que, lorsqu'elle est combinée à la raison computationnelle, cette réflexivité visuelle est par exemple observable à travers des formes abstraites propres au *raisonnement diagrammatique*.

À ce propos, Dondero a montré dans un travail sur l'image scientifique que la notion de diagramme peut aider à comprendre comment l'image et d'autres représentations visuelles sont susceptibles d'être « un support de travail qui permet de faire émerger de nouvelles formes et donc de nouvelles idées, nécessaires à la poursuite de l'enquête » (Dondero, 2010 : §10). Pour accorder un statut à l'image scientifique, il est nécessaire de considérer « la signification des techniques de production », ainsi que les manipulations que l'image scientifique rend possibles (*ibid.* : §5). En suivant Dondero, la notion de diagramme permet de considérer « le *processus exploratoire de la pensée* sur un objet qui est rendu saisissable à travers la *construction*, *l'observation et la manipulation eidétique d'une visualisation* », processus qui s'appuie la plasticité et la

[303] **Chap. 3, sect. 3.2.3. & 3.3.**

spatialité des visualisations (*ibid.* : §§52-53). L'auteure veut par là même mettre en évidence que l'image scientifique n'est pas une image à contempler, mais un « lieu d'expérimentation ». Ainsi, un instrument de visualisation élargit le point de vue par le moyen d'un processus exploratoire où le « raisonnement diagrammatique » sert à observer et à transformer les relations entre les parties d'un objet compris dans une globalité qui est toujours dynamique[304].

4.2.2. La voie ouverte par le *deep learning*

Dans les sections précédentes, nous avons vu que le traitement d'images se base sur un ensemble de techniques de programmation qui permettent de réaliser différentes tâches analytiques complexes (par exemple, d'identifier des formes, des objets dans une image numérique, de transcrire des paroles en un texte écrit, de rechercher des informations utiles dans une base de données, etc.). Par conséquent, associer les méthodes de l'analyse qualitative et celles de l'analyse quantitative offre évidemment de nouvelles manières pour analyser et pour visualiser des images (fixes ou animées). Cependant, comment cette association permet-elle de réfléchir les relations formelles entre des images qui sont classées automatiquement sur la base de représentations numériques ?

Le présent travail veut montrer que cette transition vers l'analyse computationnelle des images et du visuel ne doit pas aboutir sur le simple remplacement des anciennes méthodes par des instruments automatisés. Au contraire, si nous ne voulons pas risquer que les étapes et les résultats de ces analyses automatisées finissent par échapper aux spécialistes des images et du visuel, il faut comprendre dans quelle mesure l'informatique peut concrètement aider l'analyse d'images et les recherches sur le visuel.

De nombreux travaux affirment l'efficacité croissante des techniques d'*apprentissage automatique* qui se basent sur la combinaison du traitement informatique d'images avec les méthodes du *deep learning* (« apprentissage profond » ; cf. par exemple : Yang *et al.*, 2006 ; Nguyen & Worring, 2008 ; Schaefer, 2011).

Les méthodes du *deep learning* (et le cadre de réflexion théorique qui l'accompagne) sont connues depuis plusieurs dizaines d'années parmi les sciences de l'informatique. Cependant, à cette époque, les calculs informatiques sur de grands

[304] Nous verrons que cette fonction diagrammatique occupe de plus en plus une place centrale dans l'utilisation réflexive du numérique ; cf. **quatrième partie**.

ensembles de données étaient encore une tâche complexe et difficile. De plus, les larges bases de données informatiques n'étaient pas encore aussi nombreuses et disponibles qu'elles le sont actuellement.

Dans le domaine des sciences informatiques, nous pouvons souligner un regain d'intérêt pour ces méthodes parce que les compétences d'apprentissage et les performances des ordinateurs se sont beaucoup améliorées. Ces méthodes sont actuellement utilisées dans des analyses qui se basent sur l'extraction automatique de caractéristiques numériques qui émergent dans de larges ensembles d'images.

En effet, les ingénieurs en informatique développent actuellement des *réseaux neuronaux convolutifs* (ConvNet) qui sont des réseaux artificiels de neurones inspirés par le cortex biologique et adapté aux tâches de traitement de grands ensembles de données numériques (Goodfellow, Bengio & Courville, 2018 [2016] : 335 *sq*.). De plus en plus de travaux dans le domaine de la *computer vision* reposent sur l'exploitation de ces ConvNet, auxquels il s'agit d'apprendre à analyser et à caractériser plus ou moins automatiquement de grandes quantités d'images numériques[305]. Un ordinateur est capable d'effectuer une classification d'images en détectant des « caractéristiques de bas niveau » (*low level features*) telles que des bords et des courbes. Ces caractéristiques permettent d'élaborer des concepts plus abstraits par le moyen de ConvNet. Ceux-ci désignent un ensemble de processus informatiques qui superpose plusieurs « couches de traitement » (*layers*) et qui analysent des images dans le cadre d'un *apprentissage supervisé* (cf. **infra**). Cette superposition de couches rassemble plusieurs types d'opérations analytiques issues de l'*image processing* ; généralement, les ConvNet comprennent quatre opérations principales : (1) *Convolution*, (2) *Non lineary (ReLU)*, (3) *Pooling* ou *Sub Sampling*, (4) *Classification* (*Fully Connected Layer*).

Nous nous limiterons ici à quelques explications sur ces couches de convolution qui filtrent l'image. La première couche de convolution produit des identificateurs de caractéristiques de l'image (lignes, courbes, couleurs simples). Cette opération consiste à appliquer des filtres qui traitent successivement différentes régions de l'image en réduisant chaque région à une seule valeur numérique (ces filtres détectent des

[305] Pour un exposé théorique détaillé sur l'*apprentissage profond*, voir notamment LeCun, Bengio & Hinton, 2015 ou Goodfellow, Bengio & Courville (2018 [2016] : 115-178). Pour des travaux qui concernent la classification d'images à l'aide de ces techniques d'apprentissage, cf. Donahue *et al.*, (2014), Razavian *et al.* (2014) ou Simonyan & Zisserman (2014).

« caractéristiques de bas niveau », même si ce type de traitement informatique d'images ne se limite pas à ce type de caractéristiques). Passant d'une couche de filtrage à une autre couche de filtrage, l'*output* de la première couche devient l'*input* de la deuxième couche. Or l'*input* de la première couche de filtrage est la *surface de l'image originale*, tandis que l'*input* de la deuxième couche de filtrage n'est plus une image, mais les *résultats du traitement de la première couche* (détection des caractéristiques générales dans l'image en fonction de valeurs numériques).

Plusieurs couches sont enchainées, et chaque couche agit comme un filtre qui calcule une réponse scalaire (une réponse en rapport avec une échelle de grandeur comprenant différents « degrés ») pour chaque pixel de l'image. Une opération (*activation function*) est appliquée à chacune de ces réponses, ce qui permet de créer une « carte de caractéristiques » (*feature map*) par filtre/couche. Couche par couche, les ConvNet transforment une image originale en pixels, pour finalement aboutir à des résultats sous la forme de classes. En effet, la première couche de traitement consiste généralement à produire une image binaire, tandis que la dernière couche propose des tenseurs qui attribuent différentes valeurs à l'image par rapport à des classes prédéfinies, des « étiquettes » (*tags*) correspondantes (cf. fig. 7).

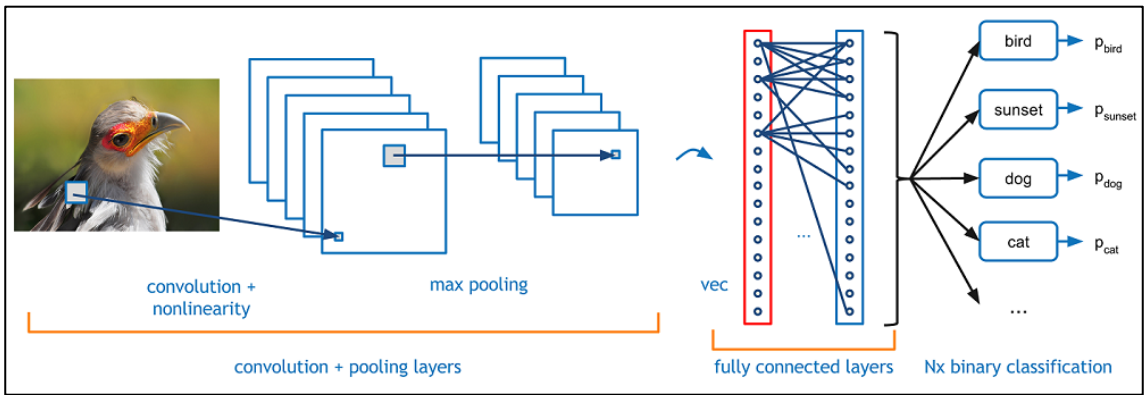


Figure 7 : Schéma représentant les principales opérations utilisées dans le traitement d'images à l'aide de *ConvNet* ; cf. https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/. L'image est traitée, d'une part, en fonction des caractéristiques plastiques de l'image originale, et d'autre part en fonction de paramètres qui analysent les résultats du traitement pour que ces résultats soient cohérents par rapport à des étiquettes déjà déterminées.

Ce processus d'analyse stratifiée filtre l'image transformée en valeurs numériques en fonction de paramètres déterminés à l'avance. Progressivement, la superposition des couches et la combinaison des résultats du traitement font ressortir des caractéristiques numériques de plus en plus complexes, jusqu'à détecter des formes particulières dans une

image. Chaque couche est reliée avec les opérations de la couche de traitement précédente, et ainsi de suite. L'analyse est, en quelque sorte, à la fois *continue* et *discontinue*.

Ces ConvNet obtiennent de bons résultats en ce qui concerne des tâches de classification d'images (cf. Goodfellow, Bengio & Courville, *op. cit.* : 450-455). Grâce à la programmation du type *General-purpose processing on graphics processing units* (GPGPU), il est désormais possible de programmer un traitement informatique avec plusieurs millions de paramètres d'analyse par le moyen d'architectures informatiques complexes. L'avantage indéniable de cette programmation est de combiner plusieurs opérations d'analyse et de les appliquer systématiquement sur de larges corpus d'images binaires. Rappelons-le, d'une part, l'objectif est avant tout la prédiction et surtout la vitesse. D'autre part, les opérations sont réalisées sur des collections de nombres et non pas sur les images en tant que telles.

Pour améliorer ces instruments et méthodes d'analyse des images, il s'agit donc d'affiner les étapes du protocole informatique, autrement dit : d'inventer des algorithmes plus performants (ce qui n'est pas de notre ressort). C'est pourquoi les informaticiens travaillent actuellement à améliorer ces techniques d'analyse en distinguant *deux techniques d'apprentissage* :

- L'apprentissage supervisé, qui vise à produire automatiquement des règles selon une base de données d'apprentissage qui contient des exemples d'images qui sont déjà classées.
- L'apprentissage non supervisé, ne nécessitant pas d'exemples de classement, celui-ci étant établi par une programmation informatique.

L'*apprentissage supervisé* consiste à apprendre à un programme une fonction de prédiction d'après une série d'exemples d'indexation « classique », donc des images annotées selon des catégories verbales. Ces exemples permettent d'établir une base d'apprentissage. Le but est par la suite de généraliser les rapports entre les images et les catégories, en partant du principe que cette base d'apprentissage est représentative d'une population d'échantillons plus large, et surtout que la machine produira des prédictions correctes sur des données non présentes dans cette base d'apprentissage.

D'un autre côté, l'*apprentissage non supervisé* consiste à apprendre sans superviseur humain. Afin d'extraire des classes ou groupes qui présentent des

caractéristiques communes, un logiciel doit détecter des structures sous-jacentes à partir de données qui ne sont pas étiquetées ou catégorisées. Autrement dit, il ne s'agit pas d'apprendre à un logiciel à classer un nouvel élément parmi un ensemble de classes prédéfinies (car les classes ne sont pas données a priori), mais c'est la machine qui doit elle-même les déterminer.

4.2.3. *Datasets* et classification d'images : persistance du langage ?

Nous avons présenté, dans les grandes lignes, comment un ordinateur exécute différentes opérations sur la base d'une programmation informatique et d'algorithmes. Différentes opérations d'analyse sont appliquées sur des images binaires (restauration, amélioration, segmentation, détection de formes, classification). Ces analyses concernent uniquement des valeurs numériques, même s'il s'agit de reconnaître des formes visuelles dans des images.

Le domaine des *data-driven approachs* (« approches pilotées par des données ») regroupe les recherches qui combine des langages de programmation et des algorithmes pour analyser de larges ensembles de données et pour faire ressortir des constantes ou des « *patterns of interest* » à l'aide d'un environnement informatique. L'évolution de ce domaine de recherche offre des perspectives prometteuses dans le domaine de l'analyse de larges ensembles d'images numériques[306] ; d'ailleurs, nous avons montré que des travaux en histoire de l'art et en sémiotique ont pris acte de ces nouvelles techniques d'analyse[307]. En effet, un ordinateur « bien entraîné » peut identifier des *régions* ou *zones d'intérêts* en fonction d'une quantification complète de la microtopographie des images. Ce processus d'apprentissage extrait automatiquement des caractéristiques à partir d'un corpus d'images et il est souvent couplé à des catégories (ou des classes) qui correspondent à des descriptions en mots d'images.

Sur ce sujet, nous avons vu que l'indexation d'images favorise l'utilisation de descriptions verbale pour organiser des archives d'images[308]. À titre d'exemple, l'inventaire d'archives d'images constitue une forme de « métadonnées descriptives

[306] À ce propos, cf. entre autres : Crowley & Zisserman (2014) ; Karayev *et al.* (2014) ; Schelbert (2017) ; Manovich (2012, 2017a & 2017b) ; Ezhilraman & Srinivasan (2018).

[307] **Chap. 3, sect. 3.1.4. & 3.2.4.**

[308] **Chap. 3, sect. 3.3.**

structurées » qui précède les métadonnées intégrées dans les bases d'images numériques. Le but de l'indexation est d'administrer des ressources en fonction d'un cadre normatif qui aboutit sur des informations stables. Aux opérations de traitement documentaire et d'indexation s'ajoute la conversion en données numériques des archives (d'une part pour disposer de copies des documents et des métadonnées ; d'autre part pour rendre accessibles ces ressources sur Internet)[309]. Cependant, nous avons également suggéré que cette manière de structurer de larges ensembles d'images est problématique, puisque les relations entre les images dépendent d'un paradigme d'interprétation des images qui vise uniquement à faire correspondre des images avec des mots-clés. Nous verrons pourtant que ces métadonnées descriptives sont largement utilisées par les institutions patrimoniales pour classer des images dans leurs services[310].

En effet, le principal défaut de la classification par mots-clés est que les descriptions utilisées ne reflètent pas les caractéristiques visuelles et les formes d'organisation propres aux images. Or nous avons montré que des progrès dans les possibilités de traitements des données ont poussé des projets scientifiques à concevoir des techniques de plus en plus sophistiqués pour indexer, récupérer, organiser et annoter des données visuelles. Ces recherches plus récentes proposent plutôt d'utiliser des *descripteurs visuels* pour analyser de grandes collections d'images à l'aide de techniques basées sur l'extraction de leurs caractéristiques plastiques (l'intensité de la couleur, les contours et les positions des formes), ou en utilisant des méthodes avancées qui s'appuient sur des algorithmes (*deep learning*).

Nous remarquerons cependant dans cette section que la question de la qualité des métadonnées descriptives persiste dans ces processus de classification automatique d'images ; notamment parce que les techniques d'apprentissage se basent encore sur une relation forte entre visuel et métalangage descriptif.

Le fait que la classification d'images est encore fondée sur un métalangage de description est observable dans un projet comme *ImageNet*, un jeu de données[311] (*dataset*) mis en place par *Stanford Vision Lab* (Université de Stanford et Université de Princenton ;

[309] **Chap. 2**.

[310] Voir **chap. 5 & 6**.

[311] Un *dataset* est un ensemble de données qui sont organisées en fonction de variables qui décrivent le même attribut.

cf. Deng, *et al.*, 2009). *ImageNet* est avant tout une ressource pour la communauté scientifique, plus précisément, pour les informaticiens qui développent des techniques de classification automatique d'images. Ce *dataset* est organisé en s'appuyant sur le modèle de *WordNet*[312], au sein duquel un concept peut être décrit par plusieurs mots ou expressions de mots qui sont appelés « ensemble de synonymes » ou *synset*. *WordNet* comprend plus de 100 000 *synsets* dont la plupart (80 000) comportent des noms. Les images sont soumises à un contrôle et sont annotées par l'humain. L'objectif est de fournir en moyenne 1000 images pour illustrer chaque *synset*. Notons qu'*ImageNet* ne possède pas les droits d'auteur des images qui sont organisées et structurées dans la base de données : *ImageNet* ne fournit que des vignettes et des adresses Web (URL) d'images ; autrement dit, *ImageNet* est une compilation, une liste précise d'images disponibles sur Internet, images qui correspondent en outre à des ensembles de concepts linguistiques (*WordNet*).

L'hypothèse est qu'*ImageNet* offrira bientôt des dizaines de millions d'images correctement triées et correspondant à la plupart des concepts de *WordNet*. La mise en place de ce type de *datasets* est la conséquence d'un besoin croissant en données structurées dans le domaine des recherches en informatiques sur l'image et le visuel (cf. **supra**, sur « l'apprentissage supervisé »). Sur ce sujet, des travaux discutent cette stratégie d'annotation par étiquettes (*labels*) des larges ensembles d'images numériques. Une étude sur l'annotation de 200 objets à partir d'un ensemble de 20 000 images a permis à des chercheurs de montrer l'efficacité d'un algorithme qui exploite la corrélation, la hiérarchie et la rareté de la distribution des étiquettes. Cet algorithme permet de réaliser plus rapidement des classifications automatiques d'images par rapport aux méthodes qui se basent sur une annotation humaine (cf. Deng *et al.*, 2014).

Une enquête plus récente sur ces stratégies fait ressortir « deux formations idéologiques distinctes, mais liées » (cf. Denton *et al.*, 2021 : 8). D'une part, l'accumulation de données comme une fin en soi et, d'autre part, la production d'une ressource collective et informatisée. En analysant les discours des principaux textes relatifs à la création et à l'impact d'*ImageNet*, cette enquête a mis en évidence trois thématiques générales : (1) l'agrégation et l'accumulation d'un maximum de données

[312] https://wordnet.princeton.edu.

numériques ; (2) la construction computationnelle du sens ; (3) l'invisibilité des opérations sur les données.

Partant des données fournies par *ImageNet*, la conception spécifique de ce *dataset* a un impact sur l'authenticité de l'ensemble de données. En effet, ces jeux de données hétérogènes sont encadrés par des stratégies d'efficience et de rentabilité de l'annotation, et surtout ces stratégies transforment progressivement cette annotation en une tâche désincarnée et décontextualisée. Les techniques informatiques qui sont employées pour collecter des données aboutissent parfois sur « un désalignement systématique entre le repère résultant et la tâche du monde réel qu'il sert à représenter » (*ibid.*). Il faut prendre en compte ce désalignement, car les annotations d'*ImageNet* s'éloignent progressivement des annotations habituellement réalisées par des humains. Ces déviations ont un impact significatif sur l'apprentissage des modèles entraînés par *ImageNet* et sur la manière dont nous percevons la progression du modèle. Selon cette enquête (*ibid.*), la précision de ces modèles est problématique. Même si les prédictions sont satisfaisantes dans les modèles les plus performants, il semble nécessaire de mettre en place des mesures d'évaluation humaine des processus automatisés afin de vérifier la précision des prédictions effectuées par le traitement informatique. Afin de mettre en évidence « les moments où des interventions réflexives peuvent être faites dans le développement de jeux de données », un questionnement sur le rôle du langage dans le traitement informatique est incontournable si nous voulons développer « des cadres holistiques et prospectifs permettant aux utilisateurs des données d'analyser de manière réflexive les éléments de leur pipeline de données » (*ibid.*)[313].

À ce stade, le traitement informatique de larges ensembles d'images nécessite une standardisation des métadonnées qui permet aux ordinateurs d'apprendre à classer des images. D'autres travaux examinent dans le même sens les conséquences de l'utilisation de la programmation informatique (cf. Tsipras *et al.*, 2020). D'aucuns craignent que le contexte actuel favorise un désintérêt de la valeur documentaire des archives d'images au profit de larges ensembles d'images prédisposées à une *logique d'exploitation* (cf. Birkin, 2020). Cette logique d'exploitation se base sur des ensembles de données structurées pour entraîner les ordinateurs et pour éprouver leurs instruments de classification. Les archives

[313] Un *pipeline* est une succession d'étapes composant un programme informatique. Voir aussi les autres travaux sur le développement et la transparence des modèles et des données : Gebru *et al.* (2018) ; Mitchell *et al.* (2019) ; Hutchinson *et al.* (2021).

et collections d'images numériques se transforment progressivement en un vaste terrain d'expérimentation pour « tester » les performances des instruments informatiques.

Somme toute, l'influence de la *data science*, de la *computer vision* et du *deep learning* dans l'étude des images aboutit sur une situation que nous continuerons d'analyser dans la suite de ce travail.

L'utilisation de ces techniques computationnelles pour analyser de larges ensembles d'images numériques d'une part, et la mise en place de techniques de visualisation des résultats de ce traitement automatisé d'autre part, ces techniques indiquent de nouvelles pratiques à venir. Plusieurs projets scientifiques étudient déjà les applications possibles de ces techniques, notamment pour classer des images à partir de leurs métadonnées descriptives ou de leurs caractéristiques visuelles.

Dans la section suivante, nous présenterons rapidement deux services Web qui permettent de chercher des images en fonction de *similitudes visuelles* ; nous présenterons d'autres projets qui se basent sur ce principe de « similitude visuelle » dans la **quatrième partie**.

4.2.4. Recherche d'images en fonction de la similitude visuelle

Le moteur de recherche créé par l'entreprise *Google* est sans aucun doute un service inévitable. Ce moteur de recherche comprend également *Google Images*[314], un moteur de recherche d'images. Ce service de recherche propose aussi une fonctionnalité pour effectuer des *recherches par images*. Contrairement à la *recherche par mots-clés*, cet outil ne contraint pas l'utilisateur à encoder des mots-clés dans le champ de recherche, mais lui permet d'utiliser directement une image comme un *input pour la recherche*. L'utilisateur peut proposer un fichier personnel (à importer dans le navigateur Web) ou utiliser l'adresse URL d'une image déjà disponible sur Internet. Ce dispositif analyse ensuite l'image pour détecter les éléments (couleurs, points, lignes et textures) les plus distinctifs. Le service utilise ces caractéristiques pour générer une demande de recherche

[314] https://images.google.com.

qui est envoyée aux *back ends* de *Google*[315]. Ce processus permet à l'image d'être comparée à des milliards d'autres images à l'aide d'un algorithme qui renvoie à l'utilisateur du service les images correspondantes et visuellement similaires sur la page des résultats finaux. Ces résultats peuvent ainsi inclure : (1) des informations concernant les objets dans l'image, (2) des images similaires, (3) des sites Web contenant l'image ou une image similaire.

En fonction des demandes de l'utilisateur, le moteur de recherche présente plus ou moins de résultats. Le service fonctionne assez bien pour retrouver certains types d'images en fonction de la composition formelle et des couleurs. Par exemple, nous proposons une image (*Sainte Anne, la Vierge et l'Enfant Jésus jouant avec un agneau* (ca. 1503-1519) de Leonardo de Vinci) comme *input*. En retour, le moteur de recherche présente une sélection d'images avec des liens vers quelques sites Web où cette image apparaît (cf. fig. 8). Il s'agit chaque fois de différentes reproductions de la même peinture ; sauf dans un seul cas, qui propose une autre image « similaire ».

Les résultats sont par contre différents lorsque d'autres types d'images sont proposées ; par exemple : une nature morte de Peter Jacob Horemans (ca. 1774 ; cf. fig. 9) ou une photographie du Stadsbrug de Kampen aux Pays-Bas (fig. 10). Dans le premier cas, le moteur de recherche propose différents tableaux du même genre (nature morte). Dans le second, le moteur de recherche retrouve différentes photographies du même monument (le pont).

[315] En informatique, un *back-end* désigne un niveau de sortie d'un logiciel dont le but est de produire un résultat ; le *back-end* s'oppose au *front-end* qui lui est la partie visible du processus informatique ; voir aussi **chap. 1 (sect. 1.3.3.) & chap. 2 (sect. 2.3.)**.

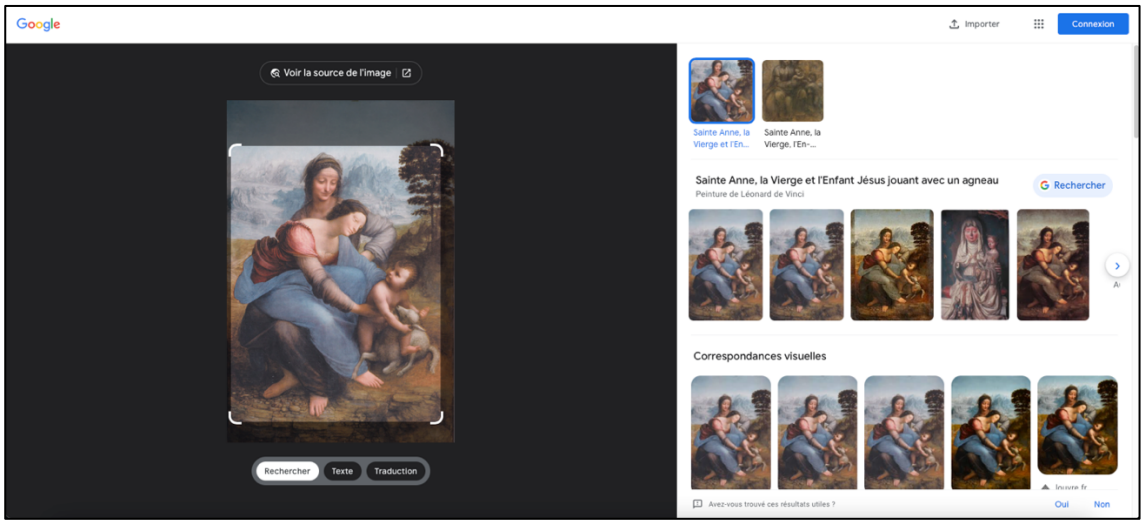


Figure 8 : Exemple de recherche par similitude visuelle (*Google Images*) ; à gauche, l'image utilisée pour la recherche, et à droite les résultats du moteur de recherche.

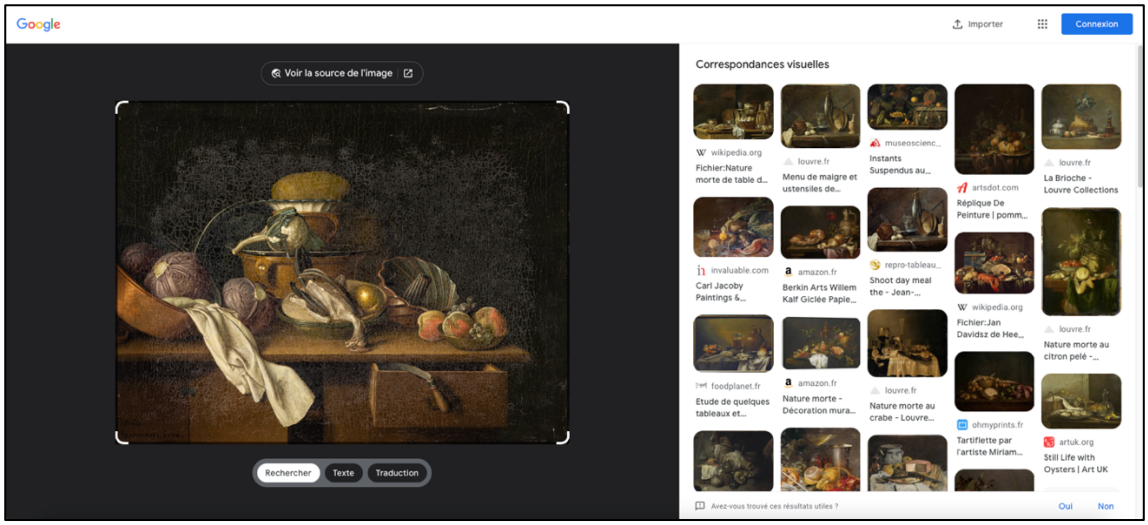


Figure 9 : Autre exemple de recherche par similitude visuelle (*Google Images*).

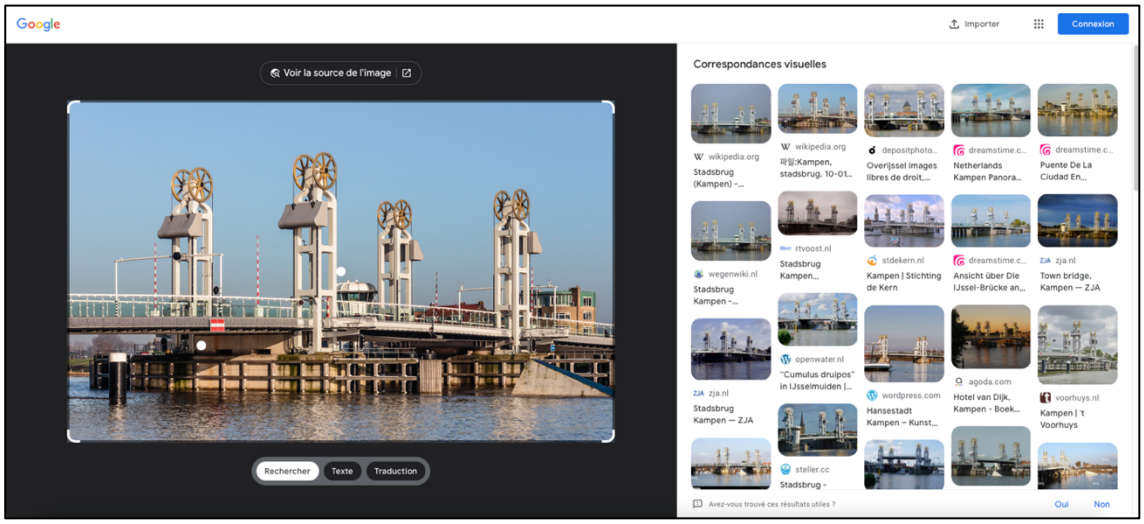


Figure 10 : Autre exemple de recherche par similitude visuelle (*Google Images*).

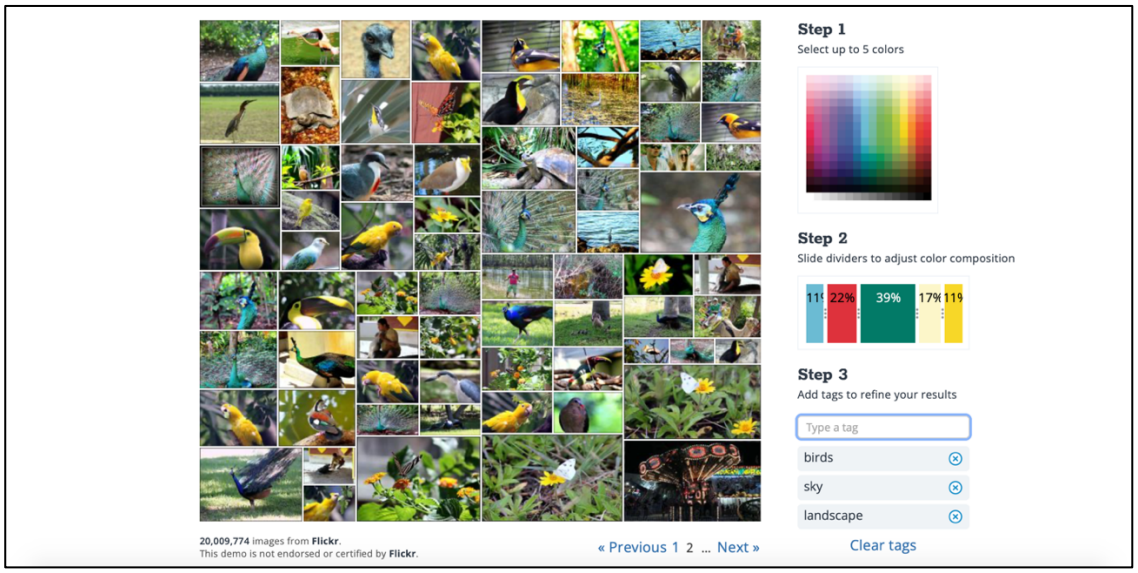


Figure 11 : Exemple de recherche en fonction des proportions de couleurs dans des images (*Multicolr*).

Dans la même perspective, *TinEye*[316] est une société canadienne créée en 2008 qui propose un service de recherche et de reconnaissance d'images. Cette société privée est spécialisée dans le traitement d'images qui se base sur la reconnaissance des formes à partir des réseaux neuronaux et de l'apprentissage automatique. Pour ce faire, cette société a conçu différentes applications pour recherche des images en fonction d'autres images, c'est-à-dire une alternative à la recherche en fonction de mots-clés. *TinEye* permet à un utilisateur d'effectuer une recherche par image (*reverse image search*) à partir de plusieurs milliards d'images indexées. Le but est de retrouver des images similaires sur Internet, soit à partir d'une image personnelle ou à partir d'un lien (une adresse URL) qui renvoie à une image accessible sur Internet. Les résultats de la recherche sont éventuellement filtrés pour réduire le nombre d'images sélectionnées, mais aussi comparées afin de mettre en évidence des différences entre des images similaires. En outre, cette société propose d'autres services.

Par exemple, en collaboration avec *MulticolorEngine*[317], *TinEye* propose l'application *Multicolr* (cf. fig. 11). Cette application permet à l'utilisateur de réaliser une recherche en fonction de couleurs. Le *MulticolorEngine* est utilisé pour extraire les couleurs de 20 millions d'images *Creative Commons* sur *Flickr* dans le but de rendre ces images consultables par le moyen d'une recherche par couleur. Le dispositif de recherche

[316] https://tineye.com.

[317] *MulticolorEngine* est un service Web qui analyse les pixels des images plutôt que les métadonnées descriptives ou les autres informations qui sont associées à des images.

Multicolr fonctionne en trois étapes : (1) choisir une ou plusieurs couleurs dans une palette graphique (maximum cinq couleurs) ; (2) ajuster les proportions des couleurs ; (3) ajouter des mots-clés (*tags*) pour affiner le résultat. Ci-dessus, exemple de sélection de résultats d'une recherche à partir de cinq couleurs et de trois tags (oiseaux, ciel, paysage).

Ceci nous permet déjà de remarquer que la recherche par similitude visuelle est un outil très intéressant, même s'il ne permet pas un paramétrage précis de la recherche. Si les fonctionnalités de recherche de *Google Images* ou de *TinEye* sont certes rudimentaires, elles mettent déjà en évidence le potentiel des méthodes assistées par ordinateur.

Pour finir, notons que l'impact de la reconnaissance des formes dans les images ne se limite pas au domaine des larges ensembles d'images. Nous la retrouvons plus récemment dans les smartphones. En effet, dans l'application *Photos* d'un iPhone, l'utilisateur peut réaliser une recherche visuelle pour identifier des objets dans les photographies ou les vidéos. Cette recherche permet d'identifier « les sites touristiques populaires, les plantes, les animaux domestiques, etc. qui apparaissent dans vos photos et vos vidéos »[318]. Grâce à ce smartphone, l'utilisateur peut aussi détacher un objet de l'arrière-plan d'une photographie ou d'une vidéo. L'utilisateur peut ensuite transformer l'objet détecté en *sticker* (« autocollant ») et surtout « le copier et le partager dans d'autres documents et apps »[319].

4.3. Conclusions

Nous arrivons au terme de cet état de l'art. Les **deux premiers chapitres** proposent une synthèse (évidemment incomplète) qui prend en compte les acquis de diverses disciplines au sujet de la conversion numérique et de la communication en ligne des images. Tandis que les **deux chapitres précédents** présentent les méthodes d'analyse d'images issues de l'histoire de l'art et des sciences du langage et les méthodes d'analyse assistées par un ordinateur.

Dans ce **quatrième chapitre**, nous avons vu que des algorithmes complexes et puissants sont mis en place pour analyser des images numériques en grande quantité. La

[318] Voir https://support.apple.com/fr-be/guide/iphone/iph21c29a1cf/ios.

[319] Voir https://support.apple.com/fr-be/guide/iphone/iphfe4809658/ios.

principale différence entre les approches en sciences humaines et en informatique est, d'une part, la manière de définir l'image, d'autre part, les opérations appliquées à l'image pour les analyser et les classer. Ces acquis concernant l'analyse d'images sont importants, car ces opérations informatiques sont actuellement utilisées dans une perspective analytique pour mettre en place des instruments de classification d'images numériques en fonction de la reconnaissance de formes ou de similitudes visuelles.

La discussion sur l'analyse des formes visuelles[320] débouche sur une question : puisque ces méthodes informatiques permettent d'appliquer un même processus d'analyse sur un grand nombre d'images à la suite de phases d'apprentissage de la machine sur des « échantillons tests » d'images, faut-il envisager que ces méthodes sont une solution possible en ce qui concerne l'indexation d'images et l'analyse des formes visuelles ?

L'utilisation du traitement informatique pour analyser des formes visuelles est actuellement limitée par trois obstacles. D'une part, ces instruments prennent seulement en compte des moyennes statistiques de caractéristiques visuelles, parce qu'ils sont justement programmés pour faire abstraction de la configuration réelle des éléments qui composent les images. D'autre part, ces similitudes visuelles identifiées ne produisent pas une analyse des formes visuelles, mais elles mettent seulement en évidence un système de similitudes. Enfin, plusieurs travaux soulignent la même nécessité de transiter par un langage de contrôle (verbal ou informatique) pour encadrer le traitement informatique des images.

Il faut donc que les sciences humaines prennent conscience du fait que, en s'éloignant volontairement de l'emprise d'un langage[321] au profit des techniques informatiques, ces mêmes sciences humaines risquent en effet de tomber sous l'emprise d'un autre langage. Dans cette perspective, nous comprenons mieux pourquoi l'un des enjeux les plus importants est effectivement celui du *semantic gap*[322]. Une conception du sens se mesure à une autre : la première est un *sens interprétable*, la seconde un *sens calculable*. Cependant, il faut rappeler que le but de la raison computationnelle est d'être au service de l'humain, c'est-à-dire qu'elle doit rester intelligible pour lui. Dans tous les

[320] Voir **chap. 3, sect. 3.4.**

[321] Voir **chap. 3, sect. 3.5.**

[322] Voir **supra, sect. 4.1.3.**

cas, nous avons suggéré que le dispositif de lecture est le point de rencontre entre ces deux conceptions du sens. En effet, l'expérience des ressources numériques repose toujours sur une interface, qui est la « face visible » du processus d'interaction entre l'humain et la machine.

Certes, nous avons vu que l'objectif de l'analyse computationnelle est d'abord de mettre en valeur les capacités de la technologie numérique plutôt que de répondre à des questions spécifiques sur les arts visuels et l'analyse d'images. Toutefois, ce passage à une analyse computationnelle peut bénéficier de certains acquis mis en évidence par les recherches en iconologie et en sémiotique sur le visuel ; en particulier des travaux sur les limites du principe de métalangage descriptif lorsque celui-ci est appliqué au domaine visuel et, par conséquent, l'idée que les images et le visuel comprennent une dimension autoréflexive qui ne se fonde pas nécessairement sur la réflexivité du langage. Dans les **chapitres suivants**, nous verrons en outre qu'une approche sémio-rhétorique de la communication en ligne aide à mieux comprendre l'*énonciation* et la construction du *point de vue* qui sont mises en scène dans les interfaces Web qui donnent à voir un large ensemble d'images.

Dans la **troisième partie**, avant de considérer le cas des techniques de visualisation dans le contexte des *big visual data*[323], nous voulons d'abord présenter plusieurs exemples d'interfaces Web mises en place par des institutions patrimoniales. Nous regroupons ces services dans une *première tendance stratégique* qui utilise l'informatique pour diffuser et pour manipuler à distance des archives d'images. Nous présentons d'abord ces dispositifs d'exploration mis en place par des institutions patrimoniales[324] pour deux raisons : d'une part, parce que ces dispositifs sont toujours employés par des institutions patrimoniales, et d'autre part, parce que les limites évidentes de ces premières stratégies institutionnelles ont contribué à un renouvellement des interfaces et de leurs fonctionnalités[325].

[323] **Chap. 8 & 9**.

[324] **Chap. 5**.

[325] **Chap. 6 & 7**.

TROISIÈME PARTIE

ANALYSE DU CORPUS (1) : BASES D'IMAGES NUMÉRIQUES ET EXPOSITIONS VIRTUELLES

CHAPITRE 5
Première tendance stratégique : les bases d'images numériques des institutions patrimoniales

Dans ce **cinquième chapitre**, nous étudions une première tendance stratégique qui concerne les bases d'images numériques. Lorsque des institutions patrimoniales utilisent le réseau Internet pour transmettre des stocks d'images numériques[326], la communication en ligne se base généralement sur un service et sur une interface, c'est-à-dire le support intermédiaire permettant d'utiliser le service[327]. Comme nous l'avons dit, nous considérons cette communication comme un processus de médiation ouverte[328] qui nécessite la participation active de l'utilisateur pour construire du sens avec le système informatique. En outre, ce processus comporte au moins trois caractéristiques :

- L'expérience de l'utilisateur s'appuie sur un dispositif de lecture qui permet de manipuler des ressources numériques[329].
- Ce dispositif de lecture combine souvent une pluralité d'objets médiatiques dans une même page-écran (structure modulaire)[330].

[326] **Chap. 2, sect. 2.2.1.**

[327] **Chap. 2, sect. 2.3.1.**

[328] **Chap. 2**, **sect. 2.2.2.**

[329] Cf. **chap. 1 (sect. 1.4.) & chap. 2 (sect. 2.3.3.)**.

[330] Voir **chap. 4, sect. 4.1.2.**

- Les stratégies de communication en ligne des larges ensembles d'images visent à aménager l'expérience des internautes à travers une relation d'énonciation et la construction d'un point de vue[331].

5.1. Des archives et des collections d'images aux bases d'images numériques

Nous avons vu que la photographie tient un rôle central dans la multiplication des archives en offrant un dispositif technique pour reproduire visuellement le patrimoine[332]. Dans le même sens, la numérisation intensive de ces photographies a favorisé l'émergence de plusieurs *bases d'images numériques* sur le patrimoine culturel. Cette importante quantité d'images constitue des ressources utiles pour la recherche scientifique dans la même perspective que les archives ou les collections de photographies qui ont été utilisées pour consigner des informations utiles pour étudier et mieux comprendre le passé ; nous verrons bientôt à quel point cet impact de la photographie est observable dans les ressources numériques disponibles dans les bases d'images des institutions patrimoniales en Europe, mais aussi ailleurs dans le monde.

Or il faut admettre que, passant du support original à une reproduction photographique, puis à une reproduction numérique des photographies, ce n'est pas tant le contenu d'une archive qui se transforme en un autre contenu, mais c'est plutôt la manière de voir et de manipuler à la fois l'archive et son contenu qui se transforme. L'expérience du patrimoine et des œuvres d'art s'appuie donc sur une double reproduction technique dont le but est de renouveler les *conditions de visibilité et de manipulation* du patrimoine culturel par le moyen de substituts, de copies.

Ces remarquent suggèrent plusieurs questions : sur quels objectifs se fondent ce second dispositif de reproduction des archives d'images ? Quels types de scénarisation des ressources numériques le dispositif de communication en ligne favorise-t-il ? Quels rôles énonciatifs occupent les institutions produisant des services Web et l'utilisateur de ces services ? Dans ces services, comment les images sont-elles classées et présentées à

[331] Cf. **chap. 1 (sect. 1.3.3.) & chap. 3. (sect. 3.2.3.)**

[332] **Chap. 1**, **sect. 1.1.**

l'aide d'un environnement informatique ? Enfin, quelles sont les fonctionnalités d'interaction prévues dans l'interface ?

Pour répondre à ces questions, nous présenterons quatre institutions patrimoniales qui proposent chacune un site Web pour diffuser un large ensemble d'images sur Internet. Ce genre d'initiative est évidemment observable au sein de nombreuses autres institutions qui proposent des solutions similaires pour diffuser en ligne des archives et des collections d'images numériques[333]. Dans une certaine mesure, les remarques qui sont présentées dans les sections suivantes peuvent être transposées sur ces autres services Web.

5.1.1. *Plateforme ouverte du patrimoine*

Le département photographique de la *Médiathèque de l'architecture et du patrimoine* (MAP) est l'un des plus grands ensembles d'archives photographiques en Europe. Ce département comprend plusieurs millions de documents photographiques[334]. Les archives photographiques de la MAP couvrent un large panel thématique : sites et monuments, expéditions archéologiques, événements historiques, arts et spectacles. En outre, le regroupement de ces archives centralise la gestion et la conservation des négatifs et tirages de nombreux photographes au sein d'une même institution, ce qui permet de favoriser leurs diffusions (après numérisation) sur Internet.

---

[333] En France, nous pouvons évoquer la banque d'images de plus d'un million d'images de la Bibliothèque nationale de France (https://images.bnf.fr/#/home/), ou encore les collections d'images numériques de l'Institut National d'Histoire de l'art (https://bibliotheque-numerique.inha.fr/collection-toutes) ; cf. aux États-Unis : *Museum of Modern Art* (https://www.moma.org/collection/), *Guggenheim Museum* (https://www.guggenheim.org/collection-online), *The Metropolitan Museum of Art* (https://www.metmuseum.org/art/collection) ; au Royaume-Uni : *Tate Museum* (https://www.tate.org.uk/art), *The British Museum* (https://www.britishmuseum.org/terms-use/copyright-and-permissions/images-and-photography) ; en Norvège : *Musée national de Norvège* (https://www.nasjonalmuseet.no/en/collection/) ; en Nouvelle-Zélande, *Museum of New Zealand* (https://collections.tepapa.govt.nz) ; en Allemagne, le site http://www.digitale-kunstgeschichte.de/wiki/Projekte regroupe plusieurs projets dont certains concernent directement l'accès à des ressources visuelles à l'aide du traitement informatique.

[334] Ces documents recouvrent entre autres les collections issues de la célèbre mission héliographie (1851), auxquelles se sont ajoutées d'autres fonds d'images provenant par exemple du Service photographique des armées (1915), les documents de l'Inventaire général, mais aussi des photographies d'Eugène Atget ou de Félix Nadar. En 2004, après la dissolution de l'Association française de diffusion du patrimoine photographique (AFDPP) qui rassemblait ces collections, la *MAP* récupère ces fonds. En outre, les photographies conservées proviennent principalement de deux autres sources : la bibliothèque et les archives du service des Monuments historiques et le service des Archives photographiques.

Dans la continuité de l'Inventaire général ou d'autres entreprises de photographies documentaires[335], l'objectif de la MAP est de rendre accessibles des reproductions visuelles des œuvres d'art et des monuments historiques français, autrement dit : de valoriser le patrimoine culturel de la France à travers de larges ensembles d'images. Depuis les années 1970, en plus d'organiser des entreprises de photographies documentaires, le ministère de la Culture du gouvernement a mis en place des bases de données informatiques sur le patrimoine (*Joconde*, *Mérimée*, *Palissy*) ainsi qu'un logiciel de recherche documentaire (*Mistral*). À la fin des années 2010, la MAP propose une nouvelle version de ce moteur de recherche sous le nom de *plateforme ouverte du patrimoine* (POP). La POP est l'interface de ce service Web permettant la libre consultation et la réutilisation de l'ensemble des ressources.

Cette « plateforme en ligne » (équivalent de « site Web ») est présentée par la MAP comme un outil au service de la connaissance du patrimoine français ; outil qui est régulièrement enrichi par la mise à jour des métadonnées existantes et par l'ajout de nouvelles images ou informations sur le patrimoine. Grâce à ce service Web, les internautes accèdent à des millions de métadonnées décrivant des œuvres d'art, des monuments, des photographies regroupées dans une même base de données. Plus particulièrement, la base *Mémoire* regroupe les collections photographiques conservées par la MAP, mais aussi d'autres services photographiques du ministère de la Culture. En retour, ces archives photographiques servent d'illustration pour les bases de données *Mérimée* et *Palissy*.

5.1.2. *Memobase*

Ailleurs en Europe, d'autres institutions patrimoniales proposent le même type d'interface que la POP. En collaboration avec divers fonds d'archives, bibliothèques et musées suisses, *Memoriav* a mis en place la plateforme en ligne *Memobase*, qui permet à l'utilisateur de rechercher des photographies, des films, des documents sonores et vidéo indexés dans une même base de données. Memoriav est un réseau d'institutions suisses et de spécialistes du patrimoine qui participent à la préservation et la diffusion du patrimoine audiovisuel suisse.

[335] **Chap. 1, sect. 1.1.**

Dans cette perspective, l'interface Web (Memobase) permet aux utilisateurs d'effectuer des recherches multilingues[336] dans une base de données afin d'accéder à des archives audiovisuelles conservées dans diverses institutions suisses. La mise en place de cette base d'images numériques répond à cinq objectifs principaux[337] : (1) « faciliter la consultation du patrimoine audiovisuel suisse dans les domaines de la photographie, du son, du film et de la vidéo » ; (2) « garantir l'accès à des métadonnées de qualité, cross-média[338] et multilingues » ; (3) « promouvoir la reconnaissance de la valeur immatérielle (culturelle, historique, politique, documentaire) et matérielle des documents audiovisuels qui font partie intégrante du patrimoine historique » ; (4) « offrir un accès aux documents audiovisuels aux secteurs de l'éducation, de l'enseignement et de la recherche » ; (5) « permettre la connexion à d'autres portails d'information en Suisse et à l'étranger ».

Mis en ligne en 2001, Memoriav a continué de développer cette base de données et ce moteur de recherche. Au départ, douze institutions fournissaient à peu près 150 000 documents. En 2020, 67 institutions suisses participent au projet, de telle sorte que les 400 000 documents organisés dans la Memobase constituent la principale base de données sur le patrimoine audiovisuel en Suisse. Encore une fois, la mission est double :

- Valoriser le patrimoine audiovisuel propre à une région déterminée.
- Garantir l'accès gratuit à des métadonnées fiables sur ce patrimoine.

5.1.3. *Belgian Art Links and Tools*

En Belgique, nous pensons à une initiative similaire concernant les archives d'images issues du service photographique des Musées royaux d'Art et d'Histoire. Créé en 1901, ce service est à l'origine d'une vaste photothèque dont les archives sont conservées à Bruxelles, au sein de l'Institut royal du patrimoine artistique (IRPA). L'une des principales missions de l'IRPA est de présenter la « mémoire visuelle de l'ensemble du patrimoine artistique de la Belgique » sur le réseau Internet[339].

---

[336] Étant donné que la langue parlée dans le pays est différente selon les cantons, les ressources de Memobase sont proposées en trois versions linguistiques (français, allemand et italien) pour la plupart des termes connus.

[337] Voir : https://memobase.ch/fr/info-memobase.

[338] Une stratégie « cross-media » a pour but de combiner différents médias (écrit imprimé, sites Web, télévision, cinéma, radio, téléphonie, etc.) afin d'augmenter l'impact d'un discours.

[339] Voir : https://www.kikirpa.be/fr/recherches-en-histoire-de-l-art-et-inventaire.

La photothèque de cette institution rassemble plus d'un million de clichés (de 1880 à nos jours) sur les arts plastiques, l'architecture, l'archéologie, voire sur des métiers d'art ou des paysages. L'inventaire est constamment enrichi de nouvelles photographies produites dans le cadre de la restauration d'œuvres ou dans le cadre d'études en histoire de l'art mis en place par l'IRPA. Ces photographies anciennes et plus récentes sont converties en ressources numériques pour les diffuser sur Internet, même si ces documents sont toujours consultables sur place (salle de lecture de l'IRPA).

En effet, depuis 2013, un portail regroupant les différentes bases de données de l'institution a été développé : *Belgian Art Links and Tools* (BALaT)[340]. L'utilisateur peut ainsi accéder à cette photothèque en ligne qui comporte à peu près 750 000 photographies ; ces ressources sont indexées à l'aide d'un thésaurus spécialisé. Cette photothèque est complétée par les ressources de bibliothèque de l'IRPA (à peu près 80 000 livres et articles spécifiques sur des problématiques qui ont trait au patrimoine, à l'histoire de l'art, à la conservation-restauration et aux analyses scientifiques). En outre, cette plateforme en ligne permet d'effectuer simultanément une recherche dans plusieurs banques de données :

- La photothèque en ligne de l'IRPA.
- La bibliothèque de l'IRPA.
- *Le Dictionnaire des peintres belges du XIVe siècle à nos jours…* (cf. Robert-Jones & De Wilde, 1995).
- Le répertoire des personnes et institutions relatives aux pratiques d'art et d'artisanat.

Comme les deux services précédents, ce service belge permet d'accéder aux images de plusieurs collections ou fonds d'archives relatifs au patrimoine belge par le moyen d'une interface couplée à une base de données[341].

[340] http://balat.kikirpa.be/intro.php. Trois versions linguistiques de *BALaT* sont disponibles néerlandais, français, anglais), mais nous pouvons remarquer que ces trois langues ne correspondent pas complètement aux trois langues nationales de la Belgique (néerlandais, français, allemand).

[341] En outre, ce service propose une rubrique d'aide qui explique en détail la procédure d'utilisation de la base de données, cf. http://balat.kikirpa.be/help.php.

5.1.4. *Getty Search Gateway*

Enfin, un autre service est similaire à celui des stratégies de communication en ligne du patrimoine français, belge ou suisse. Aux États-Unis, le *Getty Search Gateway* (GSG) est un site Web qui permet aux utilisateurs d'effectuer des recherches dans plusieurs ressources documentaires du *Getty Center*[342]. De même que les trois institutions précédentes, les principaux objectifs visent l'accessibilité à un large ensemble de ressources documentaires sur le patrimoine culturel, et l'encadrement de ces ressources par une base de données comprenant surtout des métadonnées descriptives. Le premier but du service est d'aider les spécialistes du visuel à découvrir la variété des ressources disponibles dans les ressources gérées par l'institution. Les utilisateurs peuvent utiliser ce service pour parcourir les collections par type (peintures, livres, archives) ou pour rechercher directement des ressources en fonction de thématiques précises (périodes historiques, artistes, types d'objets, etc.).

Ainsi, comme dans le cas de la POP, de Memobase et de BALaT, le GSG mobilise le numérique pour regrouper des ressources et des informations à partir de plusieurs collections d'images et bases de données :

- *J. Paul Getty Museum Collection Database*, qui rassemble notamment des photographies d'objets issus de l'antiquité grecque, étrusque et romaine, mais aussi des reproductions photographiques de peintures, de dessins, de sculptures et d'arts décoratifs européens (du Moyen Âge à la fin du XIX^{e} siècle) ou des reproductions de manuscrits enluminés (du Moyen Âge et de la Renaissance), ainsi que des photographies documentaires (du XIX^{e} siècle à nos jours) ;
- *Getty Research Institute Research Library Catalog*, qui rassemble diverses ressources documentaires (plus d'un million de livres, de périodiques, de catalogues de vente aux enchères), mais aussi des collections de livres rares et de manuscrits, d'estampes et de dessins, de photographies ;
- *Getty Research Institute Collections Inventories and Finding Aids* (notices pour des ressources rares : livres, manuscrits, gravures, dessins, dispositifs optiques, photographies, vidéos) ;

[342] Le *Getty Center* est une institution patrimoniale américaine fondée en 1997 et située à Los Angeles. Le *Getty* comprend une partie du *J. Paul Museum*, ainsi que les ressources de la *Getty Foundation*, du *Getty Research Institute* et le *Getty Conservation Institute*.

- *Getty Research Institute Digital Collection* (sélection de documents numérisés du Getty Research Institute ; plus de deux millions de documents regroupant des textes et des images).

Comme dans les services précédents, l'idée est de rassembler et de canaliser l'accès à toutes ces ressources par le moyen d'un seul service Web.

Dans les sections suivantes, nous montrerons comment ces quatre institutions patrimoniales envisagent d'aménager la communication en ligne d'un large ensemble d'images. Dans chaque cas, une base d'images numériques est rattachée à une institution patrimoniale possédant ses propres informations relatives aux images. Les fonctionnalités de ces quatre services sont relativement semblables d'un cas à l'autre. Cette similitude ne découle pas seulement du fait que la conception (le *design*) du dispositif de lecture des ressources numériques se base sur des architectures et des standards de programmation que nous avons présentés[343]. Nous postulons en outre que les institutions ont d'abord choisi de déterminer la communication en ligne en vue de certains usages, et surtout d'imposer une scénarisation de l'interface qui se base principalement sur les compétences discursive et sémiolinguistique de l'utilisateur[344].

Nous verrons que la première tendance stratégique a surtout pour but de focaliser l'attention de l'utilisateur sur deux *fonctions* du dispositif de lecture : (1) soit celui-ci fonctionne comme un *espace multimédia* qui présente rapidement l'institution, le service Web et les ressources disponibles, (2) soit le dispositif fonctionne comme un *instrument de recherche* pour explorer ces ressources.

5.2. L'interface Web en tant qu'espace graphique et multimédia

Pour préparer l'utilisation d'une base d'images numériques, les services Web sont généralement présentés par les institutions dans des *pages d'accueil* ou des *rubriques informatives*. Nous commencerons par analyser les pages d'accueil des quatre services Web que nous avons rapidement présenté, puis nous analyserons les fonctionnalités de recherche et les modalités de présentation des résultats du moteur de recherche. Même si

[343] **Chap. 2**, **sect. 2.3.1.**

[344] **Chap. 1, sect. 1.5.**

l'utilisateur peut contourner cette première phase de l'exploration, la page d'accueil doit évidemment être considérée comme le point de départ de la scénarisation conçue par les institutions patrimoniales.

Les pages d'accueil manifestent généralement un *menu*, c'est-à-dire un élément d'interface graphique, souvent rectangulaire, dans lequel sont présentées différentes listes d'accès aux fonctionnalités du service Web. De plus, le moteur de recherche et des illustrations d'images sont généralement affichés dès le départ dans ces pages. Par exemple, la page d'accueil de la POP fait ressortir une ressource à l'arrière-plan (fig. 12), mais aussi un moteur de recherche et des catégories générales qui renvoient à des ressources en fonction de différents types d'images ou de thématiques (fig. 13 ; les différentes catégories sont représentées par des images qui fonctionnent comme des liens hypertextes). Chaque catégorie renvoie vers une sélection d'images disponibles dans la base d'images. Dans la POP, cette partie est présentée comme des « accès thématiques » aux ressources en fonction de plusieurs sous-catégories (photographie, peinture et dessin, mode, voyages, artistes femme).

De même, la page d'accueil du GSG présente un moteur de recherche où l'utilisateur doit encoder des mots-clés, mais la page comprend aussi des liens hypertextes qui renvoient vers des sélections d'images (p. ex. : en fonction de la provenance des images ou d'un type particulier de ressources visuelles ; cf. fig. 16).

Enfin, en plus de présenter directement le moteur de recherche, les pages d'accueil du BALaT (cf. fig. 14) et de Memobase (cf. fig. 15) comprennent des exemples de ressources disponibles dans la base d'images numériques (la page d'accueil de Memobase accorde une place importante aux images, par exemple en utilisant des archives d'images pour constituer l'arrière-plan de la page ; cf. fig. 15). De plus, les menus de ces deux pages renvoient également à des ressources documentaires dont le but est de contextualiser les images ou encore à une rubrique d'aide (cf. fig. 14 & 15) pour faciliter la navigation sur le site Web.



Figure 12 : Page d'accueil de la *plateforme ouverte du patrimoine* ; au milieu, le « champ de recherche », en dessous, des « hyperliens » renvoyant aux différentes bases de données regroupées sur cette plateforme.

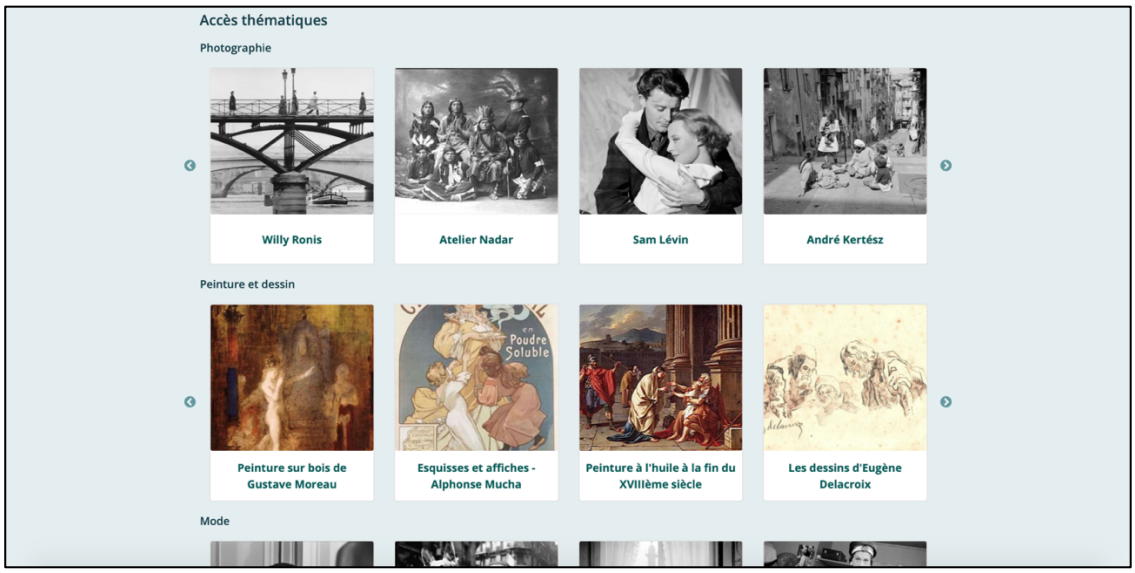


Figure 13 : Suite de la page d'accueil de la *plateforme ouverte du patrimoine* ; différents « accès thématiques » en fonction de catégories générales (photographie, peinture et dessin, mode).

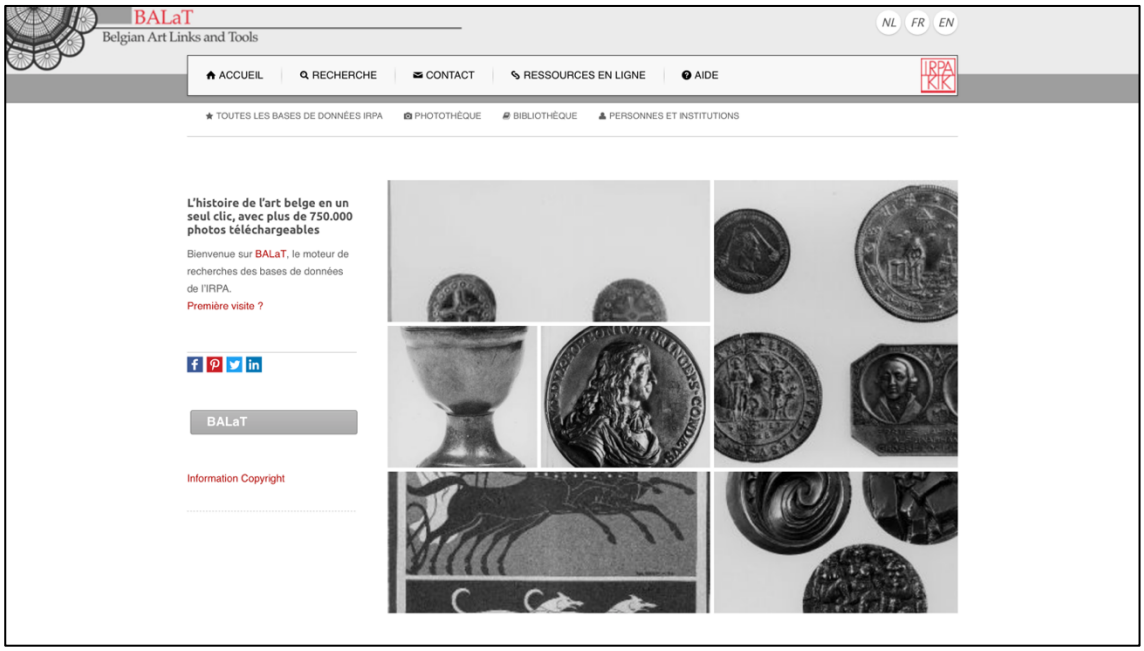


Figure 14 : Page d'accueil de *Belgian Art Links and Tools* ; cette page présente des exemples différents de ressources à chaque actualisation de la page.

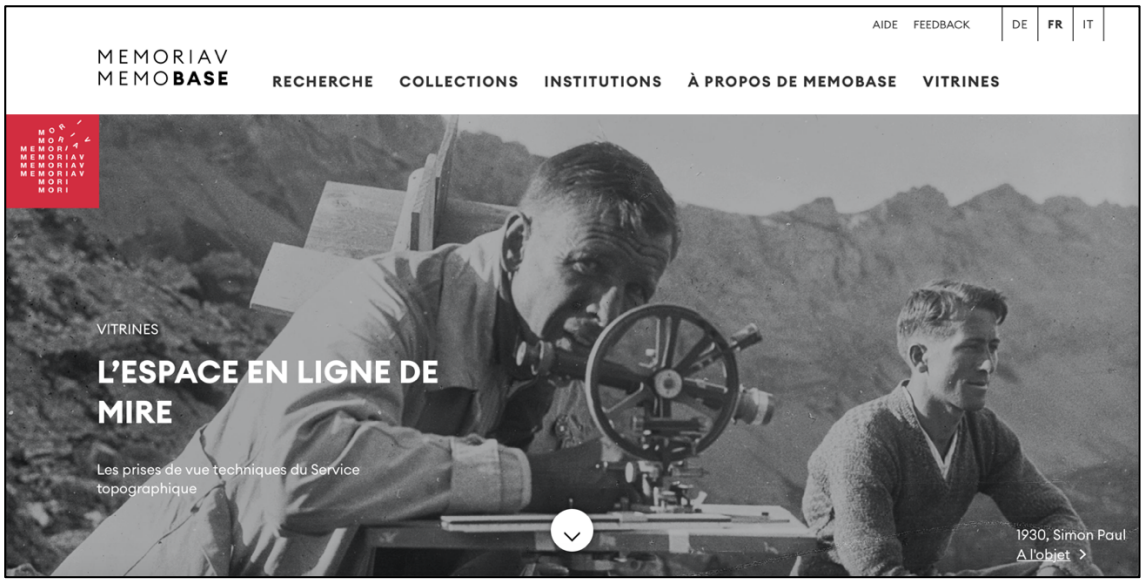


Figure 15 : Page d'accueil de *Memobase.*

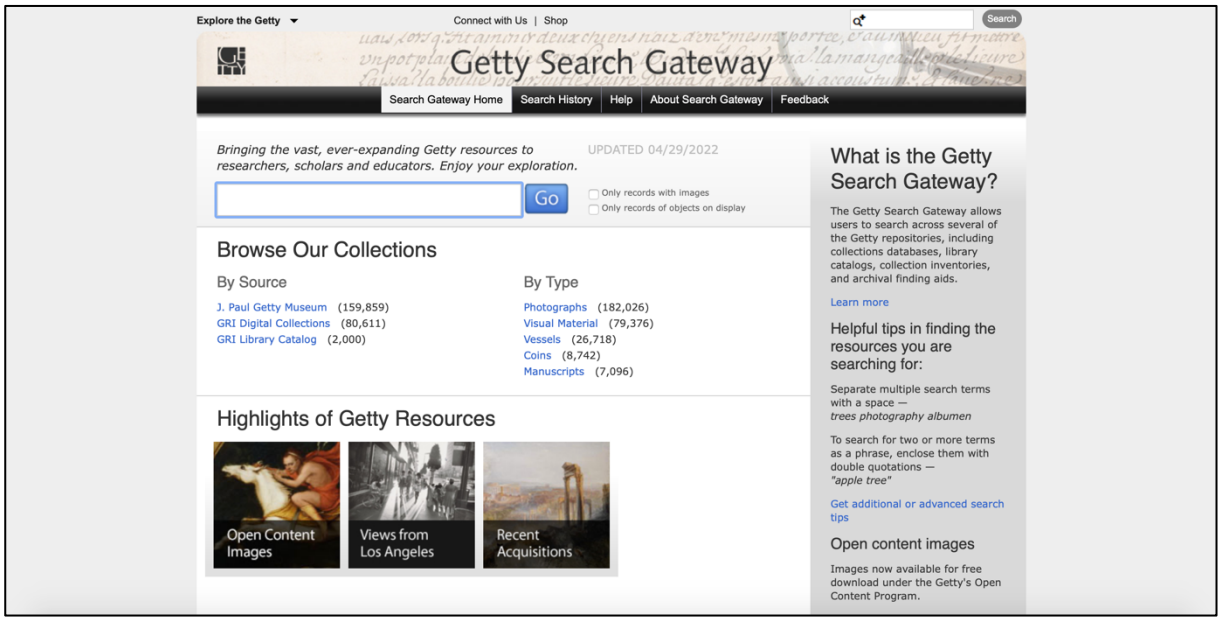


Figure 16 : Page d'accueil du *Getty Search Gateway* ; la page est composée d'un champ de recherche, de divers « accès thématiques » (*by source*, *by type*, *highlights*) et d'informations verbales sur le service Web.

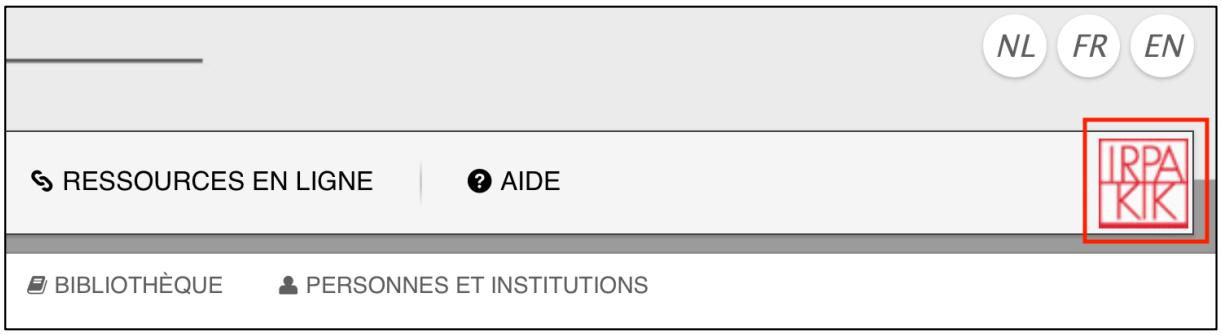


Figure 17 : Détail de la page d'accueil de *Belgian Art Links and Tools* ; le logo de l'institution fonctionne comme un hyperlien qui renvoie vers le site de l'*Institut royal du patrimoine artistique*.

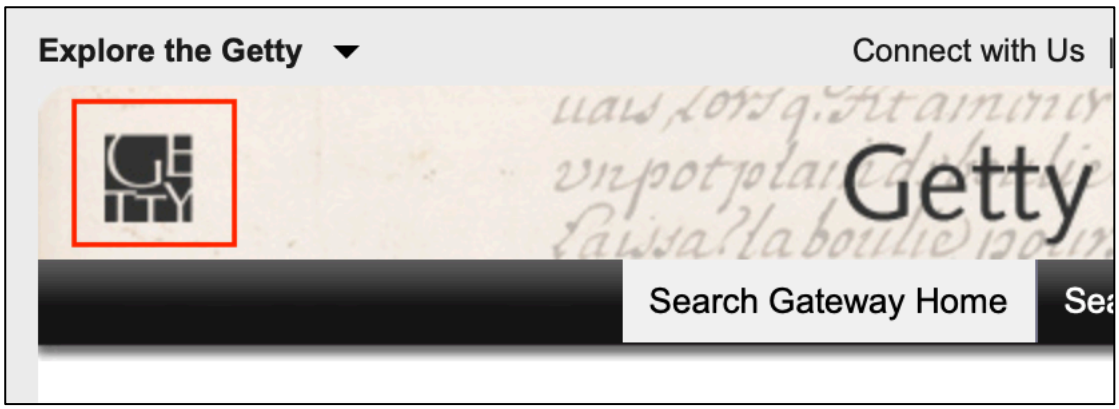


Figure 18 : Détail de la page d'accueil du *Getty Search Gateway* ; le logo fonctionne aussi comme un hyperlien qui renvoie vers le site Web du *Getty*.



Figure 19 : Rubrique « à propos » de *Memobase*.

Ces quelques exemples permettent de constater comment la page-écran est utilisée pour construire d'emblée le « point de vue » de l'utilisateur (cf. aussi **infra**). Chaque page fonctionne comme un plan d'expression dont les limites (le cadre) sont à la fois déterminées par le positionnement de l'utilisateur dans le site Web et par les bords concrets de la page-écran. L'enjeu principal est d'organiser cette page comme un espace graphique qu'il s'agit de lire. Néanmoins, cette lecture est distribuée dans différentes zones de la page. Le parcours du regard de l'utilisateur se base donc sur une mise en page différente que dans les documents écrits où le regard est habituellement invité à lire le texte de gauche à droite et de haut en bas. Cela n'empêche pas que la disposition de toutes les informations (verbales ou visuelles) dans la page-écran correspond à une mise en forme stable et cohérente.

Pourtant, dans ces exemples, la scénarisation de l'information n'aboutit pas sur un *récit*, si nous entendons par « récit » comme une suite d'événements et d'actions (une « histoire ») racontée par un « narrateur ». Cette scénarisation correspond plutôt à une stratégie d'affichage qui vise d'abord à capter l'attention de l'internaute (selon Fontanille, la « manipulation étendue ») afin de l'engager dans différentes scènes d'énonciation, très souvent des textes écrits (la « manipulation restreinte »). La lecture de ces textes ne débouche pas sur une histoire, mais sur un « produit d'information », autrement dit, un *savoir*. La fonction de l'interface est justement d'établir une *relation énonciative* entre l'institution patrimoniale et l'utilisateur du site Web. Cette relation permet, d'une part, de circonscrire l'expérience de l'utilisateur au sein d'une communauté discursive et ainsi de

garantir l'intelligibilité de l'archive[345], et d'autre part, de distribuer deux rôles énonciatifs : l'institution est définie comme le producteur d'un savoir et l'utilisateur est le destinataire de ce savoir.

Plusieurs exemples sont observables à travers des actes d'énonciations (de l'institution à l'utilisateur) formulés sur le mode de l'interrogation. Dans BALaT, l'énoncé « Première visite ? » est un lien hypertexte qui renvoie vers une rubrique informative (fig. 14). Dans le GSG, « What is the Getty Search Gateway ? » est le titre d'une rubrique informatique (fig. 16).

Dans d'autres cas, les actes d'énonciation sont plutôt formulés sur le mode de l'assertion, par exemple dans la POP (« La plateforme POP regroupe les contenus numériques du patrimoine français afin de les rendre accessibles et consultables au plus grand nombre »), dans BALaT (« L'histoire de l'art belge en un seul clic, avec plus de 750 000 photos téléchargeables ») ou encore Memobase (« Le portait du patrimoine audiovisuel suisse »).

Dans les deux cas, l'institution patrimoniale est définie d'emblée comme le détenteur d'un ensemble de connaissances que le dispositif de communication en ligne rend accessible.

D'un autre côté, les institutions projettent des marques de cette instance dans la page-écran à travers le nom ou le logos de l'institution. Dans les pages d'accueil du GSG et de BALaT, la mise en évidence des logos des institutions de l'IRPA ou du Getty indique le producteur qui est à l'origine du service en ligne (cf. fig. 17 & 18). Ces logos sont des hyperliens qui renvoient l'utilisateur aux sites Web de ces deux institutions. À l'aide de ces hyperliens, l'internaute est directement renvoyé aux pages d'accueil des sites Web des institutions patrimoniales Getty et IRPA, où une autre rubrique « à propos » (*about us*) permet d'attribuer une *identité* au producteur du service. Toutefois, cette identité est généralement réduite à la mission et aux principaux objectifs des institutions ; cette mission et ces objectifs concernent avant tout des pratiques d'archivage.

À titre d'exemple, nous pouvons lire que la mission du Getty est de permettre de mieux comprendre la société à travers la promotion des arts visuels :

[345] En effet, nous avons montré que l'archive est liée à la langue, parce que c'est toujours une pratique discursive et rhétorique qui garantit l'intelligibilité des archives ; cf. **chap. 1, sect. 1.3.**

> Getty is a cultural and philanthropic institution dedicated to the presentation, conservation, and interpretation of the world's artistic legacy. Through the collective and individual work of its constituent programs – Getty Conservation Institute, Getty Foundation, J. Paul Getty Museum, and Getty Research Institute – Getty pursues its mission in Los Angeles and throughout the world, serving both the general interested public and a wide range of professional communities in order to promote a vital civil society through an understanding of the visual arts[346].

De même, l'IRPA a pour objectif de conserver le patrimoine artistique belge, parce que ce patrimoine a une influence décisive sur société :

> L'Institut royal du Patrimoine artistique (IRPA) est une institution fédérale sans but lucratif qui prend soin du patrimoine belge : les œuvres d'art, monuments et objets emblématiques de notre passé. Ces magnifiques et irremplaçables réalisations incarnent notre histoire commune et façonnent notre identité au sein de la société. C'est pourquoi, à l'IRPA, nous sommes convaincus que l'art est l'émotion qui nous unit[347].

Cette mission des institutions est donc comprise comme une opportunité permise par le numérique[348] et surtout comme un *devoir* ou une *responsabilité.* Les autres institutions ne proposent pas des objectifs différents, au contraire, elles s'alignent à peu près toutes sur une même mission : la diffusion de ces ressources a pour but de mieux comprendre une culture visuelle locale et cette diffusion vise à fédérer une communauté sociale autour de valeurs communes. Cette mission est relative aux pratiques d'archivage. La légitimité de ces services et des discours véhiculés sur ces sites Web s'appuie entièrement sur le fait que l'institution patrimoine à l'origine du service est légitime et qu'elle a donc appliqué un traitement documentaire sur les archives d'images qui sont diffusées en ligne.

Autre exemple, la page « À propos »[349] de Memobase présente une rubrique informatique (cf. fig. 19). D'emblée, nous observons que cette rubrique mélange une énonciation verbale et une énonciation visuelle. Le discours de présentation est accompagné de reproductions numériques (dans ce cas, des ressources disponibles sur le

[346] https://www.getty.edu/about/#our-mission.

[347] https://www.kikirpa.be/fr/mission-et-vision.

[348] Cf. la discussion sur les humanités numériques dans le **chap. 2, sect. 2.1.**

[349] Cf. https://memobase.ch/fr/info-memobase. Voir aussi cf. : https://search.getty.edu/gateway/about.html.

site Web ; dans d'autres cas, il s'agit de photographies d'activités ou de bâtiments de l'institution, de reproduction d'œuvres d'art ou de monuments, etc.).

Somme toute, l'expérience de l'utilisateur des bases d'images numériques est prévue par une scénarisation de l'information dont l'objectif est l'organisation spatio-temporelle du service Web comme un « produit d'information ». Dans tous les cas, la mise en scène de la page-écran s'articule en grande partie sur une énonciation discursive. En effet, une partie importante de l'expérience interprétative de l'utilisateur est conditionnée par l'omniprésence de textes écrits (donc d'énoncés verbaux) dans la page-écran. Le site est donc compris comme un espace produit par la pratique d'un lieu qui est écrit, même s'il faut prendre en compte les activités de navigation en considérant le site *comme un tout multimédia* (orchestration de l'interface).

À ce propos, nous avons remarqué ce rapprochement entre écriture et numérique[350]. En tant que dispositif de communication, les médias informatisés gèrent et propagent d'anciennes formes écrites. Sauf que l'écrit informatique aboutit sur un texte qui est supporté à la fois par un objet matériel et par les différentes possibilités d'écriture qui sont propres au numérique. En effet, nous avons noté que les médias informatisés sont aussi compris comme des *machines à suggérer* favorisant « la création constante de nouvelles formes écrites qui suggèrent des possibilités d'expression et de pratiques » (Tardy & Jeanneret, 2007 : 214).

Le processus de communication en jeu dans les services Web des institutions patrimoniales est effectivement conçu sur ce principe de suggestion des formes écrites et d'autres formes sémiotiques à travers la page-écran. Pourtant, l'utilisation du support numérique suggère-t-elle, dans ces services, la création de nouvelles formes d'écritures et d'expression concernant les archives d'images ? Est-ce que cette médiation informatique permet d'augmenter les possibilités de compréhension des images, de faire ressortir des connexions et des relations entre elles ? Ou alors la visée de cette médiation se résume-t-elle à transmettre un stock d'images et à l'encadrer dans un scénario déterminé et écrit à l'avance lors de la conception de l'interface ? Dans ce cas, comment l'utilisateur est-il impliqué dans ce processus qui se fonde sur ces « écrits d'écran et de réseau » ?

[350] **Chap. 1, sect. 1.4.**

Il faut remarquer que l'organisation spatiale et temporelle de l'information partage des points en commun avec la « mise en page » des documents écrits. D'une manière générale, une page fonctionne comme une surface (un support) sur laquelle sont inscrites des informations. Dans les pages d'un document écrit, la disposition des signes, des mots et des phrases, voire des illustrations (figures, schéma, tableaux) qui composent le document sont organisés en fonction de certaines règles de mise en page. Celle-ci consiste à organiser formellement les éléments graphiques manifestés dans chaque page afin de faciliter la compréhension du texte. Parmi ces éléments, certains invariants (les « titres », le « corps du texte ») constituent la partie principale du texte écrit qui est manifesté sur la page ; cette partie principale est généralement bordée par des « marges régulières », ce qui contribue (avec la police d'écriture) à homogénéiser l'aspect global du texte en tant qu'ensemble signifiant. Enfin, la manifestation des « figures », au « premier plan », s'oppose à la dissimulation du « fond », c'est-à-dire « l'arrière-plan » de la page (souvent d'une couleur blanche ou plus claire que les caractères d'écriture).

Comme dans la plupart des documents écrits, une page Web est avant tout une surface-support pour inscrire et manifester un texte numérique en fonction d'une mise en page correspondant à certaines normes formelles. En effet, la page Web s'appuie également sur une organisation formelle (disposition régulière des éléments du texte) et sur l'opposition entre des « figures de la lecture » (au premier plan) et le « fond de la représentation » (à l'arrière-plan). Ces « figures » et ce « fond » participent à la construction d'un parcours qui anticipe et constitue les pratiques de lecture du texte numérique en préfigurant une place pour un « lecteur ». Néanmoins, il ne faut pas confondre complètement les fonctions de l'écrit et de l'image dans la page-écran. D'un côté, l'arrière-plan se réduit à une surface-support. De l'autre, les figures « au premier plan » soulignent la dimension verbale de la pratique énonciative qui est en jeu dans ces services Web, tandis que les images occupent seulement une fonction d'illustration dans l'espace graphique et multimédia dont l'utilisateur fait l'expérience.

Ce que nous voulons mettre en évidence, c'est que l'orchestration de ces quatre bases d'images numériques converge vers un seul but : légitimer le service Web de l'institution. Pour ce faire, la relation d'énonciation mise en scène dans et par l'interface favorise un *effet d'objectivité* au récit produit par l'institution. L'approche sémiotique que

nous avons présentée[351] (plus précisément la théorie de l'énonciation) peut aider à comprendre cet effet d'objectivité. En s'appropriant l'écriture comme principal mode d'expression, les interfaces Web mettent en avant une énonciation discursive qui s'appuie sur la *forme de l'énoncé énoncé* à travers des opérations de débrayage actantiel (projection d'un « non-je »), temporel (« non-maintenant ») et spatial (projection d'un « non-ici »).

Ces services de communication en ligne aboutissent sur une « communication impersonnelle » ; en effet, l'institution et les utilisateurs sont implicitement considérés à un niveau fortement généralisé et à travers des représentations d'actants qui sont déterminés par le récit de l'institution. Ce type de communication en ligne se fonde en outre sur des rôles énonciatifs qui ne sont pas interchangeables. Comme nous l'avons vu, certaines parties du service Web ont pour fonction d'identifier l'institution et les fonctionnalités du service qui s'adresse à l'utilisateur. D'un côté, l'institution est réduite au nom de l'institution, autrement dit, à un « je » ou à un « nous », et de l'autre les internautes sont regroupés à travers un « tu » ou un « vous ». Ainsi, le « producteur » de l'acte d'énonciation semble complètement dilué dans l'identité de l'institution patrimoniale. De même, l'identité du « destinataire » est tout aussi vague et indéfinie ; il ne s'agit pas de nommer un public cible, mais de rattacher l'utilisateur à la (très large) communauté des internautes.

La scénarisation de l'information dans les bases d'images numériques des institutions s'appuie systématiquement sur ce type d'énonciation discursive, cela même si la page-écran entremêle différents objets médiatiques dans les pages Web (structure modulaire des nouveaux médias). Du document écrit à la page-écran, c'est en effet la structure modulaire qui change certaines conventions d'expression, car le support numérique s'approprie et combine les formes sémiotiques propres à diverses pratiques médiatiques. La page articule des textes écrits, des icônes, des hyperliens et des images au sein d'un même *espace*.

Cet espace permet de « rendre matériel », ou plutôt de « rendre présent » le discours de l'institution par le moyen d'une énonciation numérique. La visée rhétorique de cet espace s'appuie sur deux caractéristiques essentielles des interfaces. D'une part, la page-écran encadre la scène d'énonciation ; chaque page d'un service Web est composée

[351] **Chap. 1, sect. 1.3.**

de différents éléments (graphiques et visuels) constituant le plan d'expression de la communication en ligne (l'utilisateur parcoure et interagit uniquement avec ce plan). D'autre part, l'interface participe à un dispositif de représentation (fonction transitive) qui impose certaines conditions de visibilité et de manipulation, notamment en s'appuyant sur une mise en scène (spatiale) et une orchestration (combinaison des différentes parties de la mise en scène) de la page-écran.

Dans les paragraphes précédents, nous avons analysé quelques aspects de l'*espace graphique et multimédia* dont l'utilisateur fait concrètement l'expérience. Comme nous l'avons déjà montré[352], un site Web prend généralement la forme d'un espace à regarder et à lire et avec lequel il faut interagir. Tout se passe comme si l'orchestration de l'interface avait pour but d'orienter et de guider la navigation de l'utilisateur à travers divers « écrits d'écran », et donc de l'engager dans des « parcours de lecture » qui concourent à la scénarisation discursive des activités de l'institution, de sa mission et de ses objectifs. La communication en ligne des archives d'images numériques impose par conséquent des contraintes sémio-pragmatiques qui déterminent une certaine distribution des rôles énonciatifs et un cadre spatio-temporel (le dispositif de lecture).

En schématisant un peu cette distribution, nous pouvons dire que l'institution patrimoniale tient le rôle de *producteur* du service et, par conséquent de toutes les scènes d'énonciation qui préparent les parcours de lecture dans un produit d'information, tandis que l'utilisateur tient le rôle de *destinataire* de ce produit. Dans cette perspective, l'interface Web est comprise comme le support intermédiaire entre l'utilisateur et une *scène énonciative globale* (manipulation étendue) qui s'ouvre sur une succession discontinue de *scènes d'énonciation locales* (manipulation restreinte). Cette stratégie favorise l'alternance entre deux niveaux de lecture de l'interface, même si ces niveaux participent à un *même acte global d'énonciation*. Cependant, cet « acte d'énonciation » n'est jamais « totalement accompli », parce que la « lecture du texte » de l'interface s'appuie sur une exploration *dynamique* (mouvement d'une page Web à une autre ; mouvement au sein d'une page Web) et *transitoire* (du fait que l'énonciation n'est achevée que lorsque l'utilisateur décide d'arrêter son exploration).

[352] **Chap. 2**, **sect. 2.2.2.**

5.3. Fonctionnalités des moteurs de recherche et présentation des résultats

La première phase a pour but de mettre en place un *espace* qui favorise surtout une relation d'énonciation verbale entre l'institution et l'utilisateur. L'exploration des bases d'images numériques se base sur une seconde phase : la recherche et la présentation des ressources numériques. Rappelons-le, l'une des principales caractéristiques du numérique est que ce support permet de manipuler des contenus grâce à des opérations computationnelles réalisées par un système informatique[353]. Comme nous l'avons remarqué, les bases d'images numériques comprennent une partie cachée, puisqu'un SGBD traite automatiquement les requêtes de l'utilisateur en fonction d'un « input linguistique » ; c'est pourquoi ce type de service met généralement en évidence une « barre de recherche » dans la page d'accueil pour inciter l'utilisateur à encoder des mots-clés.

Les bases d'images numériques qui sont conçues par les institutions patrimoniales se fondent sur le même schéma de fonctionnement (cf. fig. 20).

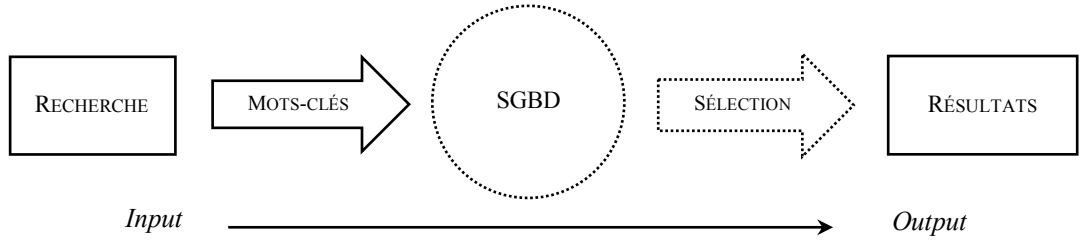


Figure 20 : Schéma typique du fonctionnement du base d'images numériques (en traits pointillés, les opérations dissimulées, et en traits pleins, les opérations visibles sur la page-écran) ; en suivant ce schéma, l'exploration d'un large ensemble d'images est conditionné par l'utilisation de mots-clés (langage verbal).

Dans cette perspective, l'espace de la page-écran est la seule partie visible du dispositif de lecture des ressources numériques. (Cet espace sert de support intermédiaire entre, d'un côté, les actions et la réflexivité de l'utilisateur, et de l'autre, les fonctionnalités du moteur de recherche ; le moteur de recherche se base uniquement sur des propositions de mots-clés et les résultats de cette recherche sont ensuite présentés dans l'interface par la page-écran.) Tandis que le système informatique est la partie invisible (le « support

[353] **Chap. 1, sect. 1.4.**

imperceptible ») qui sélectionne les ressources pertinentes uniquement en fonction des informations encodées préalablement dans un SGBD[354].

C'est dans cette perspective que les institutions patrimoniales visent à mettre en place un instrument de recherche dans une base d'images numériques. Pourtant, ce choix indique déjà une limite évidente : même si les possibilités d'exploration reposent autant sur des compétences humaines et des actes (les développeurs qui programment l'interface et l'utilisateur qui fournit l'*input*) que sur des opérations automatisées (sélections d'images par l'ordinateur, l'*output*), les possibilités d'usage de ces services Web sont quant à elles réduites, parce que les moteurs de recherche fonctionnent surtout comme un « filtre » qui donne uniquement à voir des images sélectionnées par l'ordinateur en fonction de mots-clés[355].

Par conséquent, le bon fonctionnement du service dépend entièrement de la qualité de la description et de la structuration des images, car nous avons vu que les archives d'images numériques sont généralement organisées par rapport à des informations descriptives sur les supports originaux. Il est donc attendu que l'utilisateur ait une idée assez précise de l'input linguistique à encoder[356].

Les instruments de recherche proposés par la POP, le GSG, BALaT ou Memobase sont assez semblables pour que nous les analysions ensemble. L'utilisateur peut souvent choisir entre une « recherche par champs » (conseillée pour les spécialistes) ou une « recherche générale » (plus intuitive et plus simple pour les non-spécialistes). Les termes qui sont encodés doivent être courts et précis ; pour simplifier le processus de recherche, des listes de mots-clés sont parfois disponibles.

---

[354] **Chap. 2, sect. 2.3.2.**

[355] Voir les discussions dans le **chap. 3**.

[356] Par exemple, deux utilisateurs cherchant à établir un corpus d'images : le premier sur des bâtiments détruits à cause de la guerre, le second sur des images d'architecture des ponts. Le premier encoderait les mots-clés « bâtiment », « guerre », « destruction », tandis que le second encoderait « architecture » et « pont ». Si les résultats ne sont pas satisfaisants, il encoderait progressivement des termes plus spécifiques. Or le problème est que les résultats risquent d'être extrêmement variés (non seulement le type de supports d'images, mais aussi les époques de production ou les fonctions originales des images). Pour réduire les résultats, les interfaces permettent parfois de spécifier la recherche en combinant plusieurs mots-clés répartis dans différents champs (cf. **infra**, fig. 25). En outre, elles comprennent souvent des « filtres de sélection ».

Lorsque ces services présentent les résultats d'une recherche, la mise en scène spatiale de l'interface fait ressortir *trois zones principales* :

- Une première zone (A) regroupe les filtres applicables aux résultats.
- Une deuxième zone (B) présente la barre du champ de recherche (éventuellement un bandeau avec le logo et le nom de l'interface, le nom de l'institution soutenant le projet, des menus déroulants, etc.).
- Une troisième zone (C) est l'endroit qui présente les résultats de la recherche.

Les deux premières zones (A et B) sont généralement les plus stables, c'est-à-dire que les composants graphiques se répètent indépendamment des résultats de la recherche. Ces deux zones se situent principalement dans les marges de la page, de telle sorte qu'elles forment un encadrement permanent lors de la présentation des résultats de la recherche. Le *design* de l'interface peut changer d'un site Web à un autre, mais en principe les interfaces pour explorer des bases d'images numériques sont similaires : les résultats sont présentés dans une zone, cet encadrement permanent dans une autre (cf. fig. 21-24 & 26).

Ces résultats de la recherche sont présentés sous la forme d'une liste (zone C) comprenant deux éléments principaux :

- Une miniature de l'image.
- Des informations descriptives relatives à l'image.

Il est possible d'appliquer des filtres (zone A) sur les résultats de la recherche par mots-clés, afin d'affiner et de diminuer le nombre d'images proposées en fonction de certains paramètres (créateur ; type d'objet ; matériau ; technique ; style ; etc. ; cf. fig. 21-24 & 26). Dans ces *listes d'images miniatures*, chaque entrée de la liste comporte donc, d'un côté, un substitut numérique du support d'origine, de l'autre des informations sur ce support : le type d'objet, l'auteur, la description du contenu, le numéro de référence dans la base de données et le lieu de conservation du support original. Selon les mots-clés utilisés pour la recherche, ces listes d'images peuvent comprendre un nombre très important de ressources à parcourir. Si les filtres sont mis en évidence, c'est justement pour répondre à ce problème. D'une part, ces filtres linguistiques sont des indices des modalités du traitement documentaire des images, d'autre part, ils mettent en évidence l'automatisme de la manipulation numérique.

Le choix de la *liste* est sans doute la forme de présentation la plus courante pour manifester des sélections d'images numériques par le moyen d'une interface Web ; nous retrouvons en effet cette forme dans quasiment tous les services Web mis en place par d'autres projets institutionnels similaires[357]. Comme nous l'avons souligné, le support numérique s'approprie certaines structures conceptuelles propres à l'écriture[358]. La liste est l'une de ces structures.

Sur ce sujet, des travaux ont montré que la liste est un discours délinéarisé où les unités discursives sont ordonnées dans une logique d'énumération (Goody, 1979 & 1985 ; Bachimont, 2010). Le but d'une liste est en effet de rassembler un ensemble d'informations en classant et en catégorisant ces informations autrement que dans un discours continu. D'aucuns affirment même que le développement de l'écriture a favorisé l'utilisation des listes, ce qui a « induit un rapport au monde qui procède de la raison classificatoire : penser le monde, c'est l'organiser en classes et hiérarchies, c'est l'ordonner et le ranger » (cf. Bachimont, 2010 : 165). C'est sans doute cette « raison classificatoire » qui est à la base des services Web conçus par les institutions, qui cherchent surtout à maintenir une continuité entre les pratiques d'archivage et les pratiques d'archivage numérique. Il s'agit de reproduire à distance l'expérience de consultation d'un fonds d'archives d'images, par le moyen d'une base d'images numériques qui fonctionne comme un inventaire des ressources accessibles.

À ce propos, les remarques de Marie Despret-Lonnet (2013 : 76-83) sont particulièrement éclairantes. Chaque institution est dépositaire d'un patrimoine culturel et doit procéder, durant le traitement documentaire, à des choix qui réduisent la mémoire afin de renforcer sa visibilité ; selon l'auteure, ces choix renforcent « la cohérence suggérée ou recherchée par le rassemblement préalable d'un ensemble d'objets en un même lieu au sein d'une même institution » (*ibid.* : 76). En effet, « cataloguer », « inventorier », « répertorier » ne désignent pas uniquement des opérations techniques ; ces activités du traitement documentaire sont des « manières d'englober chaque objet dans un *discours commun*, d'en faire un membre d'une collection et de lui donner un sens comme entité de cette collection » (*ibid.* ; nous soulignons).

[357] Cf. **supra, note 333**.

[358] **Chap. 1, sect. 1.4 & supra, sect. 5.2.**

De même, les bases de données ont pour but de rassembler et de diffuser (dans un même lieu et sous un même format) des « éléments textuels » qui ont été choisis pour « décrire des objets inventoriés ». L'accumulation des objets et données est déterminée par l'organisation documentaire qui rend ces données intelligibles. La construction d'un discours sur des ressources concerne plus « que la somme des parties » ; la totalité de ces ressources « doit être pensée à une autre échelle, pas uniquement technique, mais aussi politique, culturelle et symbolique » (Després-Lonnet, 2013 : 83). L'auteure envisage les résultats du traitement informatique des informations sur des archives comme « support de médiation », voire comme un « texte porteur de sens » ; sauf que, comme nous l'avons vu, la conversion et les copies numériques obligent de laisser de côté la matérialité originale de ces archives[359].

> Les campagnes de numérisation ont permis de mettre en ligne les copies numériques de millions de livres en texte intégral, de représentations d'œuvres d'art, de documents d'archives, dont les originaux n'ont plus à être nécessairement consultés hors du système documentaire. Les index servent effectivement à revenir à l'objet qu'ils désignent, mais ils permettent également de constituer des sous-ensembles d'objets porteurs des mêmes caractéristiques. Les « lots résultats » ne sont pas seulement des potentialités mais également une approche d'une des multiples dimensions de la collection. De ce fait, *l'affichage des résultats d'une recherche ne devrait plus être considéré comme un simple sous-produit intermédiaire, mais en tant que document, c'est-à-dire en tant que « texte porteur de sens et doté d'une intentionnalité »*[360] *et donc comme support de médiation à part entière*. Chaque recherche permet en effet de mettre en regard des dizaines de représentations d'objets qui possèdent des caractéristiques communes. Ces configurations textuelles, qui donnent à voir une partie des informations relatives aux objets sélectionnés, suggèrent dans le même temps l'existence des liens entre eux. Ce sont autant de documents de synthèse auxquels le lecteur donnera un sens, guidé en cela par les choix d'organisation des concepteurs (Després-Lonnet, 2013 : 78-79 ; nous soulignons).

Néanmoins, l'auteure précise dans la foulée qu'il ne faut pas réduire ces « configurations textuelles » ou ces « documents de synthèse » à de simples « répertoires d'objets » :

[359] **Chap. 2., sect. 2.2.**

[360] Voir aussi Pédauque (2003).

> Chaque catalogue est un ensemble documentaire situé qui possède une forte cohérence structurelle intimement liée aux missions de l'organisme qui l'a produit. Les choix singuliers effectués dans chaque contexte sont des prises de position sur le monde et des propositions de mise en ordre et de mise en sens des ensembles ainsi constitués. Chaque organisation documentaire est en elle-même une forme d'énonciation qui organise et anticipe les conditions d'appropriation des informations qu'elle rassemble. Les catalogues peuvent aussi être envisagés comme autant de représentations synoptiques du lieu dont ces objets proviennent, la notion de lieu n'étant pas prise ici au sens de lieu géographique, mais en tant que construction sociale. Elle désigne un ensemble d'acteurs, d'activités, de compétences professionnelles réunies autour d'un projet commun et appliquant les normes et les rites propres à ce lieu (Després-Lonnet, 2013 : 79).

Ainsi, la seconde phase, celle de l'exploration, est caractérisée par le fait que les images sélectionnées par l'ordinateur sont systématiquement présentées par le moyen de cette logique de classification qui est propre au régime graphique. Le caractère structurant de la liste fait ressortir l'un des principaux objectifs de l'archivage : l'organisation systématique d'une archive comme un ensemble cohérent de ressources documentaires qui ont été indexées. Cela révèle aussi une certaine tendance dans les archives et la documentation qui, à travers les catalogues et les inventaires, est la poursuite d'une quête démesurée, celle de la « totalité par l'inventaire » ; cette quête produit parfois des listes « vertigineuses » (cf. Eco, 2009).

Somme toute, la conception des bases d'images numériques s'approprie et manifeste la logique documentaire qui est au fondement de l'organisation des archives d'images. D'une interface à une autre, cette logique est clairement observable à travers une véritable constante : les modalités de présentation des ressources. En effet, les ressources de BALaT (fig. 27), de la POP (fig. 28) ou de Memobase[361] (fig. 29) sont systématiquement présentées par une reproduction numérique du support qui est accompagnée d'informations structurées (les métadonnées descriptives).

[361] Dans ce cas, la présentation des informations est un peu différente, mais on y retrouve toujours le même type d'informations (description ; contexte ; matériaux ; etc.). Memobase offre en outre deux fonctionnalités de recherche : la recherche en texte libre permet de proposer un terme ou plusieurs termes simples ; la recherche par opérateurs booléens, qui permet d'intégrer les signes « + » et « – » dans le champ de recherche pour relier ou exclure des mots-clés, afin d'augmenter la précision de la recherche. Dans les deux types de recherches, des filtres prédéfinis permettent également de réduire les résultats.

En plus des listes d'images, nous pouvons remarquer une autre mise en forme des résultats par le moyen d'un outil de cartographie[362] en ligne. Cet outil est une fonctionnalité proposée par la POP en combinant l'interface Web à un service de cartographie en ligne (une carte de la France) à la place des listes. Ce type de fonctionnalité correspond à des méthodes dont le but est justement de proposer un ensemble d'instruments pour schématiser de manière graphique des données statistiques[363]. En proposant un point de vue global sur une carte de France, cette schématisation a alors pour fonction de synthétiser et de mettre en évidence des relations quantitatives entre des données ou des ressources numériques (cf. fig. 30 ; Memobase propose également cette fonctionnalité cartographique). Nous observons en effet que les résultats sont manifestés à travers une représentation graphique indiquant la répartition des résultats de la recherche sur une carte ; cette représentation est accompagnée d'annotations (chiffres) indiquant le nombre de données rassemblées en chaque point. Les ressources sont regroupées dans les ronds bleus. L'utilisateur peut éventuellement parcourir la carte (agrandir, déplacement), et lorsqu'il sélectionne une région plus précise de la carte, les points se divisent et créent un nouveau réseau de points correspondant à l'endroit où sont conservés les supports originaux. Lorsque l'utilisateur change l'échelle, une nouvelle répartition est présentée, et ainsi de suite.

Dans les quatre exemples analysés, nous constatons que l'espace graphique et toutes les fonctionnalités manifestés par la page-écran convergent vers un même « effet » : la représentation des archives d'images (absentes) à travers un dispositif de lecture. Le but de ce dispositif est de pallier l'absence de l'original par le moyen d'un dispositif de représentation qui a pour seule fonction de manifester, à travers la page-écran, des sélections de ressources sous la forme de configurations textuelles (les « listes » ou, selon Després-Lonnet, des « lots de résultats » compris comme des « documents de synthèse »). Rappelons-le, Louis Marin a mis en évidence deux dimensions complémentaires de la représentation : une *dimension transitive* et une *dimension réflexive*. La première concerne la transparence du plan de la représentation et la scène (visuelle) exprimée par le moyen de ce plan. La seconde concerne la *présentation de la représentation* (l'opacité de la représentation) qui est prise, selon l'auteur, « dans la

[362] Cf. **chap. 4, sect. 4.2.1. (note 301)**.

[363] Voir aussi **chap. 8**.

structure dialogique » d'un énonciateur et d'un destinataire, tandis que le *cadre de la représentation* (ses limites, mais aussi sa structure matérielle, donc son support) fournit « un des lieux privilégiés du "faire savoir", du "faire croire", du "faire sentir", des instructions et des injonctions que le pouvoir de représentation, et en représentation, adresse au spectateur-lecteur » (Marin, 1994 : 348). Dans le même sens, l'interface est à la fois cette « présentation » et ce « cadre », autrement dit, elle est à la fois le *lieu* et le *support* d'un acte d'énonciation. Cette énonciation implique nécessairement une manière d'exprimer et de dialoguer, en fonction d'un « point de vue regardant » (l'utilisateur) et d'une « surface regardée » (la page-écran). À chaque moment de la navigation, une partie de cette surface est révélée dans le cadre de l'écran, tandis que le reste de la page est « hors-champ ».

Dans les exemples que nous avons présentés, la « face visible » de la communication en ligne construit le point de vue de l'utilisateur à travers la dimension transitive (la « présentation de la représentation ») de ce *dispositif de représentation scénarisé et programmé*. Dans tous les cas, la construction de ce point de vue est restreinte à celui d'un *utilisateur-explorateur*, voire à celui d'un *utilisateur-lecteur* qui interagit principalement avec des figures de la lecture relevant d'une compétence communicationnelle axée sur l'énonciation verbale. En effet, lorsqu'une image est sélectionnée par l'utilisateur, le dispositif d'énonciation encadre l'intelligibilité de cette image dans une « économie » qui s'appuie systématiquement sur un « métalangage descriptif »[364]. Cette description verbale est plus qu'une simple légende, puisqu'elle reprend la totalité des métadonnées sur la copie numérique qui est disponible. Ces informations sont le résultat du traitement documentaire qui a décrit les images en amont de cette classification. Dans chaque page, ces figures de la lecture sont des marques qui indiquent la manière dont les pratiques d'archivage organisent des connaissances, et non pas des marques relatives à l'instance de production (l'institution).

[364] Parmi les exemples que nous avons présentés, deux projets institutionnels issus de pays plurilingues indiquent l'un des inconvénients de cet encadrement des archives d'images à l'aide du langage verbal. En effet, dans le cas de ressources plurilinguistiques, il faut envisager le problème de la traduction à la fois des mots-clés et des descriptions relatifs aux images numériques (cf. les discussions dans le **chap. 3, sect. 3.3)**. C'est notamment le cas de BALaT (avec des versions en français, néerlandais et en anglais, mais pas en allemand), de Memobase (avec des versions en allemand, français et italien), tandis que la POP et le GSG proposent uniquement des versions monolingues (français et anglais).



Figure 21 : Présentation des résultats d'une recherche sur la *plateforme ouverte du patrimoine* ; la zone A comprend le « champ de recherche », la zone B regroupe les « filtres » pour « affiner » la recherche, et la zone C présente les résultats à travers une « liste ».

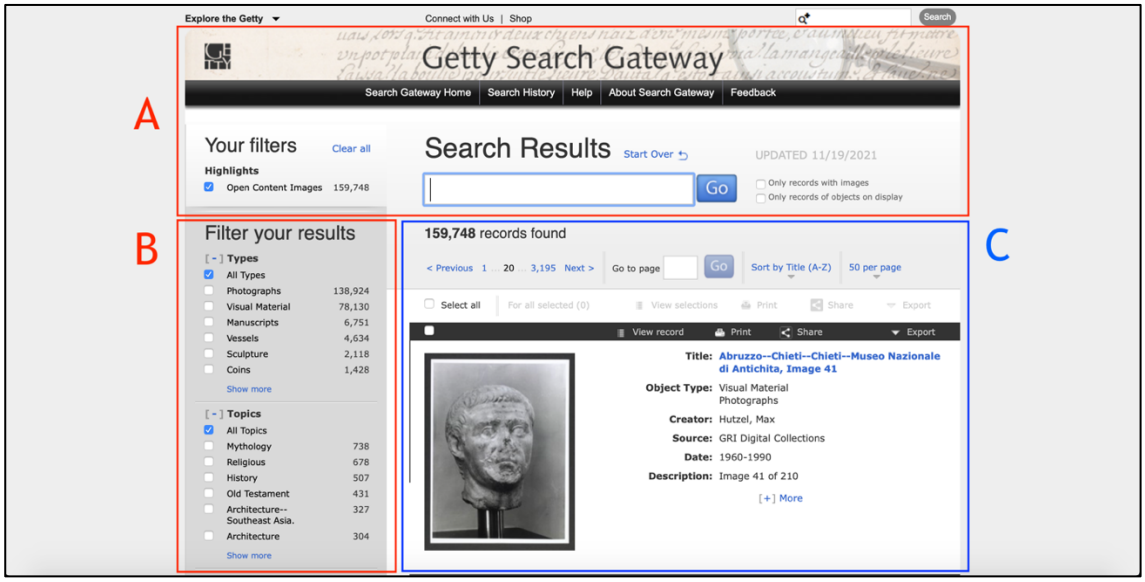


Figure 22 : Présentation des résultats d'une recherche sur le *Getty Search Gateway* ; la zone A comprend le champ de recherche, la zone B regroupe les filtres, et la zone C présente les résultats à travers une liste.

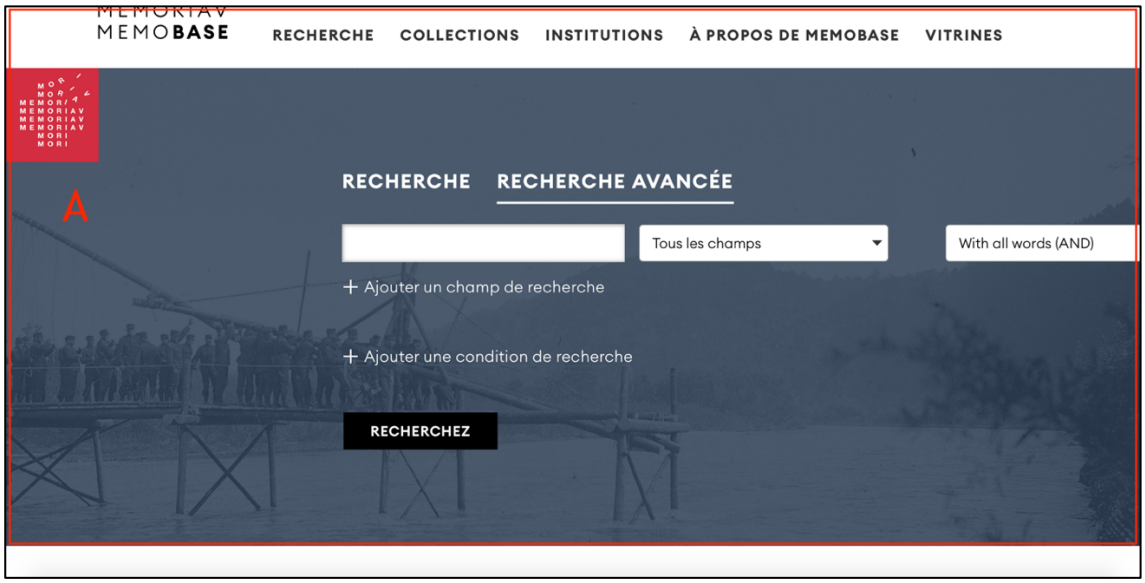


Figure 23 : Champ de recherche avancée de *Memobase* (zone A).

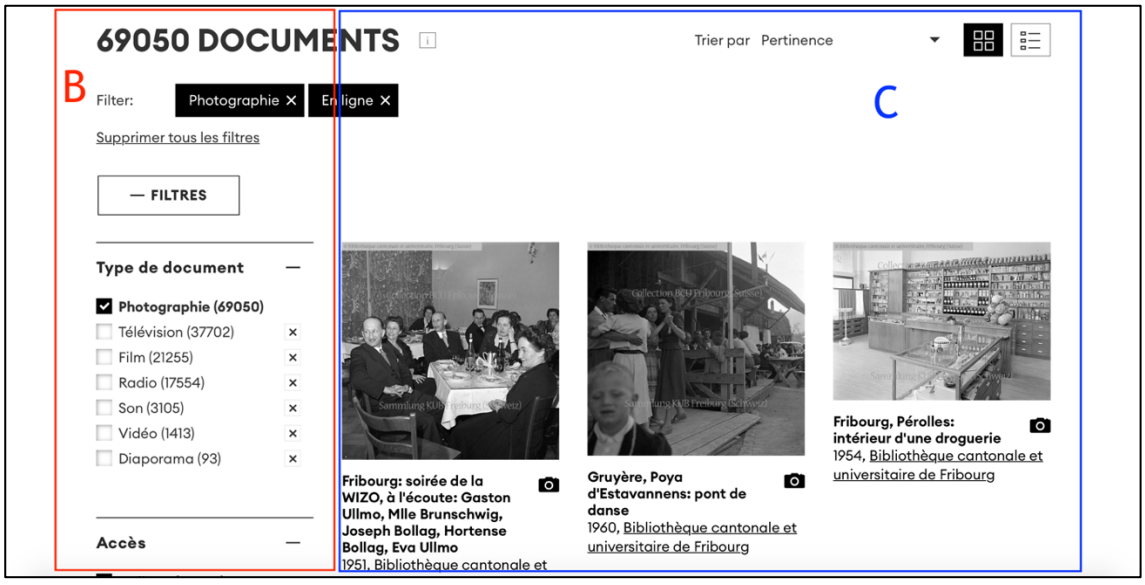


Figure 24 : Filtres (zone B) et liste (zone C) des résultats d'une recherche sur *Memobase*.

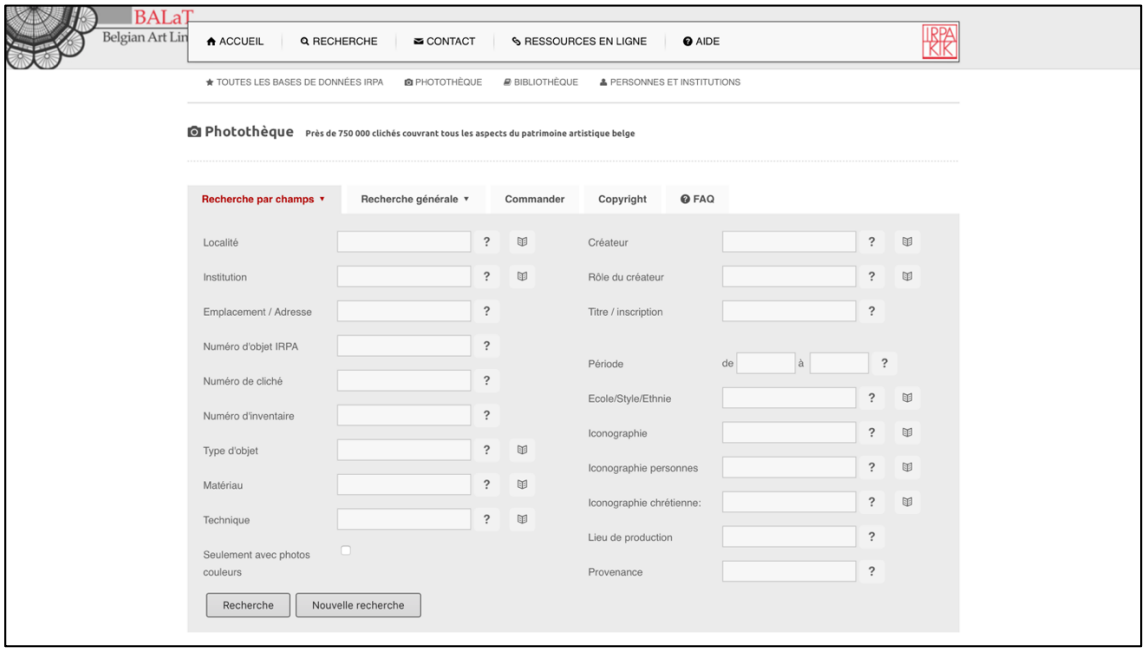


Figure 25 : Champ de recherche avancée dans *Belgian Art Links and Tools*.

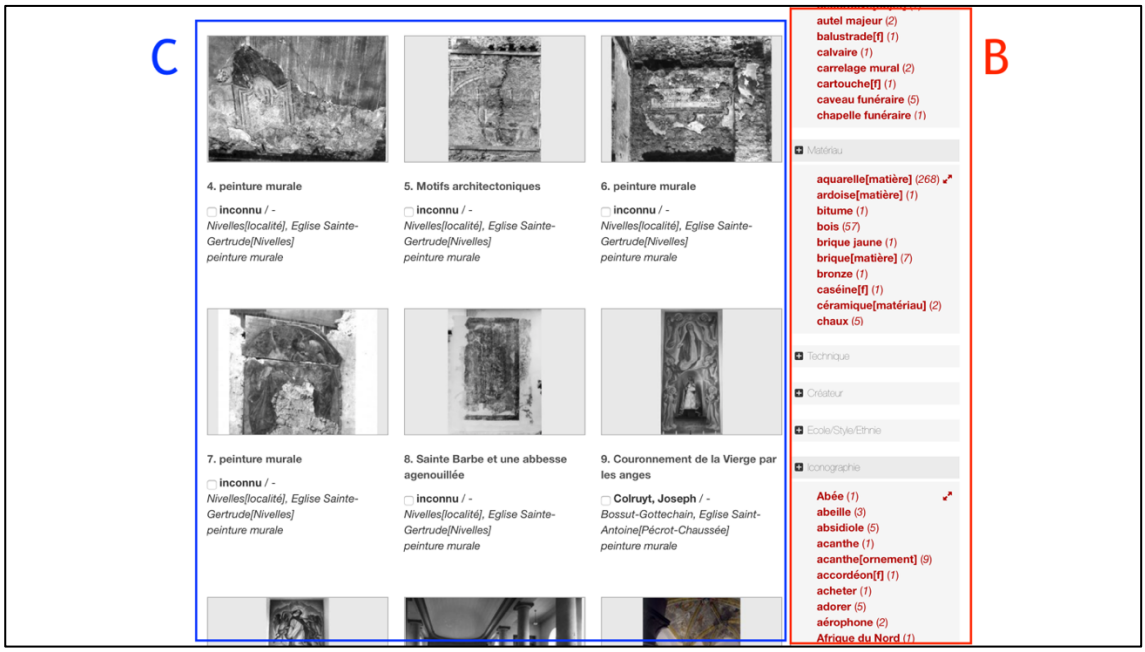


Figure 26 : Filtres et résultats d'une recherche sur *Belgian Art Links and Tools*.

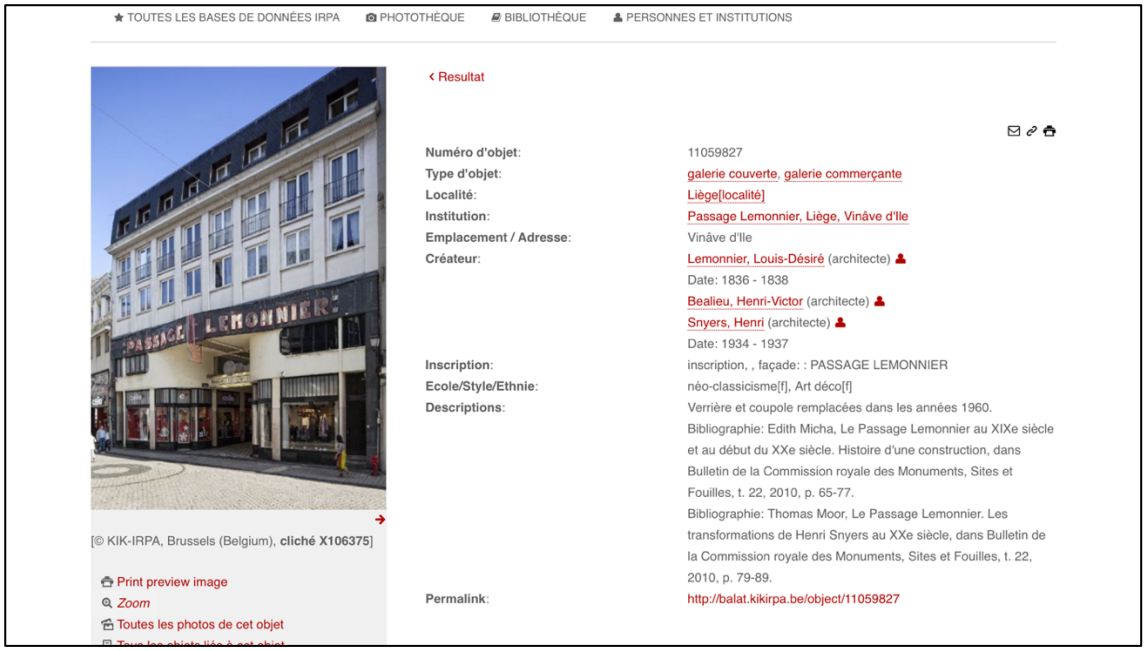


Figure 27 : Présentation d'une ressource numérique sur *Belgian Art Links and Tools.*



Figure 28 : Présentation d'une ressource numérique sur la *plateforme ouverte du patrimoine*.

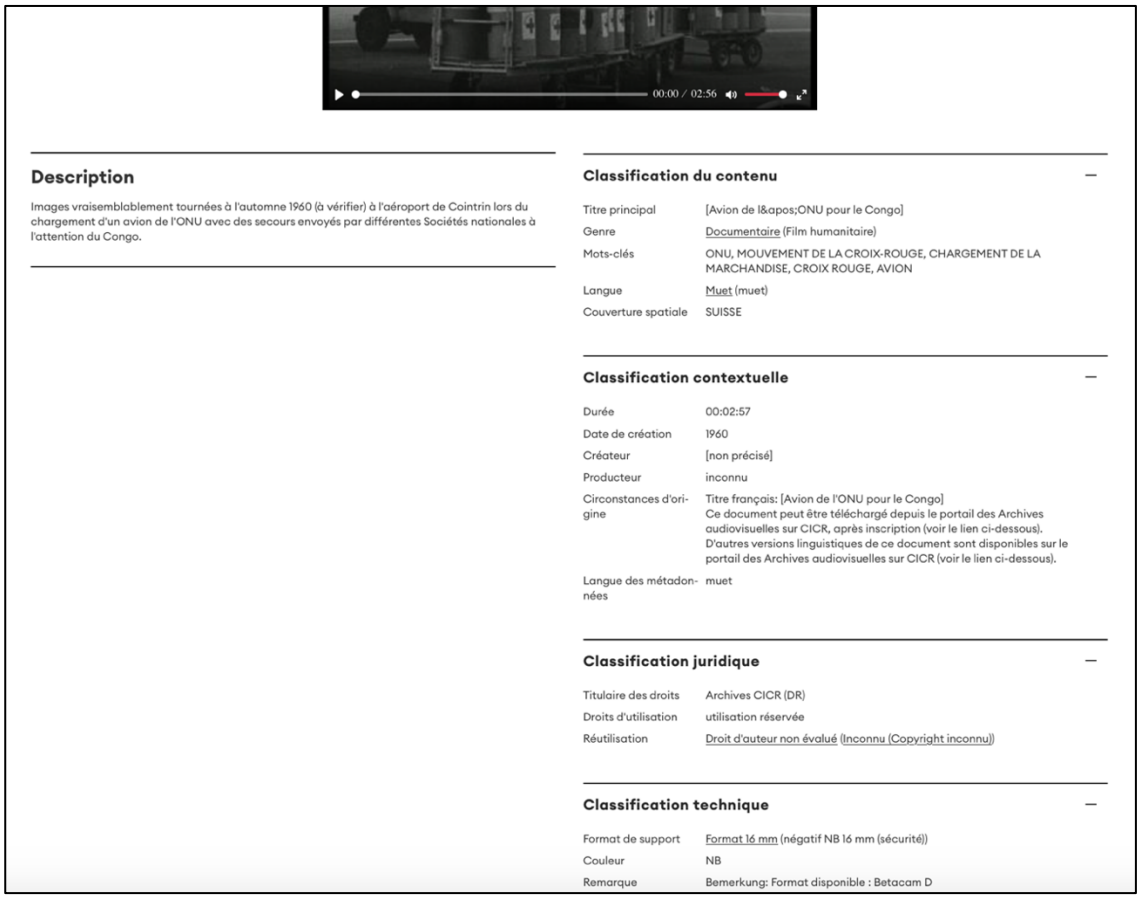


Figure 29 : Détail de la présentation d'une ressource sur *Memobase*.

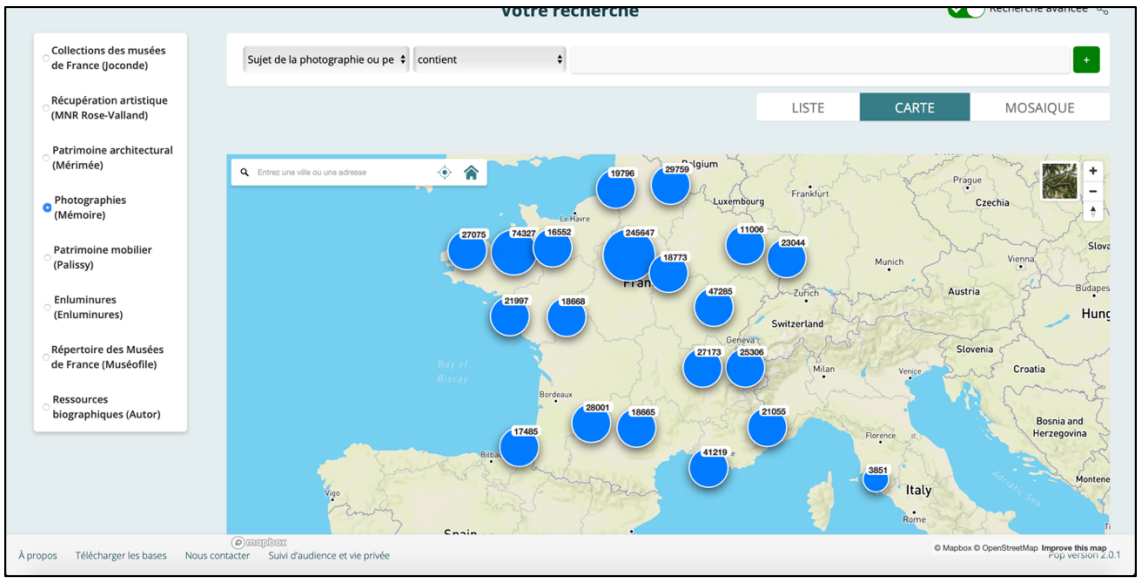


Figure 30 : Outil de cartographie en ligne de la *plateforme ouverte du patrimoine* ; par rapport aux ressources numériques disponibles sur cette plateforme, cet outil indique la quantité de ressources documentaires et les lieux où ces documents originaux sont conservés (lorsque l'utilisateur agrandit la carte, la disposition des ressources (points bleus) est redistribuée).

Somme toute, il est évident que ces services répondent en priorité à un même double objectif :

- Regrouper plusieurs archives d'images qui ont été numérisées pour constituer une base d'images numériques.
- Publier cette base d'images sur Internet par le moyen d'un site Web directement lié à une institution patrimoniale.

Dans tous les cas, l'utilisation des bases d'images numériques s'appuie sur un dispositif de lecture qui aboutit sur des conditions restreintes de visibilité et de manipulation des ressources. La visée de ces quatre services Web est de fournir un espace multimédia pour accéder à un stock d'images qui est indexé par rapport à une base de données. Cet espace participe à une scénarisation qui permet à l'utilisateur d'explorer un ensemble de ressources qui sont reliées à l'aide d'un SGBD. Cette exploration est éventuellement guidée par des discours verbaux concernant les missions et activités des institutions, tandis que la manifestation des résultats d'une recherche dans ces ressources s'appuie généralement sur la forme de listes de substituts. Par conséquent, ces services permettent surtout à l'internaute de prendre connaissance des substituts numériques des archives d'images et des informations descriptives sur ces images.

Pour reprendre d'autres termes que nous avons déjà évoqués[365] : d'un côté, le service Web peut être compris comme un dispositif de *médiation* entre l'usager et le « stock » d'images (pôle référentiel), et de l'autre, à l'aide de ce même service, la *médiatisation* a pour but de présenter l'institution patrimoniale et ses liens avec les activités qui entourent l'organisation des ressources (pôle médiatique). Pour un temps, nous pouvons remarquer que ce type de service a instauré des conditions de visibilité et de manipulation des larges ensembles d'images qui s'appuient principalement sur une *exploration* (1) *dynamique* (déplacements de page à page, ou dans une page), (2) *transitoire* (l'énonciation est interrompue lorsque l'utilisateur arrête son exploration, mais elle n'est jamais accomplie) et (3) *itérative*, puisqu'il s'agit de répéter une série

[365] **Chap. 2**, **sect. 2.2.1.**

d'actions jusqu'à l'obtention du résultat souhaité[366]. Cependant, ces conditions enferment la raison computationnelle dans une « boîte noire » (*black box*)[367].

Ce qui caractérise la première tendance stratégique, c'est que la communication en ligne est conçue comme une relation d'énonciation entre, d'une part, *le savoir que l'institution patrimoniale produit à partir des archives*, et d'autre part, *la volonté de savoir de l'utilisateur qui est guidé par la mise en scène de l'interface*. Les stratégies déployées par les institutions sont relativement standardisées, puisque nous retrouvons le même type de scénarisation dans les différents services Web que nous avons présentés. Dans tous les cas, l'interface est conçue comme un espace principalement composé d'écrits informatiques. Dans cet espace, l'expérience de l'utilisateur est guidée par des opérations de lecture (*utilisateur-lecteur*). Cette première tendance aboutit sur la construction d'un point de vue « rigide » dont le but est de persuader l'*utilisateur-explorateur* que les ressources et les services proposés sont pertinents, et donc, à légitimer la rhétorique de l'institution patrimoniale.

5.4. Conclusions

L'impact de l'informatique sur la classification d'images est indéniable. Dans ce **cinquième chapitre**, nous avons présenté une première tendance stratégique, celle qui consiste à créer un service Web pour sélectionner automatiquement des images en fonction de mots-clés. Ces services s'appuient sur un modèle de classification des images qui dépend d'une indexation préalable des images. Cette indexation repose sur un métalangage de description des archives d'images. Quatre exemples de bases d'images numériques nous ont permis d'observer que ce modèle d'indexation influence les fonctionnalités des dispositifs de lecture qui sont proposés par les institutions patrimoniales. L'expérience de ces dispositifs est guidée par une rhétorique qui

[366] Concernant cette « exploration dynamique, transitoire et itérative », nous pouvons remarquer que le numérique rend possible une fonctionnalité essentielle concernant l'exploration de bases d'images numériques (nous retrouverons également cette fonctionnalité dans les tendances stratégiques suivantes) : les interfaces Web comprennent de nombreux liens hypertextes qui facilitent la navigation de l'utilisateur en reliant des ressources à travers des « renvois » vers d'autres pages Web, voire vers d'autres sites Web, qui engagent l'utilisateur dans d'autres espaces de lecture.

[367] Une « boîte noire » est un système électronique dont le fonctionnement interne est inaccessible ou omis délibérément (seules les interactions avec le système sont observables). Au sens figuré, une « boîte noire » désigne aussi un dispositif (réel ou théorique) dont on ignore ou veut ignorer la structure et le fonctionnement.

instrumentalise le support numérique pour mettre en place des inventaires de leurs ressources ou encore des vitrines des activités de l'institution.

Dans les exemples précédents, nous avons observé que l'interface Web est utilisée pour réaliser deux fonctions : (1) présenter des espaces d'accueil qui expliquent le service ou qui identifient le producteur et le destinataire du service ; (2) pour présenter les résultats d'une recherche. Les opérations informatiques qui génèrent la page-écran et qui sélectionnent automatiquement les ressources numériques s'effacent derrière ces deux mises en scène visuelle de l'interface.

Si l'enjeu de la première tendance stratégique est d'adapter les pratiques d'archivage au support numérique, elle utilise seulement une partie réduite des possibilités qu'offre ce support. La communication en ligne est en effet conçue comme un accès *direct* et *objectif* aux ressources, et surtout comme un dispositif de communication où la distribution des rôles énonciatifs est complètement déterminée par l'institution patrimoniale (sa mission, ses objectifs, ses scénarios). Les choix stratégiques des institutions patrimoniales aboutissent sur des services qui comprennent principalement deux pôles : d'une part, un service d'aide à la recherche dans des archives d'images indexées à partir de mots-clés (pôle référentiel), et d'autre part, une « vitrine » qui vise à valoriser ce service sur Internet (pôle médiatique). Dans tous les cas, l'exploration (les parcours) et les actions (les manipulations) de l'utilisateur sont conditionnées par des *contraintes discursives et linguistiques*.

Dans une certaine mesure, ces choix stratégiques découlent des pratiques d'archivage, et surtout des outils de gestion des archives d'images qui précédent l'utilisation de l'informatique. Avant la conversion numérique des archives, les inventaires étaient habituellement déjà utilisés dans les pratiques d'archivage afin d'établir la liste des documents possédés par l'institution et de les retrouver plus facilement à travers un instrument de recherche décrivant les ressources dans un langage contrôle[368]. Les ressources sur les images sont encodées, structurées et sélectionnées en fonction d'informations verbales. Par conséquent, le traitement informatique n'est pas utilisé pour classer et pour sélectionner des images, mais pour classer des métadonnées

[368] Dans la continuité des outils d'indexation non informatique, ces inventaires font d'ailleurs l'objet d'une normalisation à l'échelle internationale (cf. notamment les normes de description archivistique proposées par le Conseil international des archives (https://www.ica.org/fr), normes qui ont été adaptées au langage informatique).

descriptives et pour présenter des images en lien avec ces descriptions. Cette constatation est importante, car elle met en évidence la principale limite de la classification d'images en fonction de mots-clés : ce type de classification permet seulement d'explorer les ressources par le biais du langage verbal. Le fonctionnement de ces bases d'images numériques est donc limité parce qu'il est seulement fondé sur une description verbale du contenu des images grâce à des ontologies qui rassemblent l'ensemble des notions et concepts utilisés pour indexer les documents.

Dans les **chapitres suivants**, nous mettrons en évidence une deuxième tendance stratégique qui révèle sans doute une volonté de renouvellement des services Web en lien avec des archives d'images, en particulier à travers les principes de coopération, de participation et d'interactivité. Nous observerons par là même une évolution en ce qui concerne des services Web donnant accès à de larges ensembles d'images. Ce changement se rapporte, d'une part, à un renouvellement de la portée des ressources numériques concernées, mais aussi le type d'images qui sont accessibles, et d'autre part, à l'augmentation des applications Web dans les interfaces, afin de mettre en scène différents points de vue sur des œuvres d'art visuel ou sur des archives d'images.

CHAPITRE 6
Renouvellement des bases d'images numériques (1) : participation et coopération

Le **chapitre précédent** a présenté une première tendance stratégique. Dans un premier temps, les institutions patrimoniales ont utilisé la communication en ligne pour mettre en place des « stocks d'images numériques ». D'une manière générale, nous avons distingué deux fonctions principales dans les services Web de quatre institutions patrimoniales. La première fonction de l'interface est de proposer un *espace d'accueil graphique et multimédia* qui favorise surtout une énonciation verbale. La deuxième fonction concerne l'*exploration de la base d'images* à travers un moteur de recherche qui présente les résultats sous la forme de listes d'images.

À ce stade, nous avons seulement présenté des projets qui sont axés sur des archives photographiques relatives au patrimoine culturel. Nous avons vu que les services mis en place par les institutions patrimoniales se limitent souvent à des archives d'images locales (ces projets concernent le patrimoine d'un seul pays ou d'une seule région géographique). Même si les archives sont souvent définies comme des connaissances en lien avec un passé révolu, il ne faudrait pas que cette définition de l'archive nous amène à mettre de côté d'autres pratiques de communication en ligne d'images.

En plus des services mis en place par des institutions patrimoniales qui gèrent surtout des images artistiques ou des archives d'images sur le patrimoine culturel, nous

verrons dans ce **sixième chapitre** que deux autres entreprises constituent également d'immenses bases d'images numériques. Ces entreprises ne se différencient pas de la première tendance stratégique par leurs visées et leurs fonctions, mais elles modifient au moins trois caractéristiques :

- Le type et la provenance des images accessibles dans ces bases d'images numériques.
- L'étendue des pays et des institutions qui fournissent des images et des métadonnées.
- Ces projets font ressortir l'importance du « faire ensemble » par le moyen du numérique.

Nous verrons que l'incitation à la participation des internautes et que la volonté de coopération entre différentes institutions caractérisent cette deuxième tendance stratégique[369]. Pourtant, même si cette deuxième tendance cherche à renouveler la relation entre les ressources et les internautes, elle est encore confrontée aux mêmes limites que la première tendance stratégique.

6.1. Médiathèques en ligne sur le modèle de l'encyclopédie

Avant de considérer le cas d'une stratégie de coopération internationale entre des institutions patrimoniales[370], nous remarquerons d'abord une autre forme de coopération. À côté des services liés à des institutions patrimoniales, nous pouvons en effet mettre en évidence une initiative particulière qui s'inspire des encyclopédies.

Wikimedia Foundation (WF) est une organisation à but non lucratif (créée en 2003) qui s'engage pour la liberté de la connaissance sur Internet. Parmi les projets issus de cette organisation, nous pouvons entre autres citer :

- *Wikipédia*, une encyclopédie libre.
- *Wiktionnaire*, un dictionnaire pour plus de 200 langues.
- *Wikibooks*, une bibliothèque de livres numériques.

[369] En outre, nous remarquerons d'autres caractéristiques : le principe d'un libre partage des ressources, la place accordée à d'autres formats et types d'images, les fonctionnalités interactives des interfaces Web.

[370] Cf. **infra, sect. 6.2.**

Dans les sections suivantes, nous analyserons deux autres projets qui sont des exemples d'encyclopédies d'images numériques : *WikimediaCommons* (WC) et *WikiArt*.

La gestion informatique des projets de WF repose sur une infrastructure informatique complexe qui est dirigée à l'aide de *Wikidata*, une base de données structurée qui permet, d'une part, l'organisation des contenus numériques en un ensemble cohérent et, d'autre part, de créer un réseau informatique pour mettre en place des relations entre les informations, les interfaces et les utilisateurs. Le but de cette base de données est donc de mettre en place un système d'hébergement pérenne (maintenance et conception des serveurs, des bases de données, des logiciels) pour les différents projets éducatifs et collaboratifs rattachés à la WF.

Depuis 2011, la (1) *stabilisation de cette infrastructure informatique* est l'une des cinq priorités stratégiques de WF, à côté de (2) l'*augmentation de la participation des utilisateurs*, de (3) l'*amélioration de la qualité des ressources*, (4) l'*étendue de la portée internationale des projets*, et (5) l'*innovation technologique*[371]. En 2017, WF a d'ailleurs proposé un nouveau plan stratégique qui s'articule principalement autour des notions de *service* et d'*équité*. L'accès aux connaissances est en effet défini comme un « service destiné aux utilisateurs », et WF s'engage à faire évoluer l'infrastructure technique pour une connaissance et des outils critiques.

D'autre part, cette connaissance prend en compte un maximum de « points de vue » pour construire des communautés solides et diverses, mais en favorisant l'échange et surtout le « libre partage ». En résumé :

> D'ici à 2030, Wikimedia deviendra la principale infrastructure de l'écosystème de la connaissance libre, et quiconque partageant notre vision pourra se joindre à nous. Nous, les communautés et les organisations Wikimedia, ferons avancer notre monde en rassemblant un savoir qui représente pleinement la diversité humaine, et en construisant les services et les structures qui permettront à d'autres de faire de même. Nous poursuivrons notre mission de création de contenu comme nous l'avons fait par le passé et le ferons davantage à l'avenir[372].

[371] Voir notamment *Wikimedia Strategic Plan : A collaborative vision for the movement through 2015* ; cf. https://upload.wikimedia.org/wikipedia/commons/c/c0/WMF_StrategicPlan2011_spreads.pdf?uselang=fr.

[372] Voir : https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Direction.

Cette vision ambitieuse influence l'ensemble des projets qui se basent tous sur ce principe de libre partage des connaissances, notamment en défendant une stratégie qui se fonde sur l'*encyclopédisme.* Rappelons-le, l'*encyclopédie*[373] concerne un ensemble organisé de connaissances, mais elle se différencie clairement de l'*inventaire* (liste exhaustive) ou encore du *dictionnaire* dont l'objectif concerne plus précisément l'explication du sens et de l'usage des mots d'une langue. Sur ce sujet, Alain Rey précise : « Alors que le "dictionnaire" est enfermé dans les structures d'une langue et de son lexique, l'encyclopédie est ouverte sur le monde. Qu'elle soit rédigée dans une langue ou une autre importe peu : on peut toujours la traduire »[374].

Ainsi, une encyclopédie est un ouvrage de référence dont le but est de synthétiser et d'organiser un ensemble de connaissances générales. Dès que des connaissances sont jugées dignes « de mémoire et de transmission », elles sont susceptibles d'être rassemblées dans une encyclopédie. L'encyclopédie est en outre caractérisée par un *discours* qui doit être le plus neutre possible, c'est-à-dire un discours qui efface les marques de l'énonciateur dans le texte. Ainsi, le « discours encyclopédique » est proche du « discours de vulgarisation »[375].

De plus, l'encyclopédie vise à organiser des connaissances. Au départ, le mode d'organisation le plus courant est l'organisation thématique par rapport à différentes disciplines scientifiques. Le classement alphabétique a été généralisé par la suite. Le « classement alphabétique » est plus adéquat aux habitudes de lecture qui se développent aux environs du XVIII^e^ siècle (voir par exemple l'encyclopédie de Diderot, 1985). Du reste, l'encyclopédie s'est transformée durant son histoire et les approches d'organisation des ouvrages encyclopédiques ont suscité des débats (cf. notamment Pinault, 1993).

À l'origine, ce genre d'ouvrage est habituellement publié sous la forme d'un livre ou d'un ensemble de livres. Plus récemment, des encyclopédies sont également

---

[373] La production de ce type d'ouvrage est attestée depuis le XVIII^e^ siècle : *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers* (Denis Diderot & Jean d'Alembert) de 1751 à 1772 ; *Encyclopédie méthodique* (Charles-Joseph Panckoucke) de 1782 à 1832 ; *Encyclopédie nouvelle* (Pierre Leroux & Jean Reynaud) de 1834 à 1841 ; *Grand Dictionnaire universel du XIX^e^ siècle* (Pierre Larousse) de 1866 à 1888 (il s'agit d'un « dictionnaire encyclopédique ») ; *La Grande Encyclopédie* (Henri Lamirault) de 1886 à 1902 ; *Encyclopédie française*, (Anatale de Monzie & Lucien Febvre) de 1932 aux années 1960 ; *Grand Larousse encyclopédique* de 1960 à 1975 et *La Grande Encyclopédie* de 1970 à 1985.

[374] Voir Alain Rey, « Encyclopédie », dans *Encylopædia Universalis* [en ligne : https://www.universalis.fr/encyclopedie/encyclopedie/].

[375] Cf. **infra**, notamment la discussion sur la place de la vulgarisation dans le projet *Europeana*.

disponibles sur Internet[376]. Dans cette perspective, *Wikipédia* est sans aucun doute le service le plus connu de WF et l'encyclopédie en ligne la plus consultée sur Internet. Sur ce sujet, Sémir Badir et Nicolas Couégnas considèrent l'encyclopédie en ligne comme un objet étrange du fait de « la nature collaborative de la production de ses entrées » et « par le statut et la fiabilité incertaine des savoirs collectés » (Badir & Couégnas, 2020 : 243). Cependant, les auteurs remarquent qu'une analyse sémiotique peut mettre en évidence les éléments qui permettent de confirmer que *Wikipédia* est une encyclopédie ou de particulariser *Wikipédia* par rapport à d'autres encyclopédies (*ibid.* : 244) ; le but de cette analyse est de montrer « ce que les traits spécifiques ou distinctifs de *Wikipédia* disent des discours et de genres avec lesquels, ordinairement, une encyclopédie entre en rapport » (*ibid.*).

Pour ce faire, Badir et Couégnas proposent de considérer *Wikipédia* comme un « objet sémiotique » relevant du niveau de pertinence des pratiques culturelles ; ces pratiques lient une *composante discursive* et une *composante médiatique*. Selon les auteurs, *Wikipédia* et les encyclopédies ont en commun qu'ils se fondent sur une « forme d'organisation du savoir » et sur une « manière de présenter ce savoir ».

À ce propos, rappelons que l'analyse énonciative des images a mis en évidence qu'un énoncé visuel est souvent un « lieu de savoir » (l'énoncé est donc en relation avec un « positionnement » d'où des choses sont vues et à partir duquel l'observateur construit sa propre vision). Chaque énoncé visuel est porteur d'un questionnement sur le « savoir partagé » qui circule et qui est parfois disputé entre l'énonciateur (caractérisé par une activité de production) et l'énonciataire (caractérisé par une volonté de connaissance). C'est dans ce sens que nous avons vu que les institutions patrimoniales aménagent la communication en ligne comme un *dispositif, graphique et multimédia,* qui propose une manière de présenter un savoir (en tant qu'énonciateur, ces institutions sont caractérisées comme les détenteurs d'un savoir, tandis que les internautes, caractérisés par une volonté de savoir, en sont les destinataires)[377]. Ce dispositif détermine le positionnement du spectateur en préparant la « place » que l'observateur doit occuper pour reconstruire le parcours de la signification.

[376] Par exemple, l'encyclopédie payante *Universalis*, cf. https://www.universalis.fr.

[377] Voir **chap. 5**.

Lorsqu'ils considèrent la complexité discursive de *Wikipédia* par rapport à la question du *genre*, Badir et Couégnas estiment que ce service est à la fois un *forum encyclopédique* et un *réseau social encyclopédique*, mais un forum et un réseau « où le caractère encyclopédique tient davantage du thème que du genre » (*ibid.* : 254). En effet, certaines caractéristiques des discours véhiculés sur *Wikipédia* ne se confondent pas avec celles que nous retrouvons, par exemple, dans l'encyclopédie en ligne *Universalis* (ou encore dans la version papier). Dans le second cas, nous pouvons notamment remarquer que (1) la rédaction et la longueur des articles est plus homogène, et que (2) les auteurs (souvent des spécialistes du sujet traité dans l'article) sont mentionnés.

Ce deuxième point est important, car il souligne la relation (implicite) entre les connaissances et le discours scientifiques (ou académiques). Au contraire, *Wikipédia* refuse explicitement cette affiliation à la communauté scientifique, même si de nombreux articles comprennent des références bibliographiques qui renvoient à des ouvrages scientifiques[378]. Même si la composante discursive de ce service peut être liée à la pensée scientifique, l'objectif est avant tout de rendre accessibles ces connaissances auprès d'un large public. Badir et Couégnas rappellent en effet que le discours scientifique n'a pas sa place dans *Wikipédia*. Néanmoins, ce service se base sur une procédure rigide de production des articles, ainsi que sur une administration complexe et un formatage simplifié des ressources numériques (cf. *ibid.* : 256-257).

Ainsi, la composante discursive de *Wikipédia* vise à transmettre des opinions et à éduquer la communauté d'utilisateurs du service. L'objectif est de *vulgariser* des connaissances en proposant des ressources qui sont compréhensibles par un public plus large que les seuls experts. En définissant ses propres « conventions de style », ce service Web collaboratif pose probablement les jalons d'un autre genre discursif, que l'expression « forum encyclopédique » désigne assez justement[379].

D'un autre côté, plusieurs traits distinctifs permettent de différencier *Wikipédia* des autres encyclopédies (cf. *ibid.* : 252-253 & 256) :

---

[378] À ce propos, la documentation disponible sur le site Web définit le « style encyclopédique » en fonction de six points essentiels (le style de rédaction doit être *neutre*, *impersonnel*, *clair*, *précis*, *compréhensible*, *non académique*). Voir notamment : https://fr.wikipedia.org/wiki/Wikipédia:Style_encyclopédique & https://fr.wikipedia.org/wiki/Wikipédia:Conventions_de_style.

[379] En effet, sur Internet, le « forum » est un service de discussion sur un thème donné et sous forme d'articles (chaque utilisateur peut lire les interventions de tous les autres et apporter sa propre contribution).

- L'hétérogénéité d'échelles des sujets (certains articles sont très longs, tandis que d'autres sont très courts).
- Un fonctionnement à travers un réseau de renvois hypertextes (ce réseau rend possibles à la fois la catégorisation et la représentation hétérogène du savoir ; cf. **infra**).
- L'homogénéité et le nombre limité des formats sémiotiques (inscriptions textuelles stables, parfois accompagnées de quelques illustrations).
- L'information réfutable (aucune information présentée sur *Wikipédia* n'est objective ni subjective, mais l'information peut être révisée ou éditée à l'infini).
- L'accessibilité généralisée (le trait spécifique qui est le plus mis en évidence par *Wikipédia*).
- Le principe de neutralité dans la rédaction des articles.
- L'absence d'un consensus entre les différentes versions linguistiques[380].
- L'authentification des contenus à l'aide de sources.

Nous verrons que certains de ces principes (réseau hypertexte ; homogénéité des formes sémiotiques ; accessibilité généralisée ; neutralité) sont également observables dans les bases d'images de WF que nous analyserons (cf. **infra**).

Par rapport au genre de l'encyclopédie « classique », notons que, dans les encyclopédies accessibles en ligne, la question du « classement alphabétique » tend à devenir obsolète étant donné que l'utilisateur du site Web explore l'encyclopédie par le moyen d'*hyperliens* (ou liens hypertextes). Comme nous l'avons déjà remarqué[381], un hyperlien est l'une des fonctionnalités centrales des sites Web. Cette fonction informatique permet de passer automatiquement d'un document numérique à une autre partie de ce document ou à un autre document numérique. D'une manière générale, ce type de liens est généralement mis en évidence dans la page (en gras ou souligné, ou encore à l'aide d'une couleur différente du corps du texte). Un hyperlien peut être situé

[380] Notons que *Wikipédia* se décline en à peu près 300 langues et dialectes, dans le but de permettre au plus grand nombre de communautés d'utiliser les ressources de cette encyclopédie en ligne ; toutes langues confondues, *Wikipédia* rassemble actuellement plus au moins 174 millions de pages, chacune correspondant à un article encyclopédique ; voir notamment les statistiques publiées par *WF*, cf. https://stats.wikimedia.org/#/all-wikipedia-projects/content/pages-to-date/normal|line|all|~total|monthly.

[381] **Chap. 2**, **sect. 2.3.1.** Dans les bases d'images institutionnelles, cette fonctionnalité est souvent utilisée dans les « pages d'accueil », les « menus » et les « rubriques informatives », cf. **chap. 5**, **sect. 5.2.**

sur différents types de composants graphiques de la page Web (texte, icône, image). Pierre Lévy (1991 : 66-67) a déjà caractérisé l'hypertexte à travers six principes :

- Le principe de métamorphose (le réseau hypertextuel est constamment en construction et en renégociation, car il est un « enjeu permanent pour les acteurs concernés »).
- Le principe d'hétérogénéité (les nœuds et les liens d'un réseau hypertextuel sont hétérogènes).
- Le principe de multiplicité et d'emboîtement des échelles (parce que « n'importe quel nœud ou n'importe quel lien » peut lui-même être composé comme un réseau, et ainsi de suite).
- Le principe d'extériorité (le réseau hypertextuel dépend toujours d'un « extérieur indéterminé »).
- Le principe de topologie (le « réseau n'est pas dans l'espace, il *est* l'espace »).
- Le principe de mobilité des centres (le réseau hypertextuel « n'a pas de centre, ou plutôt, il possède en permanence plusieurs centres »).

De ces principes, nous retenons surtout le « principe de topologie » et l'idée que le réseau hypertexte *est* un espace. Celui-ci n'est pas observable. Il est pourtant indéniable que c'est à travers cet espace invisible des hyperliens que se déploie le parcours de l'utilisateur. Nous remarquerons que les larges ensembles d'images mises en place par WF proposent une perspective parallèle à celle des institutions patrimoniales, notamment en s'appuyant sur *cette spatialité du système hypertextuel*.

Dans le cas de WF, la communication en ligne repose plutôt sur ce que nous pouvons appeler « la structure non linéaire et hyperliée du Web » (Zimmer, 2009 : 107-108). Michael Zimmer remarque à ce propos que l'hypertexte fonctionne comme « assemblage non séquentiel d'idées » servant des parcours de lecture qui sont déterminés par les besoins et les intérêts du lecteur (*ibid.* : 110). Le mode d'exploration à travers des « renvois » est déjà fondamental dans les encyclopédies, de telle sorte qu'il a naturellement été intégré dans les encyclopédies en ligne afin de rendre possible « une navigation ouverte[382] et non déterminée de l'information » (*ibid.*). C'est pourquoi, selon Zimmer, il faut étudier les rôles et les fonctions des liens hypertextes dans la continuité

[382] Voir aussi les discussions sur la médiation ouverte dans le **chap. 2, sect. 2.2.2.**

des renvois encyclopédiques, car cela permettra de comprendre « la manière dont nos nouveaux systèmes médiatiques encadrent l'information et la façonnent » (*ibid.* : 110)[383].

Ces « hyperliens » occupent une place centrale dans de nombreuses activités de navigation Web[384]. Si cette fonctionnalité des sites Web est déjà présente dans les bases d'images numériques des institutions patrimoniales, cette structuration à travers un réseau d'hyperliens constitue à peu près les seules composantes graphiques qui sont manifestées dans les services Web de WF, et surtout les seules composantes avec lesquelles l'utilisateur peut interagir. Dans les sections suivantes, nous verrons que ce réseau hypertexte est à la base du fonctionnement des deux bases d'images de WF :

- Une médiathèque en ligne sur le modèle de l'encyclopédie en ligne.
- Une encyclopédie des arts visuels en ligne.

Dans le premier cas, nous verrons que la différence principale entre ce service et les bases d'images institutionnelles concerne le type d'images qui sont rassemblées dans ce service Web, mais aussi l'origine de ces images. Plus généralement, ces deux cas indiquent que les images sont principalement classées en fonction d'un système de relations entre des *catégories générales* et des *sous-catégories*. Même si cette notion de « catégorie » est proche de celle de « mot-clé », nous verrons que l'association systématique des hyperliens avec ce principe de catégorie met en scène des conditions de lecture et de visibilité différentes de celles qui sont prévues par les institutions patrimoniales.

6.1.1. *WikimediaCommons* : une stratégie participative

Parmi les services mis en œuvre par WF, WC est présenté comme une *médiathèque multilingue* comprenant plus de 80 millions de ressources numériques[385] (photographies,

[383] En effet, nous avons vu que le numérique influence la manière dont les connaissances sont manipulées, cela même si ce support s'approprie les caractéristiques d'autres régimes médiatiques, ainsi que les formes sémiotiques d'autres pratiques médiatiques pour que ces connaissances restent intelligibles ; cf. **chap. 1, sect. 1.4.**

[384] À ce propos, cf. notamment les travaux de Jeanneret & Souchier (1999), de Vandendorpe (1999 & 2015), de Boullier *et al*. (2004) ou encore les contributions dans Salaün & Vandendorpe (2004).

[385] *Commons* agglomère et catégorise seulement des fichiers multimédias. Concernant ces fichiers multimédias, nous pouvons noter que certains formats sont acceptés ou non selon le type de fichier. À titre d'exemple, *Wikimedia Commons* ne stocke pas des données binaires ou des programmes exécutables (p. ex : des *widgets* ou des *applets*). Les formats pris en charge sont limités, mais un large panel de formats est accepté, notamment : les formats images : SVG, PNG, JPEG, GIF, TIFF, WebP, XCF ; les formats audio : Ogg, MIDI, WAV, FLAC, MP3 ou Opus ; les formats vidéos : WebM ou Ogg ; les formats

diagrammes, animations, vidéos, musiques et autres fichiers audio). La particularité de ces ressources est qu'elles sont partagées par des utilisateurs volontaires et non pas des institutions patrimoniales. En effet, n'importe quel utilisateur peut ajouter des contenus, mais pour intégrer un fichier il faut respecter strictement deux conditions principales. D'une part, le fichier doit être sous une licence libre ou dans le domaine public. Un contenu qui n'est pas protégé par le copyright est généralement considéré dans le domaine public[386]. D'autre part, le service n'héberge que des contenus dont le but est éducatif, instructif ou informatif. Pour intégrer des fichiers, il s'agit donc de justifier leurs usages éducatifs potentiels. Les collections de photographies privées ou les photographies personnelles, les créations artistiques créées par des contributeurs (sans usage éducatif) et les images publicitaires ne sont pas acceptées.

Cette médiathèque en ligne sert de lieu en commun pour le stockage d'images et d'autres médias visuels en commun. Ces ressources servent à illustrer les divers projets de WF, mais ils constituent aussi des ressources en libre partage au service des internautes (pôle référentiel)[387]. Cette organisation veut devenir un acteur central de la diffusion de connaissances sur Internet, comme le suggère la documentation sur WC :

> Imaginez un monde dans lequel chaque être humain pourrait librement accéder à l'ensemble des savoirs. Vos contributions peuvent faire partie de cela. Lorsque vous partagez des photos ou d'autres fichiers sur Commons, et illustrez Wikipédia grâce à elles, votre travail peut être vu par des milliers – et même des centaines de milliers – de personnes de par le monde. Vous aidez aussi à créer une ressource commune qui touche un public encore bien plus grand ; les fichiers de Commons sont utilisés sur des sites web éducatifs, dans les médias d'actualités, par les bloggers, les artistes, les réalisateurs de films, les étudiants, les enseignants, et beaucoup d'autres personnes[388].

d'animation : GIF, APNG, SVG animé ; formats documents multipages (livres, journaux, magazines, etc.) : PDF et DjVu ; des formats de structures 3D : STL.

[386] La communauté de WF insiste sur le fait qu'un contenu numérique accessible au public ne signifie pas forcément qu'elle soit tombée dans le domaine public ; dans certains cas, pour ajouter un contenu, il faut prouver que ce contenu appartient au domaine public.

[387] Voir **chap. 2, sect. 2.2.1.**

[388] https://commons.wikimedia.org/w/index.php?title=Commons:First_steps/fr&uselang=fr.

Si WC fonctionne comme un instrument public de stockage de contenus multimédias, ces pratiques « collectives » d'archivage numérique sont-elles en concurrence avec les autres pratiques d'archivage d'images « institutionnelles » ?

D'emblée, la réponse négative semble la plus évidente, puisque les institutions patrimoniales et WC ne concernent pas le même type de ressources et le même type de public cible. Néanmoins, ces pratiques entrent en concurrence à un autre niveau, celui de la fréquentation ou, plus précisément, de l'audience. Ce point sur la concurrence est important, car c'est sans doute cette concurrence qui a motivé le renouvellement stratégique des institutions[389].

Comme nous l'avons suggéré, WF propose un cadre général et des normes techniques afin d'intégrer les images dans un ensemble cohérent de ressources. En retour, la participation active des internautes est fondamentale, car ce sont eux qui fournissent les images. Par conséquent, le rôle de ces utilisateurs est différent. L'apport des utilisateurs concerne avant tout l'ajout de médias, tandis que le processus d'organisation des médias dans les bases d'images est géré par la communauté de WC.

À la différence des services des institutions, cette communauté est composée par n'importe quels utilisateurs du service désirant participer au projet, même s'il faut noter que cette communauté d'utilisateur est modérée par des « administrateurs ». Ces administrateurs sont des utilisateurs qui ont des autorisations particulières en ce qui concerne la modération et l'éditorialisation des ressources. En simplifiant, l'administrateur vise : (1) à gérer la suppression ou la restauration des fichiers ou des pages, mais aussi renommer les fichiers ; (2) à protéger des pages ou annuler cette protection ; (3) à bloquer ou débloquer des utilisateurs, à ajouter ou à supprimer des groupes d'utilisateur ; (4) à modifier certains aspects et informations de l'interface Web et envoyer des messages à des utilisateurs. Mis à part ces rôles de modération technique, les administrateurs n'ont pas d'autorité sur les contenus qui sont diffusés par le service[390].

[389] Cf. **infra, sect. 6.2. & chap. 7**.

[390] Selon la documentation du site, si l'influence des administrateurs a beaucoup d'impact sur la communauté, cette influence découlerait de la large confiance acquise par ces administrateurs ; cf. https://commons.wikimedia.org/wiki/Commons:Administrators/fr.

6.1.2. Neutralité de la page-écran, hyperlien et classification par catégories

La caractéristique la plus remarquable de WC (et d'autres services de WF) est sans doute le *design* du site Web. En effet, nous constatons que la page-écran est réduite à un appareil simple (textes ; liens hypertextes ; images ; cadres) qui contribue à proposer une mise en scène de la page sans ornements complexes. Cette sobriété de l'interface correspond vraisemblablement à la « neutralité du point de vue »[391], qui est l'un des principes fondamentaux de l'encyclopédie libre *Wikipédia*, et plus généralement, de tous les projets de WF. Dans ces projets, ce principe est présenté comme une caractéristique du discours encyclopédique ; ce discours s'appuie sur le même effet d'objectivité que le discours scientifique, bien que ces deux genres ne se confondent pas.

Or nous avons vu que le numérique et la communication en ligne ne sont pas des dispositifs « neutres » et « transparents ». Cette neutralité du point de vue est plutôt le résultat de la visée rhétorique qui englobe tous les projets de WF : le numérique et Internet sont une opportunité permettant à tout un chacun de participer à la construction d'un ensemble de connaissances *libres* et *partagées*. Pour ce faire, WC ne modifie pas l'idée centrale dans les encyclopédies d'une connaissance organisée en fonction de catégories, voire de thématiques. Dans la même perspective que l'encyclopédie en ligne *Wikipédia*, c'est plutôt la provenance et le type de connaissances accessibles d'une part, et le public cible d'autre part, qui différencient cette médiathèque des bases d'images institutionnelles. En effet, les informations sur ces médias ne sont pas le produit d'un traitement documentaire, de même qu'elles ne visent pas à l'objectivité du discours scientifique. La totalité des ressources médiatiques hébergées sur ce site Web est composée de contenus sous une licence libre, de telle sorte qu'il peut être partagé gratuitement par n'importe quel autre utilisateur, à condition de respecter les termes de cette licence. Ce principe de licence libre est au centre de l'objectif de WC ; afin de rendre publiques des ressources qui peuvent être utilisées par tout le monde, ce projet accepte uniquement des médias dont les auteurs autorisent l'utilisation libre[392].

---

[391] https://fr.wikipedia.org/wiki/Wikipédia:Neutralité_de_point_de_vue.

[392] Même s'il existe d'autres licences acceptées sur le site, les licences libres les plus courantes sont *Creative Commons Attribution Attribution*, *Creative Commons Attribution Share-Alike Attribution*, *Share-Alike*, *Non-Commercial* et *No-Derivatives*. À titre d'exemple, ces deux derniers types de licence ne sont pas acceptés sur *Wikimedia Commons*, de même qu'il est évident que les contenus soumis au *copyright* ou à des droits d'auteurs ne sont pas acceptés sur le site.

La participation collective et le libre partage servent de garantie en faveur de cette rhétorique de la neutralité du point de vue, puisque celui-ci comprend, en théorie, le point de vue de tout le monde. Pourtant, cette sobriété de l'interface Web évoque aussi un autre élément : la conception de la page-écran correspond au *webdesign* de la fin des années 1990 et du début des années 2000. La *Gallery of Web Design History* du *Web Design Museum* permet en effet d'observer certaines similitudes entre les projets de WF et d'anciennes versions de services Web aujourd'hui très connus, tels que *Google*, *Youtube*, *Yahoo !* ou encore *eBay*[393]. Jusqu'au début des années 2000, les sites étaient généralement composés de peu de composants graphiques, hormis un arrière-plan blanc, les inscriptions verbales et les hyperliens, plus rarement des images et des icônes. Dans la même perspective, nous observons que les pages de WC sont majoritairement composées de textes et de liens hypertextes, avec quelques images.

Dans la page d'accueil (fig. 31), en plus d'un exemple aléatoire d'image (image du jour ; média du jour), le cadre « sommaire » résume les principales catégories qui sont utilisées pour organiser les ressources (fig. 32 & 33). En outre, la page d'accueil propose un cadre « chefs-d'œuvre » qui regroupe trois types de sélections d'images :

- Les « images remarquables » (*Featured Images* ; environ 15 000 images).
- Les « images de qualité » (*Quality Images* ; environ 300 000 images).
- Les « images de valeur » (*Valued Images* ; environ 40 000 images).

Ces corpus d'images rassemblent des images auxquelles la communauté a attribué un label de qualité ; les images remarquables sont les plus « extraordinaires » ou « impressionnantes » (fig. 34), tandis que les images de qualité sont des images de contributeurs qui respectent des normes de qualité techniques mises en place par la communauté de WC. Les « images remarquables » sont également catégorisées en fonction de catégories générales qui sont accompagnées d'un pictogramme (fig. 35).

D'emblée, nous pouvons remarquer une continuité avec les autres projets que nous regroupons dans la première tendance stratégique[394] : l'organisation et la classification des images se basent uniquement sur des notions et des concepts relatifs à une langue. En effet, les ressources accessibles sur WC sont regroupées en fonction de « catégories » (*Categories*) et elles sont présentées dans des « galeries » (*Galleries*).

[393] https://www.webdesignmuseum.org/all-websites.

[394] **Chap. 5**.

Les catégories et les galeries sont complémentaires et visent un même objectif : aider l'utilisateur à trouver des médias. Pourtant, elles réalisent différemment cet objectif.

Une catégorie contient tous les fichiers qui sont liés à un sujet, ce qui n'est pas le cas d'une galerie. Autrement dit, tous les fichiers doivent figurer dans au moins une catégorie, mais tous les fichiers ne doivent pas figurer dans une galerie. Les *catégories* sont le résultat d'une indexation automatisée des ressources (ces catégories servent à ressembler tous les fichiers en rapport avec une même catégorie). Tandis que les *galeries* permettent d'organiser ces catégories et de présenter simplement un échantillon des ressources (idéalement, les images de qualité d'une catégorie). Ces ressources sont affichées dans la page avec une légende et différentes résolutions de l'image, voire des hyperliens vers d'autres services de WF. Lorsqu'un média est sélectionné par l'utilisateur, ce média et les métadonnées sont présentés par le moyen du *Media Viewer* (cf. fig. 37 ; la même interface de présentation est utilisée dans les différents projets de WF). En résumé, ce navigateur permet de (1) lire la description de l'image, (2) d'agrandir l'image (ou la visualiser en plein écran), (3) d'en savoir plus sur l'image et son auteur, (4) d'obtenir plus de détails sur le fichier, (5) de partager ou de télécharger l'image.

Ainsi, WC se base sur une structuration hiérarchique à partir de plusieurs catégories générales (*Topics*, *Copyright statuses*, *Media by source*, *Media types*, *Commons*, *User categories*)[395]. Durant la navigation sur WC, des « accès thématiques » (p. ex. : arts, histoire, politique, géographie, religion, sciences, technologies, etc.) renvoient progressivement vers d'autres catégories plus spécifiques. Cette continuité est sans doute à mettre en lien avec le projet *Wikipédia*. En s'appuyant sur le modèle de l'encyclopédie, *Wikipédia* associe en effet les différents articles à une ou plusieurs catégories pour que l'utilisateur puisse facilement accéder aux autres articles correspondant à la même catégorie. Cette catégorisation repose sur un système de classement thématique des articles, de telle sorte que chaque article de *Wikipédia* est classé dans une ou plusieurs catégories (éventuellement dans des sous-catégories) en fonction du sujet qu'il traite ; la relation entre deux catégories s'appuie sur la distinction entre une « catégorie mère » et une « sous-catégorie ». Chaque catégorie renvoie l'utilisateur vers une autre page où sont présentés d'autres sous catégories, et ainsi de

[395] La plupart des catégories sont générales et sont mises en place par WF. Les utilisateurs ont également la possibilité de proposer leurs propres catégories, même si la communauté WF encourage les utilisateurs à réutiliser des catégories déjà présentes dans la base de données.

suite, pages après pages. Si l'utilisateur peut aussi naviguer dans les ressources à travers cette structuration en catégories et sous-catégories, il peut évidemment chercher des images à l'aide de mots-clés dans la barre de recherche.

D'un autre côté, l'aspect de la page-écran illustre bien l'impact de ces catégories sur la structuration en réseaux de cette médiathèque. Par exemple, sur la version anglaise, lorsque nous encodons « *dog* » ou « *cat* », un ensemble de médias concernant des chiens et des chats sont proposés en retour de cette requête. Ces résultats sont éventuellement filtrés (notamment en fonction des licences de réutilisation, des types de fichiers, des formats d'images) et sont répartis en cinq volets par rapport au type de ressources : *Images*, *Audio*, *Video*, *Other Media*, *Categories and Pages* ; dans les volets « *Images* » et « *Video* », les résultats sont présentés sous la forme d'une mosaïque d'images, tandis que les autres volets présentent les résultats par le moyen de listes. Le volet « *Categories and Pages* » (fig. 38-39) renvoie vers de nombreuses *Galleries* d'images en fonction de catégories plus spécifiques et qui rassemblent différents corpus d'images en fonction de plusieurs autres catégories (*anatomy* ; *dogs by colour/pattern* ; *dog behavior* ; *puppies* ; *dog breeds* ; *dogs in art*). Ce volet renvoie aussi vers les pages de la « *Category: Dogs* » ou de la « *Category: Felis silvestris catus* » qui représentent toutes les différentes sous-catégories relatives. Ainsi, la page de la *Category: Dogs* regroupe toutes les sous-catégories relatives dans un système d'arborescence (fig. 40). En outre, cette page apporte quelques informations générales sur le contenu de la catégorie.



Figure 31 : Page d'accueil de *WikimediaCommons*.



Figure 33 : Détail de la page d'accueil (sommaire regroupant les « catégories principales » du service).

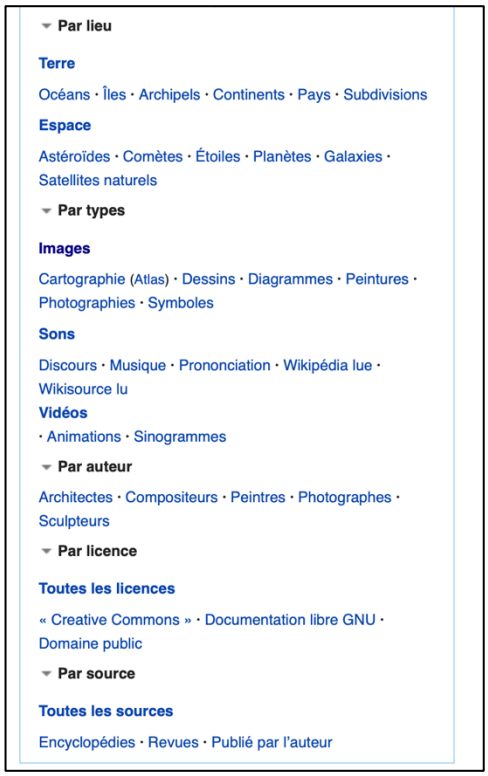


Figure 32 : *Idem* que figure 33.

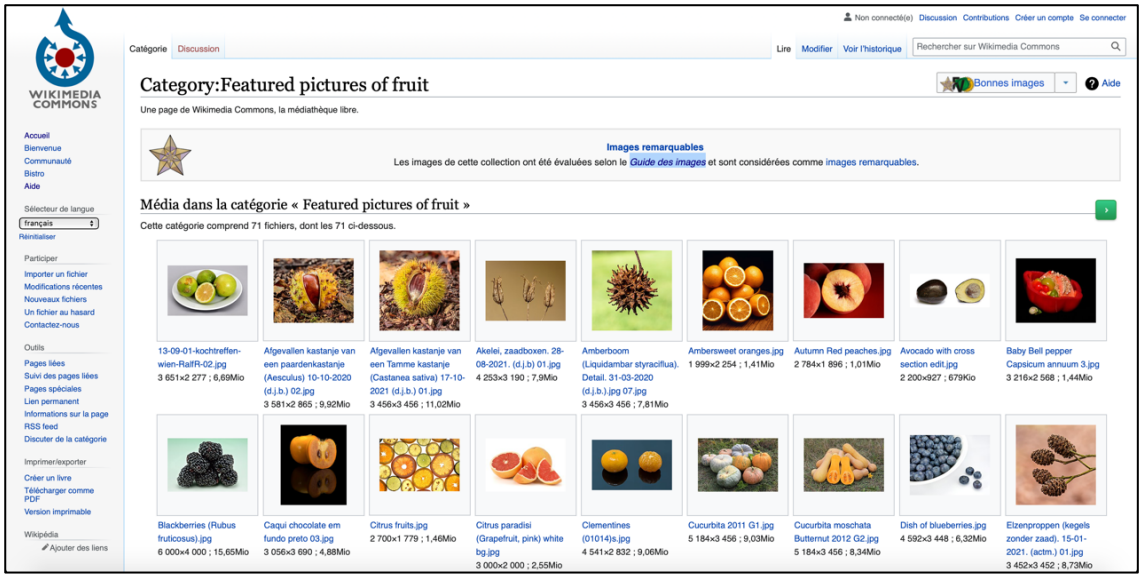


Figure 34 : Exemple d'une page de catégorie (*pictures of fruit*).

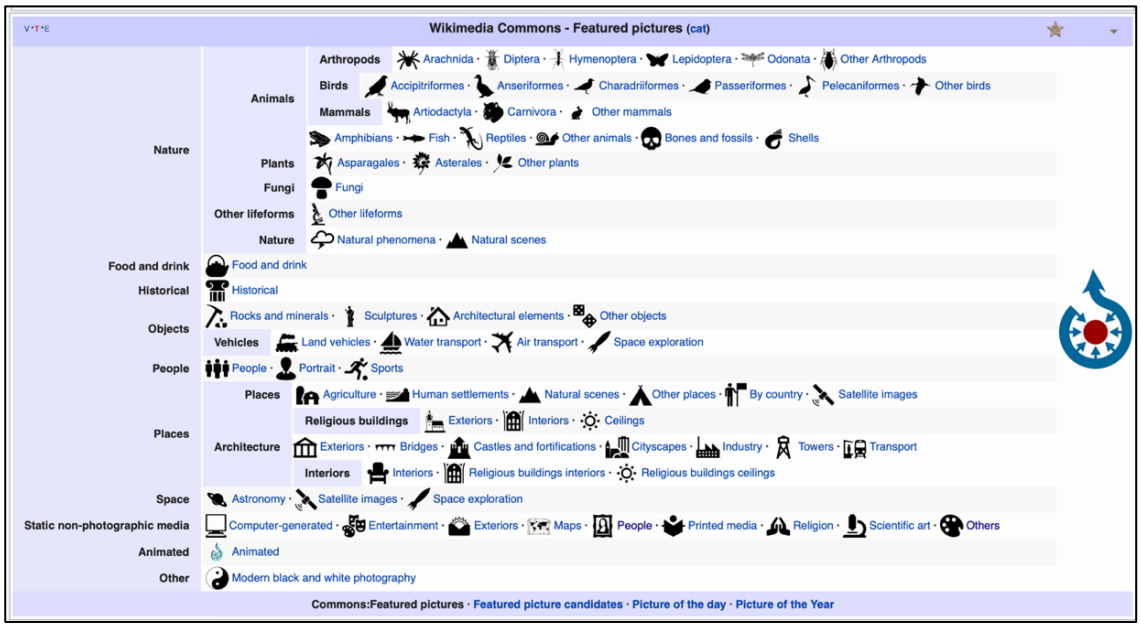


Figure 35 : Cadre (situé en bas de la page de *WikimediaCommons*) regroupant les « images remarquables » (*featured pictures*) en fonction de diverses catégories générales.

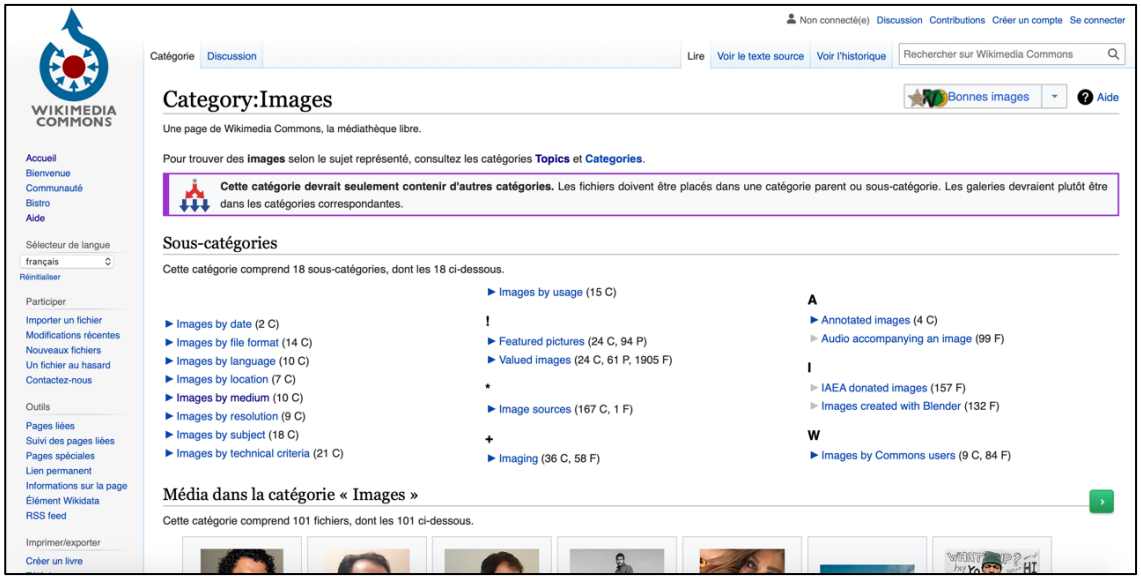


Figure 36 : Page de la catégorie *Images* ; dans le service *WikimediaCommons*, les catégories sont complétées par des « sous-catégories » (images par date, par format, par support, par sujet, etc.).



Figure 37 : *Media Viewer* ; cet outil de présentation est utilisé dans la plupart des projets de la *Wikimedia Foundation*, notamment *Wikipédia.*

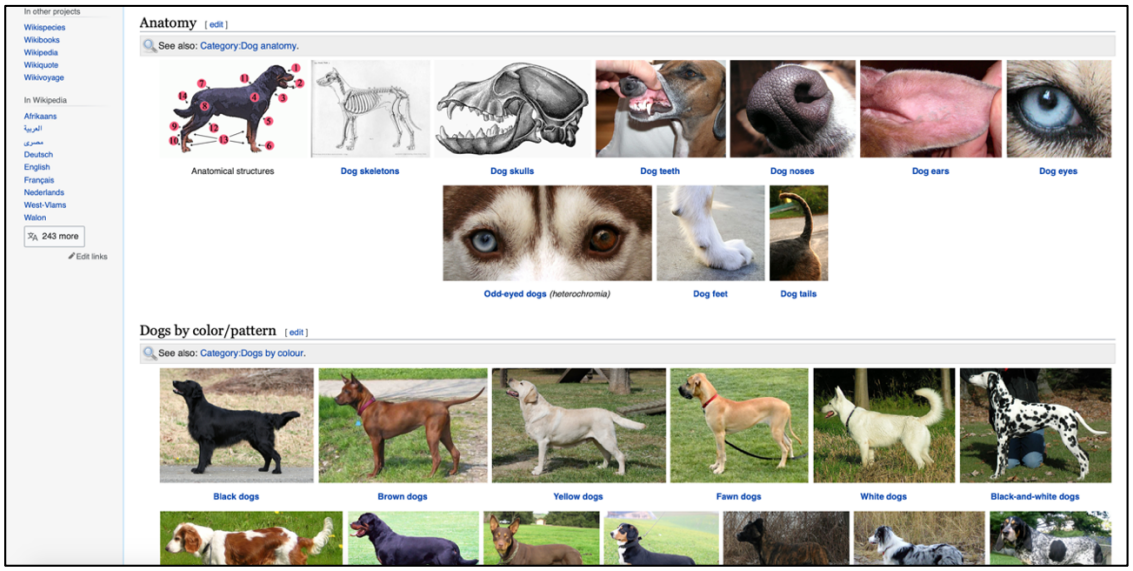


Figure 38 : Page de la catégorie *Dogs* avec différentes *Galleries* correspondant à des sous-catégories.

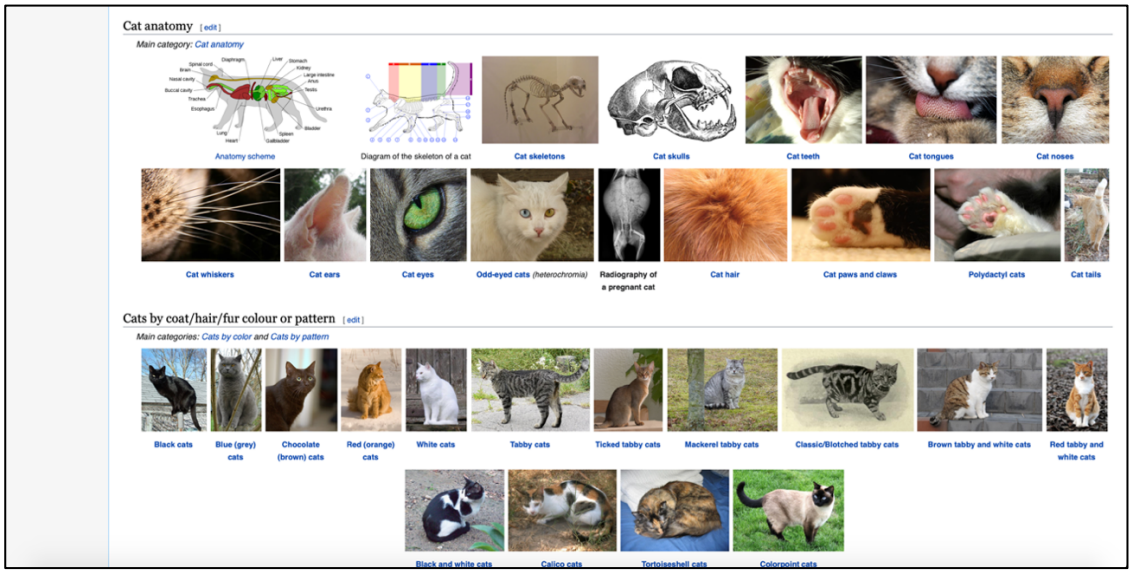


Figure 39 : Page de la catégorie *Felis silvestris catus* avec différentes *Galleries* correspondant à des sous-catégories

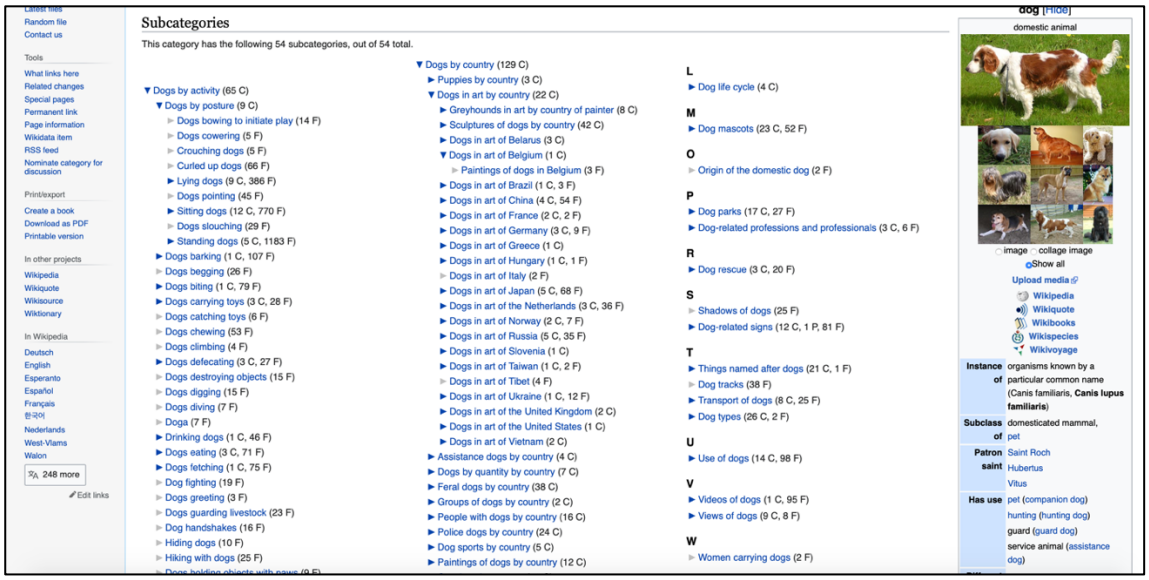


Figure 40 : Liste des sous-catégories de la catégorie *Dogs* ; à droite, quelques informations générales sur cette catégorie.

Comme nous l'avons suggéré, les différentes interfaces Web de WF se basent toutes sur une infrastructure commune aux différents projets. L'avantage est que toutes les ressources sont rassemblées dans un environnement qui se base sur un cadre et des normes techniques partagées. Cette organisation requiert cependant que toutes les catégories soient reliées entre elles au sein d'une même hiérarchie, c'est-à-dire que les catégories sont structurées à travers un réseau en arborescence. Par exemple, dans la catégorie « Images », plusieurs sous-catégories sont distinguées en fonction de paramètres en relation avec l'image (fig. 36). On peut ainsi accéder à des images catégorisées en fonction de la date, du lieu, du support, du format du fichier, de la résolution de l'image, du sujet, de l'usage, etc.

Pourtant, le principal problème concernant la classification des images subsiste : les images sont organisées en fonction de leurs métadonnées descriptives, c'est-à-dire que le support numérique est encore une fois uniquement employé pour organiser automatiquement des ressources numériques en fonction de descriptions simples. Comme les *ontologies dans les outils d'indexation*[396], les *catégories* sont souvent comprises à travers le rapport entre l'*être* et le *langage*.

Le questionnement sur les catégories trouve notamment son origine dans les réflexions d'Aristote, plus précisément, dans le traité des *Catégories* (cf. la traduction de Jules Tricot, 2011 : 15-86). Ce traité qui « a pour objet l'étude du terme et des différents genres de l'Être » (*ibid.* : 7). Benveniste (1966) a lui aussi considéré la conception des catégories défendues par Aristote, surtout pour montrer que celui-ci comprend les catégories comme des « catégories de langue » (*ibid.* : 66).

> Pour autant que les catégories d'Aristote sont reconnues valables pour la pensée, elles se révèlent comme la transposition des catégories de langue. C'est ce qu'on peut dire qui délimite et organise ce qu'on peut penser. La langue fournit la configuration fondamentale des propriétés reconnues par l'esprit aux choses. Cette table des prédicats nous renseigne donc sur la structure des classes d'une langue particulière. Il s'ensuit que ce qu'Aristote nous donne pour un tableau de conditions générales et permanentes n'est que la projection conceptuelle d'un état linguistique donné (*ibid*. : 70).

Plus loin, l'auteur ajoute ceci :

[396] **Chap. 3, sect. 3.3.**

> Il est de la nature du langage de prêter à deux illusions en sens opposé. Étant assimilable, consistant en un nombre toujours limité d'éléments, la langue donne l'impression de n'être qu'un des truchements possibles de la pensée, celle-ci, libre, autarcique, individuelle, employant la langue comme son instrument. […] L'autre illusion est à l'inverse. Le fait que la langue est un ensemble ordonné, qu'elle révèle un plan, incite à chercher dans le système formel de la langue le décalque d'une "logique" qui serait inhérente à l'esprit, donc extérieure et antérieure à la langue (*ibid*. : 73).

Il ne s'agit pas de discuter ici du contenu de cet ouvrage ni de la place de l'étude de l'être dans de la philosophie de l'antiquité. Nous voulons plutôt remarquer à quel point le paradigme d'interprétation logico-grammatical est assimilé dans les outils de classification d'images. Par conséquent, le *style de visibilité* des images est entièrement conditionné par la relation au langage (compétences discursive et sémiolinguistique), de même que l'absence de la matérialité originale est remplacée par des métadonnées descriptives. Ancrée dans ce paradigme, la stratégie de WF accorde une place centrale aux catégories dans ce projet. Partant du présupposé que la langue « est un « système formel ordonné et inhérent à l'esprit », il s'agit de mobiliser des « catégories de la langue » pour organiser des médias visuels. La classification des ressources visuelles est toujours organisée par l'intermédiaire de la langue. Dans WC, cette classification se base sur la manière dont les ressources sont structurées, en l'occurrence, en fonction d'une hiérarchie de catégories (de catégories thématiques à des catégories de plus en plus spécifiques).

Dans ce sens, cette deuxième tendance stratégique perpétue une limitation importante de la première[397], même si nous avons suggéré que d'autres corpus (dans la page d'accueil) font ressortir des images que la communauté estime remarquables en vertu de leurs qualités graphiques et de la composition esthétique des images. Néanmoins, nous pouvons mettre en évidence deux caractéristiques qui distinguent cette médiathèque en ligne des bases d'images des institutions patrimoniales.

(1) La spécificité de ce service est que les images sont fournies par une communauté d'usagers et non pas par une institution patrimoniale. Dans ce cas, s'agit-il du même type de dynamique unidirectionnelle (entre d'une part une institution et d'autre part

[397] Pourtant, cette deuxième tendance est aussi l'occasion de trouver des solutions pour s'éloigner du paradigme linguistique de la compréhension du visuel en mettant en avant les possibilités interactives des sites Web ; cf. **chap. 7**.

l'utilisateur) ? La dynamique n'est plus exactement la même, car l'institution fournit seulement un environnement informatique qui sert de cadre pour une pratique. Le moteur de cette dynamique repose entièrement sur la participation des utilisateurs ; ce sont les internautes qui créent un stock d'images qui est accessible, de même que ce sont les internautes qui produisent les ressources et les métadonnées sur ces ressources. *Orientée des utilisateurs à l'institution*, *la visée rhétorique de ce service semble donc renversée*. Cependant, l'apparente « neutralité » de l'interface n'évacue pas la question de la communauté discursive ni celle des pratiques discursives au sein desquelles ces encyclopédies d'images numériques font sens. Encore une fois, cette communauté est comme diluée par le fait que la communauté (les « internautes ») n'est pas clairement identifiable. En outre, la répartition des rôles énonciatifs est autant problématique que dans les projets précédents : qui est l'énonciateur ? Qui est le destinataire ? Faut-il présupposer qu'il s'agit à la fois de tous et de personne, puisque n'importe quel usager peut participer à l'élaboration de cette archive d'images d'un nouveau type ?

(2) L'importance fondamentale de la participation des utilisateurs est également en rapport avec une autre spécificité de WC : le type d'images qui sont rassemblées au sein de cette médiathèque est différent de celui des bases d'images des institutions. Ce type de service est vraisemblablement laissé de côté par les chercheurs pour une raison que nous avons évoquée dans la **deuxième partie**. En effet, nous avons suggéré que les approches en histoire de l'art se sont d'abord focalisées sur l'étude des images artistiques ou les photographies documentaires rattachées à des institutions patrimoniales[398]. Autrement dit, les analyses et les discours produits par ces approches ne concernent pas des images qui sont récentes ou actuelles, mais plutôt des archives ou des collections d'images concernant un passé révolu.

Cela étant dit, nous présenterons rapidement un second cas parmi les projets de WF. Ce projet concerne des images artistiques, en particulier des peintures, et il s'apparente à une autre forme de communication en ligne que nous analyserons dans un chapitre ultérieur[399]. Dans les paragraphes suivants, nous verrons que les principes de *catégorisation* et de *réseau hypertextuel* sont également omniprésents dans ce projet.

[398] Cf. **chap. 1 (sect. 1.1.) & chap. 3 (sect. 3.1.1., 3.2.1. & 3.2.2.)**.

[399] **Chap. 7**.

6.1.3. *WikiArt* : une encyclopédie en ligne des arts visuels

Si nous pouvons remarquer la publication d'encyclopédies sur les arts visuels à la fin du XIX^e^ siècle – par exemple : les cinq volumes de l'*Encyclopédie des beaux-arts plastiques* (1872-1880) d'Auguste Demmin[400], encyclopédie accompagnée de 6000 photographies d'illustration concernant l'architecture, la sculpture, la peinture et la gravure – notons que des entreprises plus récentes envisagent d'utiliser le numérique pour mettre en place le même type d'encyclopédie, mais sur Internet.

WikiArt est une encyclopédie en ligne des arts visuels, mais à l'échelle du monde entier, afin de concerner l'histoire mondiale de l'art mondial (cf. fig. 41 : page d'accueil). Il s'agit d'un projet sans but lucratif, mais qui utilise toutefois la publicité comme source financière pour couvrir les frais d'hébergement des données numériques et les dépenses liées à des services techniques de maintenance. En effet, le projet défend le même principe que les autres projets de WF, puisque les ressources numériques sont gérées par des bénévoles et les utilisateurs de l'interface Web sont invités à contribuer au projet. L'objectif principal du projet est de rendre accessibles à peu près 250 000 œuvres artistiques de 3 000 artistes, mais encore des versions en plusieurs langues des informations sur ces œuvres[401].

Dans leurs versions physiques, ces œuvres sont conservées et exposées dans des musées, des fonds d'archives d'images, des universités ou d'autres institutions patrimoniales de différents pays. Cependant, la plupart de ces œuvres ne sont pas nécessairement accessibles pour le public dans ces institutions. Cette encyclopédie des arts visuels rassemble à la fois des œuvres du domaine public et des œuvres protégées par des droits d'auteur, mais ces œuvres protégées sont également soumises au principe du *fair use*[402] lorsque ces œuvres sont « historiquement significatives », « utilisées à des fins

[400] Cf. https://fr.wikisource.org/wiki/Encyclop%C3%A9die_des_Beaux-arts_plastiques.

[401] Voir aussi le projet *Google Arts & Culture* (anciennement *Google Art Project*) qui est un service Web de *Google* qui propose des visites virtuelles de différents musées ou projets sur le patrimoine culturel. L'objectif est d'agglomérer et de visualiser des reproductions numériques d'œuvres d'art ou de photographies en fonction de thématiques variées (artistes, supports, mouvements artistiques, événements historiques, personnages historiques, etc. ; cf. https://artsandculture.google.com.

[402] Aux États-Unis, le *fair use* est un ensemble de règles de droit qui apportent des limitations et des exceptions aux droits exclusifs d'un auteur sur son œuvre. Ces règles sont consignées dans le code légal des États-Unis (https://www.law.cornell.edu/uscode/text/17/107). Le but est de prendre en compte les intérêts des bénéficiaires des droits d'auteur, mais aussi l'intérêt public, en ce qui concerne certains usages (principalement non commerciaux et éducatifs) d'œuvres qui sont protégées par le droit d'auteur. Selon certaines conditions, il est donc possible d'utiliser des œuvres protégées sans pour autant enfreindre la loi.

informatives et éducatives », « facilement accessibles sur Internet », « en tant que copies en faible résolution », et « ne convenant pas à un usage commercial ».

Ce projet est un exemple d'hybridation entre « musée virtuel »[403] et « encyclopédie en ligne » (cf. **supra**). En suivant le principe du musée, WikiArt donne à avoir des substituts d'œuvres d'art dans un espace virtuel (un musée « sans murs »). En suivant le principe de l'encyclopédie, WikiArt renvoie infiniment l'utilisateur vers d'autres pages par le moyen d'un réseau de liens hypertextes qui, nous l'avons vu, est au fondement de la stratégie qui aménage les interfaces de WF.

Comme dans WC, nous observons une mise en scène neutre et une organisation des ressources en fonction de catégories générales (fig. 41). Ces catégories portent d'une part sur les artistes (mouvements artistiques ; écoles et groupes ; genres ; champs ; nationalités ; siècles ; institutions) et d'autre part sur les œuvres (styles ; genres ; médias). De telle sorte que l'exploration des ressources est uniquement guidée par une compréhension « catégorielle » des œuvres. Cette organisation est le résultat d'une conception en arborescence des relations entre des catégories générales et plusieurs sous-catégories (fig. 43 : WikiArt repose également sur un système de renvois par le moyen de nombreux liens hypertextes vers d'autres pages de WikiArt, mais aussi vers d'autres projets comme *Wikipédia*). À la différence, de la première tendance stratégique, qui présente surtout les ressources par le moyen de « listes », la présentation des images est dans ce cas-ci réalisée par le moyen de « mosaïques d'images » ; c'est-à-dire que les images sélectionnées par l'ordinateur sont juxtaposées au centre de la mise en page et ces images sont accompagnées d'une légende qui présente quelques informations générales (titre ; auteur ; date ; cf. fig. 42).

[403] **Chap. 2**.



Figure 41 : Page d'accueil de *WikiArt*.

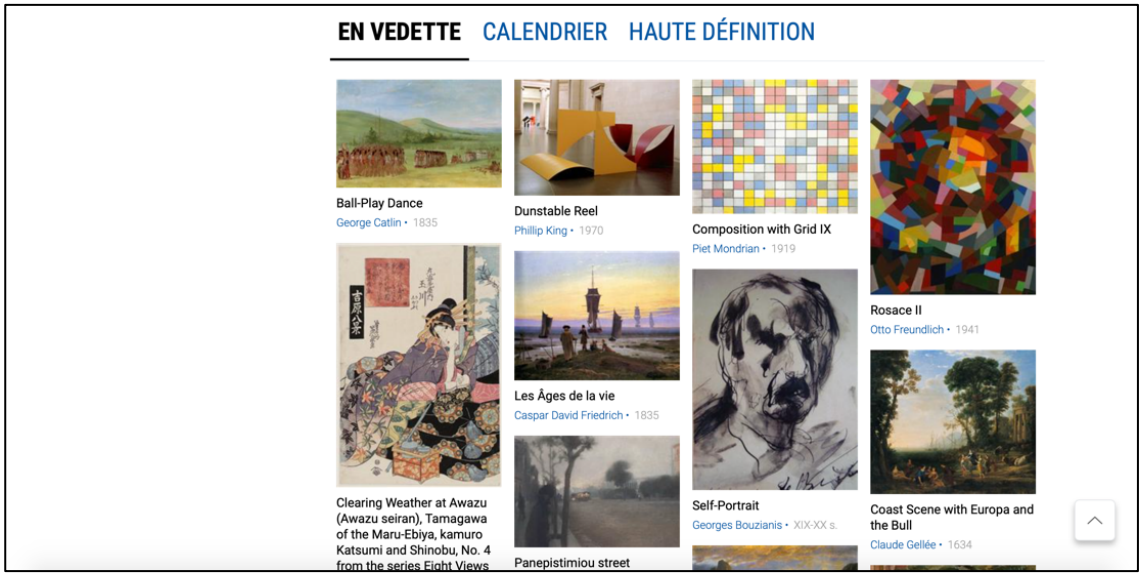


Figure 42 : Présentation de quelques ressources disponibles sur *WikiArt*.

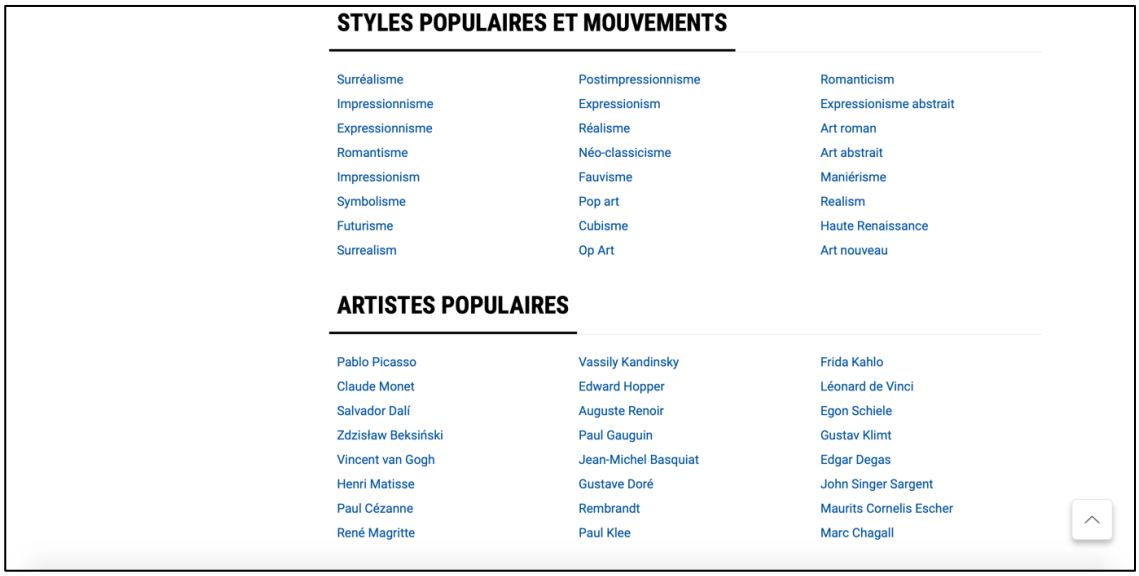


Figure 43 : liste de catégories sur *WikiArt* ; en haut, les catégories concernent des styles et des mouvements artistiques, en bas, les catégories concernent des artistes.

Somme toute, la visée de ces deux services est quelque peu différente de celle des institutions patrimoniales. WC et WikiArt nous ont permis d'observer que des pratiques numériques s'approprient certaines caractéristiques de l'encyclopédie (rassembler un maximum de connaissances ; classification par thèmes ; neutralité du point de vue) pour organiser des bases d'images numériques. Ce projet est également un bon exemple du *principe de structure modulaire* des nouveaux médias[404]. Étant donné que tous les projets de WF se basent sur une même infrastructure informatique, les projets peuvent plus facilement constituer un réseau cohérent de services Web et, par conséquent, optimiser la navigation à travers l'interconnexion des différents projets et interfaces de WF. Cet apport est observable dans l'utilisation intensive de renvois hypertextuels intégrés dans l'interface Web, ce qui favorise une multiplication des « chemins » à parcourir.

WF fait ressortir des choix qui visent davantage à mettre en scène des ressources à travers un *réseau hypertextuel*. Ceci indique une métamorphose importante en ce qui concerne la conception de la scénarisation des parcours qui sont préparés pour l'utilisateur : la généralisation du modèle hypertextuel au sein du réseau Internet aboutit sur d'autres dispositifs de médiation, qui aboutissent sur de nouvelles pratiques d'exploration d'une base d'images numériques.

[404] **Chap. 4, sect. 4.1.2.**

Nous observons notamment que la particularité de l'un de ces services (WC) est que les ressources sont fournies par les utilisateurs du service Web. Si le cadre et l'infrastructure informatique sont le produit d'une institution, c'est le principe de participation des internautes qui est au fondement même de cette médiathèque. Ainsi, nous comprenons mieux les motivations de WF, à savoir : impliquer directement l'utilisateur dans la constitution d'une « base d'images numériques collective » et l'organiser à partir du modèle de l'encyclopédie. L'expression « libre partage » prend ici tout son sens : ces archives d'images sont le résultat d'une participation collective, ce qui n'induit pas que les ressources sont de qualité moindre par rapport aux archives institutionnelles.

Ce déplacement est important et doit être pris en compte pour comprendre l'évolution vers les tendances stratégiques qui suivent, car nous verrons l'effort qui est produit par différents projets et institutions pour concevoir une communication en ligne dont le but est de faire participer le public.

6.2. Agglomérer et standardiser les bases d'images numériques

Dans la continuité de ce que nous avons dit à propos de la communication en ligne d'images[405] et de l'indexation d'images[406], nous constatons un renouvellement stratégique à travers une volonté de standardisation du processus de description des images. Étant donné que la communication informatique repose sur une programmation informatique, il s'agit de proposer un cadre en commun pour faciliter la conception des services Web qui permettent de rendre accessibles de larges ensembles d'images numériques. À titre d'exemple, les propositions de l'*International Image Interoperability Framework* (IIIF) regroupent des normes permettant de diffuser à grande échelle des archives et des collections d'images.

Depuis 2015, plusieurs institutions culturelles ont en effet constitué une communauté internationale de spécialistes qui développent des solutions pour mettre en place des logiciels et des applications informatiques pour améliorer la communication en ligne d'images. Le but de l'IIIF est par conséquent de normaliser les pratiques de diffusion d'images et de fichiers audiovisuels à partir d'une infrastructure qui est adéquate

[405] **Chap. 2**.

[406] **Chap. 3, sect. 3.3.**

au Web. Dans la continuité de la conversion des archives d'images en ressources numériques, il est nécessaire d'envisager la standardisation de ce type de ressources à une échelle internationale, de telle sorte que les institutions mettent désormais en place cette conversion dans un cadre de normes partagées. Pour ce faire, les normes de l'IIIF concernent principalement deux dimensions : la (1) *transmission* et la (2) *visualisation de ressources numériques*. Dans la plupart des cas, ces informations proviennent des institutions locales qui, nous l'avons vu, traitent les archives d'images et utilisent un système et des logiciels informatiques pour organiser toutes les métadonnées descriptives et autres informations sur ces archives. L'IIIF propose un schéma général[407] qui fait ressortir les principaux éléments et le fonctionnement du cadre de normes proposé. Dans ce schéma (cf. fig. 44) :

- *Image Server* correspond à l'ordinateur ou au service où sont stockées les ressources numériques en attendant que les utilisateurs y accèdent.
- *Image API* paramètre les serveurs d'images afin de transmettre tous les pixels d'une image à un utilisateur (image en taille réelle ; image réduite ; rotation ; version en noir et blanc).
- *Presentation API* associe des métadonnées aux ressources numériques qui sont présentées dans l'interface utilisateur ; des fichiers JSON regroupent tous les principaux éléments descriptifs d'une image comme le titre, la description et les informations sur les droits, à travers un *Manifest*, qui regroupe et liste toutes les informations sur les ressources, mais en indiquant aussi comment ces ressources et ces informations doivent être affichées dans l'interface[408].

Afin de connecter et de rassembler ces ressources au sein d'un environnement ergonomique, le protocole informatique mis en place par l'IIIF permet de visualiser et de manipuler ces ressources à travers un dispositif de lecture. Ces dispositifs sont plus ou moins spécialisés en fonction du type de contenus ou des fonctionnalités prévues dans l'interface Web (par exemple : effectuer des panoramiques, des zooms ou des rotations, redimensionner ou annoter des images, comparer des images d'une même collection côte à côte, etc.).

---

[407] https://iiif.io/get-started/how-iiif-works/.

[408] En plus d'afficher des formats d'images (JPEG) ou vidéo (MPEG-4) en fonction de tailles définies, il s'agit d'un cadre de normes et de protocoles pour aider les institutions dans la conception de service Web : cf. la discussion sur les *frameworks* dans le **chap. 2, sect. 2.3.1.**

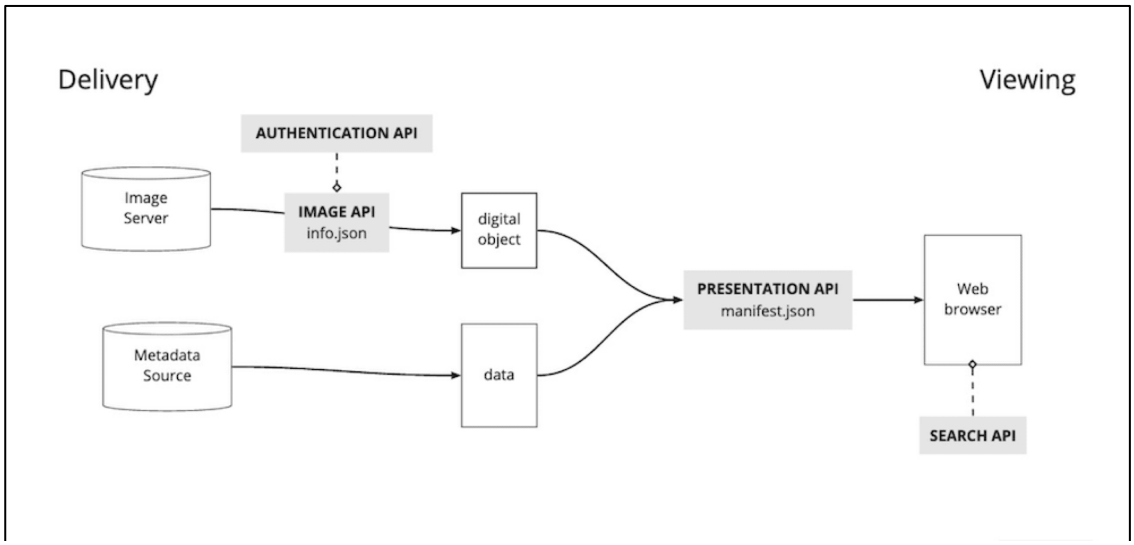


Figure 44 : Schéma du processus et des normes de l'*International Image Interoperability* ; cf. https://iiif.io/get-started/how-iiif-works/.

Dans les sections suivantes, nous présenterons un autre exemple de stratégies de coopération. Au contraire de WC, nous verrons que cette stratégie est étroitement liée à une mission et des objectifs scientifiques. Le principal but de cette stratégie est l'agglomération de données et d'images numériques pour servir la communauté scientifique. Nous verrons que le service Web qui en résulte renouvelle la conception des bases d'images en intégrant des discours de vulgarisation. Cependant, ce service maintient les mêmes fonctionnalités des services conçus par les institutions patrimoniales, notamment en ce qui concerne la classification/sélection/présentation des images.

Nous avons déjà remarqué la place centrale du principe de coopération dans la deuxième tendance stratégique. Ce principe de coopération est également observable dans le projet *Europeana* dont l'objectif est justement d'agglomérer des ressources numériques issues de différentes institutions patrimoniales. Le but évident du projet est de renforcer l'identité européenne en regroupant le patrimoine numérique de plusieurs institutions patrimoniales de plusieurs pays européens[409]. D'autre part, nous remarquerons aussi que

[409] Au sujet d'Europeana, Marie Després-Lonnet (2013) propose des remarques critiques très intéressantes au sujet de ce sentiment européen : « Décider de rendre visible le "patrimoine culturel européen" est un acte politique fort. Disposer d'un patrimoine culturel commun est l'un des fondements d'une société. Concevoir un site dédié à la mise en valeur d'un patrimoine à l'échelle européenne ne peut donc consister à rassembler ce que chaque pays et chaque institution dans ce pays, prise dans une logique locale, auront reçu et désiré transmettre localement. La réussite d'un tel projet passe par la revendication et la construction de ce patrimoine global. La première question n'est donc pas technique. Elle concerne la part de culture que nous voulons, en tant qu'Européens, conserver et transmettre ou parfois oublier... […] Penser notre patrimoine culturel en tant que patrimoine culturel européen, c'est d'abord décider de ce qui fonde notre

ce projet est l'occasion de standardiser le processus de structuration et d'organisation des métadonnées de ces institutions à travers une coopération internationale.

6.2.1. *Europeana project* : le choix stratégique de la coopération

Le service Web d'*Europeana* est créé en 2008 par la Commission européenne et les Ministères de la Culture de 21 États membres. Le projet est présenté comme le point de rencontre entre les pays membres de l'Union européenne (UE) qui devient par conséquent la communauté globale garante de cette agglomération de ressources numériques. L'idée défendue par l'UE est que le patrimoine constitue des ressources publiques dont nous avons tous intérêts de faciliter et d'entretenir le libre accès. Ce projet s'appuie sur un constat. Le patrimoine européen est actuellement consigné sur toutes sortes d'images et d'objets numérisés qui ont à la fois un intérêt culturel et scientifique (livres, manuscrits, photographies, peintures, émissions de télévision, films, sculptures, partitions, enregistrements sonores, etc.). Pour réagir à cette situation, la mission d'Europeana est de diffuser les ressources numériques de nombreux musées, bibliothèques, archives et institutions en Europe à partir d'un même service Web. Nous verrons que le public ciblé est la communauté scientifique internationale.

À peu près 54 millions de documents provenant de plus de 3700 institutions à travers l'Europe sont accessibles sur ce site[410]. Grâce à cette gigantesque quantité de ressources numériques, Europeana vise à améliorer les techniques de communication en ligne de la documentation multimédia à travers la numérisation et la mise en ligne du patrimoine culturel et scientifique européen (cf. Dugenie, Freire & Broeder, 2017). Pour ce faire, une communauté de spécialistes (des « agrégateurs » ou *aggregators*)[411]

culture commune et ce qui l'incarne. C'est ensuite procéder à des choix, qui feront de certains objets associés à certaines pratiques des objets de patrimoine, que les générations d'Européens se transmettront en tant que tels. Le changement d'échelle nécessite de passer du plan local au plan européen et donc de concilier des représentations parfois antinomiques. L'Europe est riche d'une histoire commune dont nous ne partageons cependant pas tous la même vision. L'institution d'un patrimoine à cette échelle ne peut se faire que dans une subtile dialectique entre mémoire et oubli. C'est un processus qui passe notamment par le choix des objets qui sont porteurs de notre mémoire commune et des discours qui seront tenus sur eux, afin de donner à chaque citoyen européen la possibilité de se l'approprier en tant que fondement de sa propre identité européenne » (cf. Després-Lonnet, 2013 : 83).

[410] https://www.europeana.eu/en/collections/organisations.

[411] Le champ d'action de ces « agrégateurs » concerne différents domaines du secteur culturel (musées, archives, bibliothèques) ou encore différentes thématiques (art, mode, audiovisuel, etc.).

travaille avec ces institutions pour rassembler des données authentiques et fiables à partir des institutions patrimoniales de différents pays ou régions de l'Europe.

À partir de 2012, Europeana s'oriente vers une stratégie du libre partage. En considérant que les utilisateurs et Europeana appartiennent à un même réseau d'activités et qu'ils constituent, par conséquent, une même communauté d'individus, il s'agit de construire un service Web dans une perspective *collective* et, par la même occasion, de *connecter les publics européens à la culture européenne à l'aide de l'informatique*.

La stratégie d'Europeana s'articule autour de cinq principes généraux :

- La *mutualité*, puisqu'il s'agit de créer une communauté qui se base sur l'idée d'un bien commun, agissant de bonne foi, comme elle présume celle des autres.
- L'*accès*, car le projet veut proposer des contenus réutilisables de haute qualité et des outils qui permettent aux usagers d'être créatifs et innovants dans l'utilisation des données.
- L'*attribution* en s'engageant à respecter les droits de propriété intellectuelle et de copyright à travers des avertissements et des attributions légales.
- La *cohérence*, afin de construire un modèle informatique qui correspond aux exigences et aux standards propres au domaine du patrimoine et de la documentation.
- L'*engagement*, c'est-à-dire que les membres de cette communauté d'utilisateurs s'engagent à utiliser ce service et ces outils avec prudence et créativité.

Avec WC et WikiArt, nous avons vu que les institutions patrimoniales ne sont pas les seuls acteurs dans le domaine de la communication en ligne d'images. De plus, ces institutions sont poussées au renouvellement, car les bases d'images numériques sont en concurrence avec d'autres services Web. En effet, afin de concurrencer l'émergence de réseaux sociaux où les images occupent également une place centrale (p. ex. : *Pinterest* ; *Pixabay* ; *Flickr* ; *Instagram* ; *Snapchat*), la question est donc de savoir quelle stratégie employer pour « séduire » les internautes et pour susciter l'intérêt et la participation du public. Comme WF, la visée à long terme d'Europeana est de devenir un acteur majeur dans la diffusion du patrimoine culturel numérique en Europe. Dans les deux cas, l'idée centrale est en effet d'organiser et de « structurer » ce patrimoine comme un bien en commun. Dans le cas d'Europeana, la structuration des données se base sur une

standardisation des données à l'aide d'un cadre normatif comme celui proposé par l'IIIF[412]. Nous présenterons rapidement ce cadre dans la section suivante.

6.2.2. Modélisation et standardisation des données

La deuxième tendance stratégique a pour but de constituer un environnement coopératif. Concernant Europeana, nous avons suggéré que cette coopération entre différentes institutions patrimoniales vise à améliorer la circulation des connaissances et du patrimoine culturel européen sur Internet. Pour ce faire, l'*Europeana Data Model*[413] (EDM) est un cadre normatif qui permet aux institutions de concevoir un service Web qui se base sur une standardisation des métadonnées (cf. Doerr *et al.*, 2010). En effet, l'idée principale d'Europeana est l'accumulation et l'exploitation des informations relatives à des reproductions d'objets ou d'artefacts culturels qui proviennent de contextes institutionnels parfois différents. Ainsi, sur la base de l'EDM, les différentes institutions participant au projet structurent, décrivent et présentent de la même manière les informations servant à contextualiser des archives et des collections d'images numériques. Ce modèle s'accorde sur des processus de standardisation des métadonnées qui utilisent le langage XML[414], ou encore sur d'autres cadres normatifs utilisés dans les pratiques d'archivage numérique : *Lightweight Information Describing Objects* (LIDO)[415], *Encoded Archival Description* (EAD)[416], *Metadata Encoding and Transmission Standard* (METS)[417]. Néanmoins, comme dans la première tendance, ce modèle favorise uniquement une *contextualisation sémantique* des ressources qui sont traitées par le système informatique.

À ce propos, une enquête en ligne a notamment permis de considérer un échantillon des « demandes types » de recherche des utilisateurs, en fonction des types

---

[412] Cf. **supra, sect. 6.2.**

[413] https://pro.europeana.eu/resources/standardization-tools/edm-documentation

[414] L'*Extensible Markup Language*, généralement appelé XML, est un métalangage informatique de balisage générique qui dérive du *Standard Generalized Markup Language* (SGML). Ce langage est dit « extensible » car il permet de définir différents langages avec chacun leur vocabulaire et leur grammaire. L'objectif est de favoriser l'interopérabilité, c'est-à-dire de faciliter l'échange automatisé de contenus informatiques complexes entre des systèmes d'informations hétérogènes.

[415] http://network.icom.museum/cidoc/working-groups/lido/what-is-lido/

[416] https://www.loc.gov/ead/

[417] https://www.loc.gov/standards/mets/

de recherches, des motifs typiques dans cette activité de recherche, mais aussi, les utilisations courantes des informations trouvées (Clough *et al.*, 2017 : 207). D'une manière générale, les utilisateurs n'emploient pas le site comme une fin en soi, mais plutôt dans le cadre de leurs tâches professionnelles. De plus, les différences individuelles des utilisateurs (âge, sexe, connaissance du domaine) affectent la façon dont les gens recherchent des informations (cette enquête met en évidence qu'il est nécessaire de comprendre cette diversité d'utilisateurs à travers les recherches effectuées par le moyen de l'interface Web). Enfin, les résultats de cette enquête indiquent aussi deux défis pour l'avenir de ce type de service Web :

- Étudier les activités de recherche et d'exploration du patrimoine numérique.
- Créer des dispositifs permettant des analyses plus approfondies de l'exploration et de la manipulation de larges ensembles de données numériques.

En effet, les objectifs et les besoins des utilisateurs du service Web sont variés, et c'est pourquoi il est nécessaire de « fournir des systèmes qui vont au-delà de la recherche par mots clés et qui prennent en charge des comportements de recherche d'information plus larges, comme la navigation et l'exploration » (*ibid.*).

Ce constat rejoint sans aucun doute la problématique présentée dans la première partie du présent travail : la communication en ligne des archives et des collections d'images prolonge et impose un paradigme d'interprétation des images qui influence les conditions de visibilité et de manipulation des images. À ce stade, nous constatons que ces choix concernent d'abord la concurrence avec d'autres services, et non pas les limites de l'indexation informatique des images. C'est dans ce sens que les futurs « choix stratégiques » des institutions seront décisifs.

Du reste, nous présenterons les fonctionnalités actuelles du service dans la section suivante. Auparavant, afin de contrebalancer cette dernière critique, soulignons qu'un nouveau plan stratégique propose une « vision » et une « mission » pour la période 2020 à 2025[418]. Cette vision projette qu'Europeana sera un acteur central dans le secteur du patrimoine numérique en Europe. En parallèle, la mission est de développer « l'expertise, les outils et les politiques pour embrasser le changement numérique » et d'encourager « les partenariats qui favorisent l'innovation » dans le but de faciliter l'utilisation du patrimoine numérique par des chercheurs, ou encore pour faciliter l'utilisation

[418] Cf. https://pro.europeana.eu/files/Europeana_Professional/Publications/EU2020StrategyDigital....

personnelle de ces archives d'images dans le cadre d'autres activités et pratiques culturelles. Ce plan stratégique fait ressortir plusieurs objectifs prioritaires :

- Promouvoir les avantages de la description de données par rapport à un cadre de description standardisé des ressources numériques et développer des agrégateurs plus performants.
- Améliorer, d'une part, la qualité des données et des métadonnées, et d'autre part, les processus le traitement automatique des métadonnées.
- Mettre en évidence les avantages du libre partage des ressources.
- Renforcer les capacités de transformation numérique et les infrastructures nationales des différents pays participants au projet.
- Développer la formation et les compétences des acteurs et des réseaux qui participent à la diffusion sur Internet du patrimoine numérique.
- Proposer davantage de services et applications Web aux institutions ou encore de perfectionner les instruments de recherche de données.

Ce dernier point concerne les limites de l'indexation d'images que nous venons d'évoquer ; nous pouvons donc imaginer que, dans les années suivantes, Europeana améliorera les fonctionnalités de recherche dans son service Web. Du reste, la coopération des institutions patrimoniales n'est pas tant l'occasion de renouveler le *design* des interfaces Web ou d'y intégrer de nouvelles fonctionnalités. Au premier abord, ce service semble identique aux bases d'images que nous avons déjà présentées[419], parce que la scénarisation de l'information et les fonctionnalités de l'interface font ressortir les mêmes caractéristiques (espace d'accueil ; champ de recherche ; listes d'images ; filtres et métadonnées linguistiques).

[419] **Chap. 5**.

6.2.3. De l'exploration à la vulgarisation : écrits d'écran et rhétorique multimédia

Nous pouvons tout de même faire ressortir deux différences entre Europeana et les bases d'images institutionnelles (chap. 5).

Premièrement, le dispositif de lecture comprend *trois formes de présentation des images*, même si les images sont toujours classées et sélectionnées en fonction d'une correspondance avec un mot-clé ou en fonction de filtres (thème ; type de support ; agrégateurs ; institutions ; pays ; etc.) :

- Dans la « vue liste » (fig. 46) , les résultats sélectionnés sont présentés les uns à la suite des autres, de haut en bas (ces résultats peuvent également être réduits à l'aide de différents filtres linguistiques, comme dans les quatre exemples de bases d'images numériques institutionnelles).
- Dans la « vue grille » (fig. 47) , les images sont présentées en juxtaposition les unes par rapport aux autres et avec un minimum d'information : le titre, la localisation et l'institution qui fournit la ressource numérique (il s'agit de la même forme de présentation que dans le projet WikiArt).
- Enfin, la « vue mosaïque » (fig. 48) propose un point de vue un peu différent sur les ressources, dans le sens où cette juxtaposition exprime une sélection d'images à partir de descriptions linguistiques, mais sans afficher ces descriptions dans le dispositif de lecture.

Ainsi, certaines caractéristiques de la mise en scène de la communication en ligne sont semblables à la mise en scène des stratégies des institutions patrimoniales locales :

- Un champ de recherche est mis en évidence dans une zone (au-dessus), les résultats sont présentés dans une deuxième zone (à gauche), les filtres dans une troisième zone (à droite).
- Lorsque l'utilisateur sélectionne une image proposée dans les vues « liste », « grille » ou « mosaïque », il est invariablement renvoyé à une page Web spécifique qui présente la version numérique de l'image, ainsi qu'une partie réduite (« bon à savoir ») ou l'ensemble des métadonnées (« toutes les métadonnées ») sur cette image (cf. fig. 49).

Par là même, nous remarquerons que, même à ce niveau plus global (Union européenne), il s'agit toujours de considérer les images et la culture visuelle en fonction

d'une traduction linguistique, et donc de réduire la classification des images à une sélection d'images en fonction de mots-clés qui correspondent à des métadonnées descriptives. En effet, même si chaque forme de présentation correspond à un « point de vue » différent pour présenter et pour visualiser des images, les résultats du traitement informatique sont uniquement sélectionnés en fonction d'un métalangage de description et non des images en tant que telles. Dans un travail sur ce sujet, Carlotta Capurro et Gertjan Plets ont pourtant souligné que les institutions patrimoniales acceptent par là même de sacrifier la précision des descriptions pour que leurs informations soient intégrées dans une base de données à l'échelle européenne (cf. Capurro & Plets, 2020).

En outre, nous pouvons mettre en évidence deux problèmes concernant la gestion des données et des métadonnées qui proviennent de plusieurs institutions.

D'une part, étant donné que les données et les informations agglomérées ont été produites dans plusieurs contextes institutionnels, il est parfois difficile de faire dialoguer plusieurs communautés linguistiques (les métadonnées concernent différentes langues selon leurs provenances) ; pour répondre à ce problème, Europeana est disponible dans les différentes langues des institutions qui participent au projet (fig. 45).

D'autre part, les institutions patrimoniales ont parfois utilisé, selon les régions et les époques, différentes normes documentaires dans le traitement des archives physiques ; pour répondre à cette seconde difficulté, nous avons suggéré que le projet a conçu un modèle de standardisation (EDM).

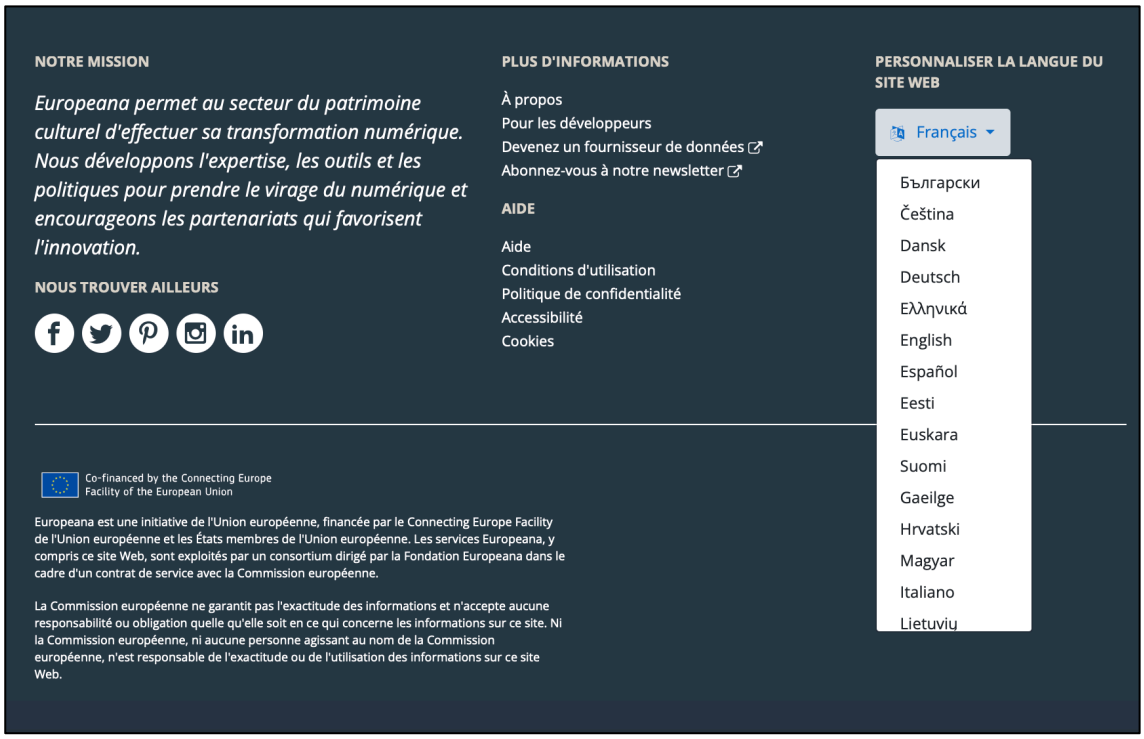


Figure 45 : Information en bas de la page d'*Europeana* avec des informations générales sur la mission et la proposition de personnaliser la langue du site Web.

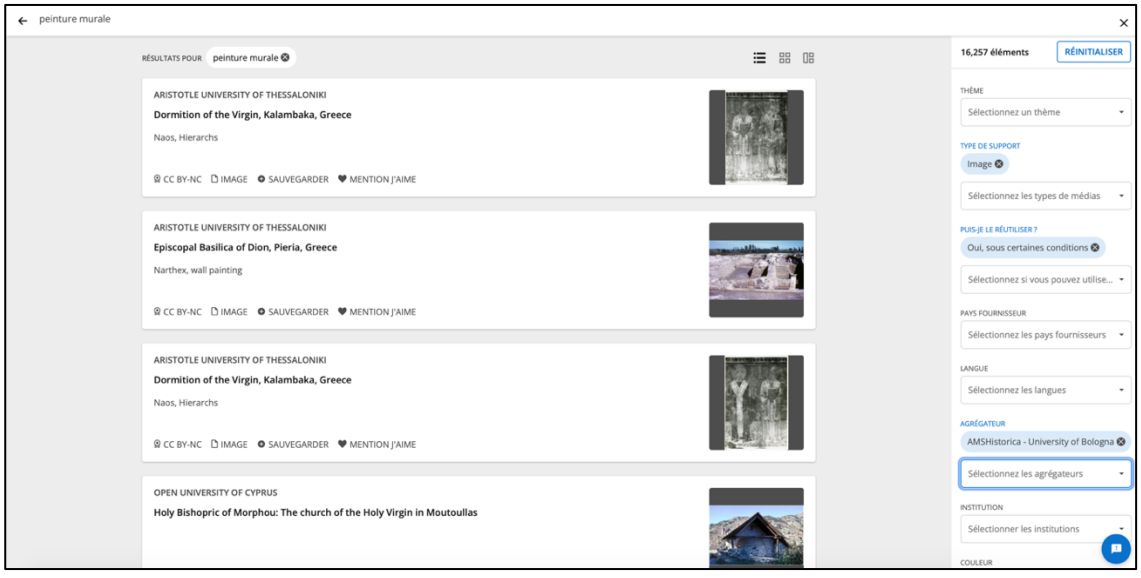


Figure 46 : Exemple de présentation des résultats d'une recherche sur *Europeana* (« vue liste »).

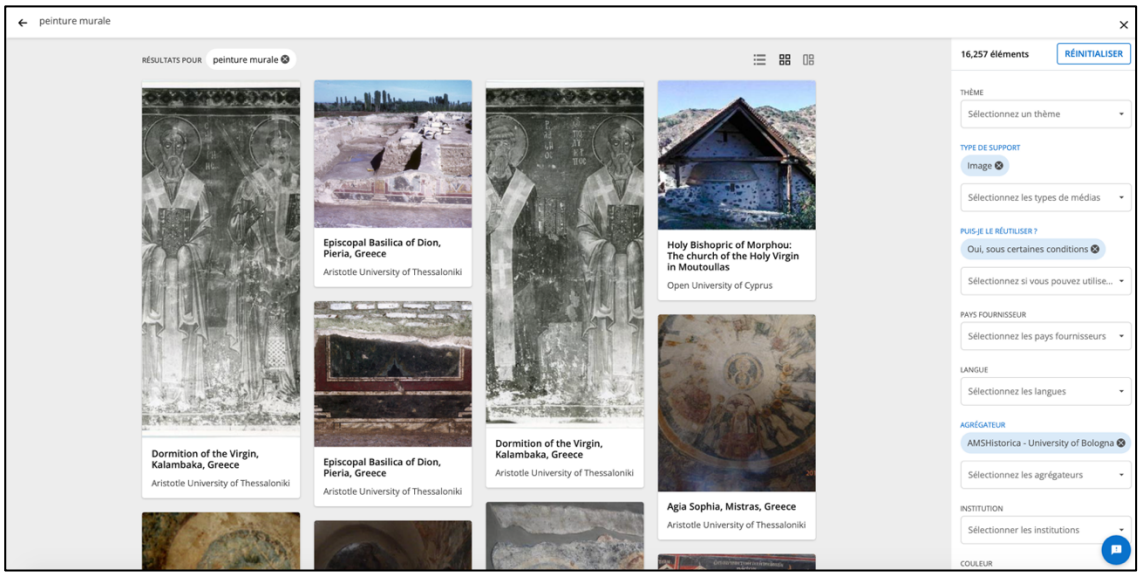


Figure 47 : Exemple de présentation des résultats d'une recherche sur *Europeana* (« vue grille »).

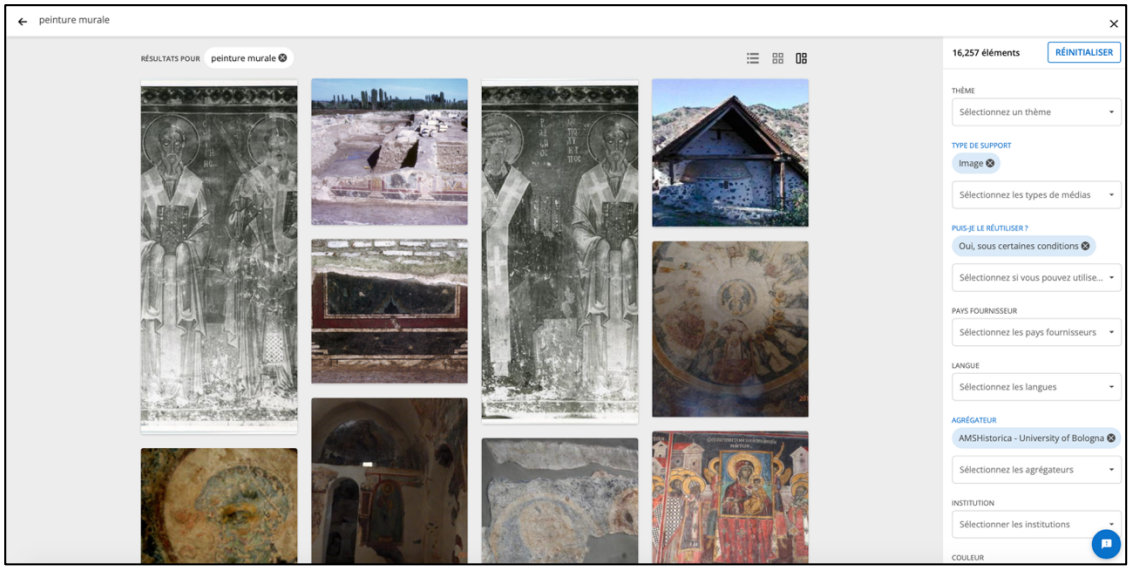


Figure 48 : Exemple de présentation des résultats d'une recherche sur *Europeana* (« vue mosaïque »).

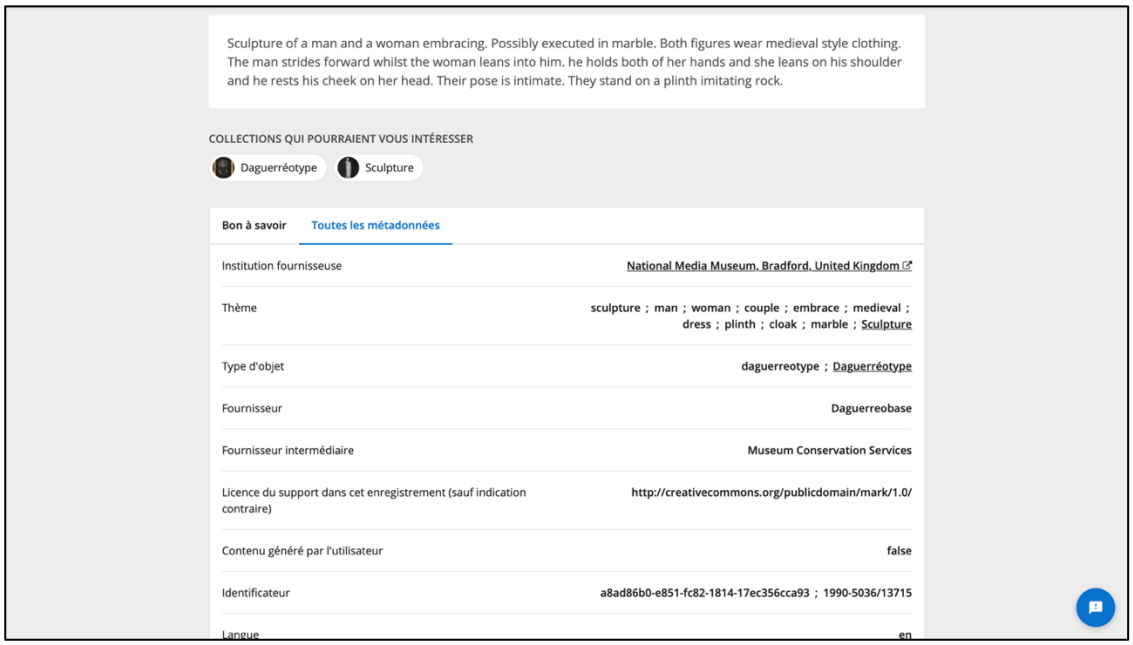


Figure 49 : Détail de la page de présentation d'une ressource d'*Europeana* avec toutes ses métadonnées descriptives.

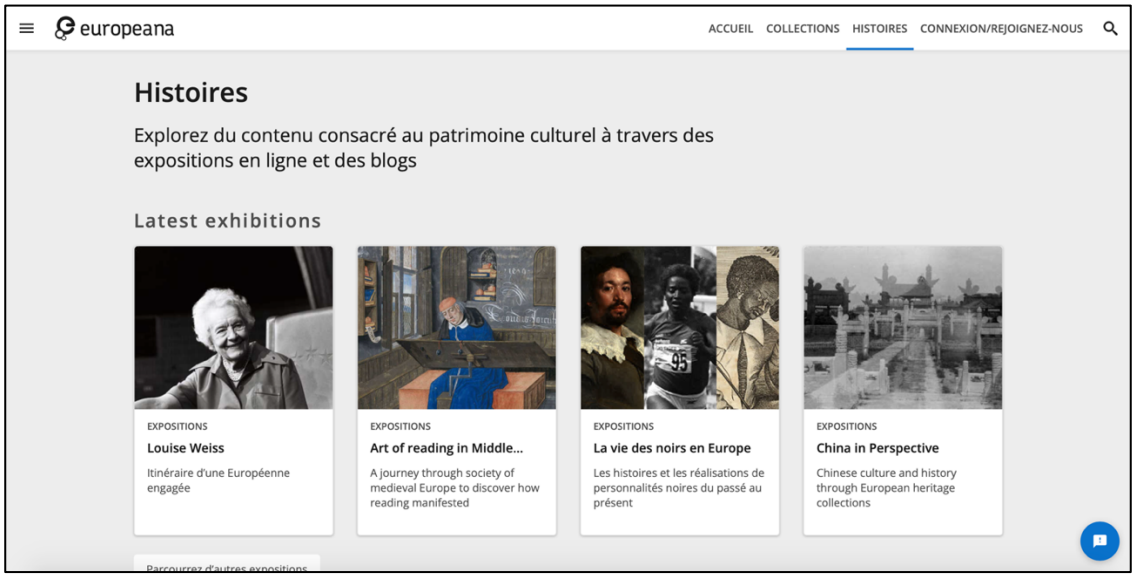


Figure 50 : Page d'accueil de la rubrique « histoires » (*Stories*) regroupant différentes « expositions ».

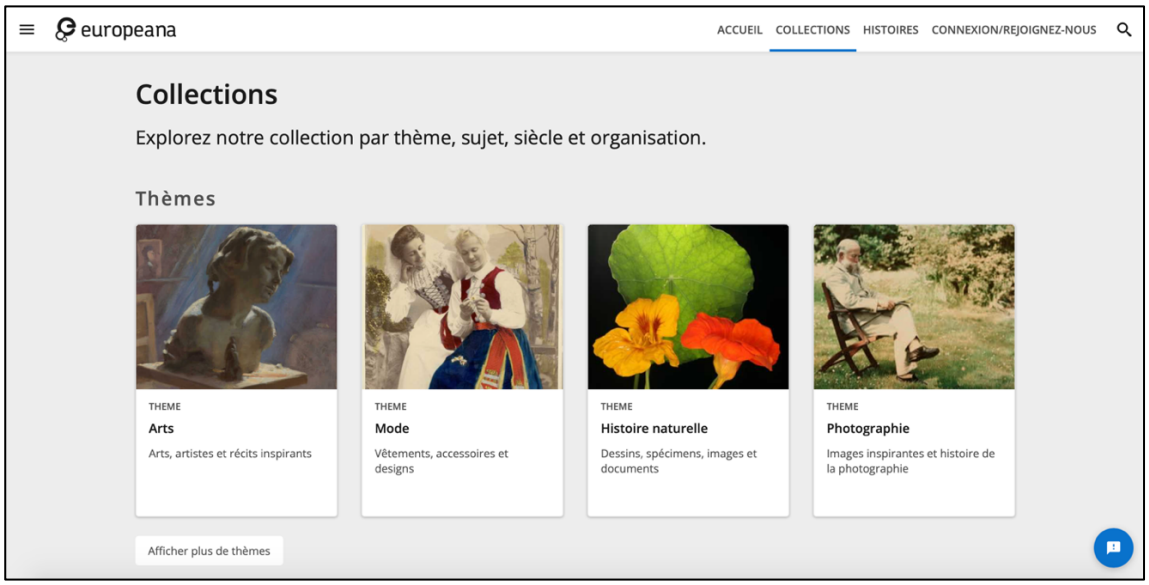


Figure 51 : Page d'accueil de la rubrique « collections ».

Deuxièmement, nous avons suggéré que le service d'Europeana est développé en vue d'un usage scientifique des ressources agglomérées par le projet. L'objectif est notamment d'inciter les chercheurs à découvrir ces ressources numériques et d'encourager la communauté scientifique à utiliser ce service dans leurs travaux et projets. En retour, ces usages valorisent les activités et services d'Europeana au sein de la communauté scientifique.

Cet objectif est observable à travers la composante discursive du service. En plus des sélections d'images numériques qui sont regroupées en fonction de thématiques ou de périodes historiques qui correspondent à des catégories générales (p. ex. : arts, mode, histoire naturelle, etc. ; fig. 51), plusieurs discours sont répartis dans les parties « expositions » (*exhibitions*) ou les « histoires » (*stories*) de ce service Web (fig. 50).

Dans la rubrique « histoires », Europeana fournit un espace Web où des *discours de vulgarisation* sont utilisés pour mettre en valeur les ressources documentaires agglomérées par le projet. Ces discours sont intégrés dans une scénarisation qui a pour but de promouvoir les activités des institutions ou des chercheurs grâce à un travail de réflexion et d'écriture. Dans ce sens, cette partie du projet Europeana est un bon exemple de « médiatisation », c'est-à-dire du « pôle médiatique » qui consiste à valoriser des

activités de recherche à travers une diffusion informatique[420]. Mais cette médiatisation s'appuie sur une autre pratique discursive : la vulgarisation, qui prend dans ce cas-ci la forme d'un discours verbal qui est éventuellement réparti en différents chapitres, et où des images sont intégrées dans le texte pour illustrer le propos.

Par exemple, dans un article en français concernant une exposition sur l'Art Nouveau, nous pouvons observer que le discours est divisé en sept chapitres, où le texte associe l'écrit avec l'image pour illustrer le propos sur l'Art Nouveau (cf. fig. 52-55). L'énonciation est neutre et impersonnelle (effet d'objectivité au discours)[421], et les textes comprennent parfois des hyperliens qui renvoient vers d'autres pages contenant des informations descriptives (étant donné que les images proviennent de la même base d'images, la plupart des informations sur les images sont facilement accessibles). De plus, pour approfondir le sujet considéré par l'article, le texte est accompagné de diverses ressources bibliographiques, voire d'autres ressources disponibles sur Internet qui sont complémentaires (fig. 53). Par rapport à la première tendance stratégique, Europeana renouvelle la conception des bases d'images en intégrant davantage la vulgarisation dans le service. Pour s'en convaincre, il suffit de comparer les rares textes présentés sur les sites Web des institutions patrimoniales aux nombreuses « histoires » d'Europeana (à peu près 1500, toutes langues confondues)[422].

La « vulgarisation » est l'adaptation « des notions, des connaissances scientifiques ou techniques afin de les rendre compréhensibles au non-spécialiste », la « reformulation d'un discours spécialisé qui consiste généralement à le débarrasser de ses difficultés spécifiques, de ses caractères techniques afin de le rendre accessible au grand public »[423]. Concernant, la diffusion de la culture scientifique auprès du plus grand

[420] **Chap. 2, sect. 2.2.1.**

[421] **Chap. 5, sect. 5.2.2.**

[422] Notons que certaines histoires sont traduites en différentes langues. La répartition des histoires en fonction des langues est cependant inégale, et elle est largement au profit de l'anglais (à ce propos, voir les discussions dans le **chap. 3, sect. 3.3.**) : anglais (963) ; français (51) ; italien (48) ; espagnol (48) ; allemand (43) ; polonais (34) ; néerlandais (33) ; roumain (24) ; suédois (20) ; grec (18) ; hongrois (18) ; lituanien (18) ; letton (17) ; finnois (17) ; slovène (16) ; danois (15) ; croate (15) ; portugais (15) ; russe (15) ; maltais (15) ; estonien (15) ; gaelige [irlandais] (14) ; slovaque (14) ; basque (7).

[423] Voir : https://www.cnrtl.fr/definition/vulgarisation. Pour Daniel Jacobi (1985), le rôle de la vulgarisation scientifique est pourtant sous-estimé par la recherche scientifique. En effet, si nous admettons que « le but de la science est d'imposer une connaissance du vrai », il faut alors remarquer que « tous les moyens de diffusion des concepts visant à les faire utiliser et employer par d'autres ont un rôle à jouer » (Jacobi, 1985 : §101).

nombre, Daniel Jacobi (1999) propose de mobiliser une approche sociolinguistique pour étudier les habitudes de lecture propres à la vulgarisation scientifique. En remarquant que les sciences de la communication et la linguistique permettent d'étudier en profondeur les formes sémiotiques et les procédés rhétoriques qui sont employés dans la vulgarisation scientifique, Jacobi a mis évidence que les discours de vulgarisation ne sont pas autonomes de la rhétorique scientifique. En simplifiant, celle-ci peut être comprise comme la production de discours en relation avec une activité scientifique et épistémique où des énoncés et des formes discursives participent à la construction et à la validation des savoirs qui fondent les réflexions scientifiques (pour une autre approche sociolinguistique sur le discours scientifique, cf. par exemple Latour & Fabbri, 1977, ou plus récemment, cf. Rinck, 2010).

Europeana et WF se fondent sur la même idée principale. Les pratiques numériques et le réseau Internet favorisent des activités de « communication numérique écrite », notamment à travers des blogs qui permettent de publier des discours de vulgarisation scientifique. Des internautes peuvent utiliser ces blogs pour partager les résultats de leurs recherches et pour communiquer leurs avis sur un sujet ou sur une question scientifique. Cette vulgarisation scientifique sur Internet est « une forme hybride de discours de transmission des connaissances » où le support numérique « permet à l'interaction de se mettre en place avec des échanges via ordinateur présents dans les commentaires et aussi par l'utilisation des liens hypertextes qui offrent des parcours de lectures infinis » (Reboul-Touré, 2020 : 11). En parallèle du discours scientifique, les *discours de vulgarisation* sont considérés comme un « type de discours », tandis que le *blog de vulgarisation scientifique* peut être compris comme un « genre de discours » (*ibid.*). De même que l'encyclopédie en ligne (ou « forum encyclopédique ») ne se confond pas complètement avec le genre de l'encyclopédie, le blog de vulgarisation ne se confond pas avec la vulgarisation scientifique en tant que genre. Dans le blog, des caractéristiques du discours de vulgarisation ou celles du discours scientifique s'intègrent vraisemblablement dans un nouveau genre discursif.

Dans ces travaux, Ingrid Mayeur s'est interrogée sur les caractéristiques génériques du blog de recherche en sciences humaines, notamment en prenant en compte les contraintes formelles propres à ce type de médiation (cf. Mayeur, 2018). Pour l'auteure, en tant que dispositif, le blog permet d'intégrer les « discours sur le savoir » dans une « mise en scène personnelle de la recherche ». De plus, l'étude des

pratiques d'écriture dans un environnement informatique permet d'engager une réflexion concernant de nouveaux « gestes discursifs » qui concernent une multiplicité d'instances énonciatives.

En effet, Mayeur (2020) a montré comment « l'enrichissement hypertextuel des énoncés » et les commentaires remettent en question le billet scientifique en tant que forme textuelle. Selon l'auteure, les billets sont le résultat d'actes énonciatifs complexes et qui sont susceptibles d'être produits par différents acteurs. Ces actes peuvent aussi être considérés comme des « gestes discursifs du savoir », à savoir un geste discursif qui n'est pas « nettement circonscriptible dans l'énoncé », mais un geste qui permet de « saisir à travers des traces la manière figurale dont un discours progresse » (cf. Lttr13, 2016 : 94-95). Mayeur envisage le commentaire comme « geste discursif du savoir du type discussion, tandis que l'enrichissement hypertextuel manifesterait un partage de ressources » que le lecteur peut choisir d'actualiser en fonction de ses motivations et de son parcours (Mayeur, 2020 : §27).

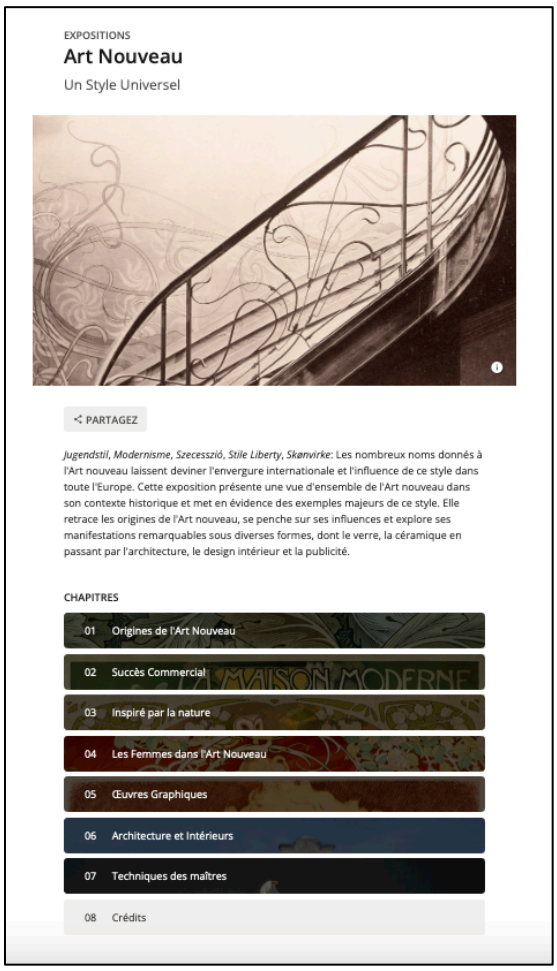


Figure 53 : Exemple d'une « exposition » sur l'Art Nouveau (*Europeana*).



Figure 52 : Cette exposition comprend une rubrique qui renvoie vers de nombreuses ressources sur l'Art Nouveau.



Figure 55 : Même si elle présente quelques images, l'exposition s'appuie surtout sur le discours verbal.



Figure 54 : Voir fig. 55.

Somme toute, la mission générale d'Europeana est accomplie. D'une part, ce projet a mis en place un cadre de normes propres à la communication en ligne de larges ensembles d'images dans le but de collecter, d'interconnecter et d'enrichir des données et des métadonnées qui concernent le patrimoine culturel européen. D'autre part, ce projet propose un service coopératif qui rend possible la recherche dans la plus large agglomération d'images numériques en Europe. Ce service propose trois formes de présentation des résultats ; pourtant, même si ces formes offrent différents points de vue, elles sont conditionnées par la même logique de sélection des images.

Dans le **chapitre précédent**, nous avons déjà montré que (1) les choix stratégiques des institutions patrimoniales favorisent une rhétorique unidirectionnelle (de l'institution à l'usager) et que (2) le support numérique est principalement utilisé pour mettre en place un espace multimédia (où la dimension graphique est toutefois dominante) et qui a deux fonctions principales : soit l'interface fonctionne comme un inventaire soit comme vitrine des activités de l'institution. Dans cet espace, l'institution est caractérisée par le rôle de producteur d'un savoir sur des archives d'images, tandis que l'utilisateur est caractérisé par le rôle du récepteur et par une volonté de savoir.

Europeana est influencée par le même type de rhétorique : le dispositif de lecture encadre l'expérience de l'*utilisateur-explorateur* en lui proposant surtout une exploration

par le moyen d'une compréhension verbale des images. Sauf que, dans ce cas-ci, la place des discours verbaux est même plus importante que dans la première tendance stratégique. En effet, la présence de discours verbaux ne se limite pas seulement à des textes de présentation du service ou des métadonnées descriptives. Comme nous l'avons suggéré, le point particulier de ce projet est l'augmentation considérable du *pôle médiatique* à travers des « blogs de vulgarisation » qui sont intégrés dans l'interface afin de mettre en évidence le potentiel des ressources agglomérées au sein de ce service Web. Ce changement stratégique vise à construire un autre point de vue, celui de l'*utilisateur-lecteur*.

Avant de conclure, nous pouvons mettre en évidence plusieurs caractéristiques des stratégies de communication en ligne qui sont conçues pour encadrer des bases d'images numériques. Nous en distinguons cinq.

(1) Comme nous l'avons remarqué, une base d'images numériques nécessite de concevoir un processus informatique qui gère la sélection des ressources numériques. L'exploration des bases d'images se fonde sur une structuration préalable des relations entre les ressources à partir des informations descriptives qui sont encodées dans un SGBD. Par conséquent, nous avons vu que les outils d'indexation des bases d'images se fondent systématiquement sur un métalangage de contrôle. L'avantage de cette conception de l'indexation est que les ressources numériques forment un ensemble structuré et cohérent d'informations, et surtout, que ces informations sont rassemblées dans un même format technique (les données numériques). Pourtant, l'inconvénient majeur est que l'utilisateur peut uniquement explorer les bases d'images des institutions à travers des moteurs de recherche en fonction de mots-clés. Cette manière de sélectionner et de classer des images est le résultat d'un paradigme d'interprétation des images qui est étroitement lié à une conception réductrice de l'analyse d'image que plusieurs travaux invitent explicitement à dépasser[424].

(2) Le fait de rendre publiques des archives ou des collections d'images à travers des bases d'images numériques contraint les institutions à produire un dispositif de lecture des ressources. En effet, la conception des bases d'images numériques est nécessairement

[424] **Chap. 3**.

complétée par la conception d'une interface Web permettant de manifester et de manipuler les ressources numériques (substituts d'images, métadonnées, textes de présentation du service). Ces ressources sont manifestées dans une page-écran fonctionnant comme une scène d'énonciation qui comprend diverses formes sémiotiques. Nous avons montré que l'énonciation des bases d'images comprend deux acteurs principaux : d'un côté, l'institution patrimoniale, en tant que productrice du service, des métadonnées descriptives et de tous les discours verbaux qui sont énoncés dans l'interface Web ; et de l'autre, les internautes, c'est-à-dire les utilisateurs de la base d'images qui, motivés par une « volonté de savoir », doivent construire leurs propres parcours d'exploration à travers le dispositif de lecture.

(3) Grâce à ce dispositif de lecture, les actions de l'utilisateur sont en interaction simulée avec les fonctionnalités du système informatique. En raison de la première caractéristique (cf. **supra**), ces actions se limitent dans la plupart des cas que nous avons présentés à un « input verbal » ou à des opérations de lecture. À ce propos, nous avons vu que ce dispositif de lecture favorise souvent une scénarisation standardisée à travers une mise en scène de l'espace de la page-écran, où il s'agit toujours de distinguer, d'une part, des scènes d'énonciation locales (textes, liens hypertextes, icônes, images, etc. autrement dit, les parties statiques et invariantes de la page) et, d'autre part, une scène d'énonciation globale (l'énonciation de l'interface comprise comme une « médiation ouverte »). Dans les différents services que nous avons présentés, même si cette énonciation globale mélange constamment une composante discursive avec une composante multimédia, la raison graphique est dominante dans cette énonciation.

(4) Concernant les usages des ressources numériques, nous avons vu que la scénarisation des bases d'images numériques oscille entre deux fonctions assignées au dispositif de lecture.

D'un côté, les services Web que nous avons présentés fonctionnent surtout comme des *inventaires* qui permettent d'accéder rapidement à des reproductions numériques d'archives d'images (*pôle référentiel*). Pour ce faire, l'interface est couplée à un moteur de recherche qui s'appuie sur un SGBD. Les possibilités d'exploration des bases d'images numériques sont ainsi réalisées à l'aide d'un instrument utile, quoique rudimentaire, qui sélectionne les images en fonction de mots-clés. Dans chaque cas, il s'agit de présenter les images sélectionnées par l'ordinateur dans une même zone dynamique de la mise en page. Cette zone peut prendre différents aspects, mais les résultats de la recherche sont

généralement présentés dans l'interface à travers des listes d'images. En outre, les résultats de la recherche peuvent être affinés à l'aide de filtres (liste de termes et de notions prédéfinies en fonction des métadonnées).

D'autre part, les services Web des institutions patrimoniales fonctionnent également comme une *vitrine* des activités de l'institution (*pôle médiatique*). Dans ce cas, les services préparent des parcours de lecture qui engagent clairement l'internaute dans un environnement qu'il s'agit de lire. En effet, les stratégies de communication en ligne des institutions patrimoniales se fondent en grande partie sur un principe de lisibilité. Ce principe de lisibilité de l'interface est étroitement conditionné, d'un côté, par la place accordée au métalangage de description dans le traitement et l'analyse des images, et de l'autre, par l'encadrement de la communication en ligne au sein d'une communauté discursive déterminée.

(5) Le **sixième chapitre** montre que lorsque les institutions patrimoniales ont tenté de renouveler cette deuxième fonction des services Web, ces institutions se sont d'abord orientées vers une stratégie que l'expression « ère de l'accumulation » de Gervereau caractérise assez justement. En effet, l'un des objectifs de ces services est aussi d'agglomérer et de standardiser une grande quantité d'images et de ressources documentaires sur ces images à l'aide de l'informatique. Ce renouvellement ne modifie pas l'idée centrale des bases d'images, à savoir que l'utilisateur est immergé dans un espace graphique et multimédia où la scénarisation de l'information est complètement contrôlée par une ou plusieurs institutions ; encore une fois, il s'agit de mettre en évidence que cette institution est à l'origine de la conception du service Web et surtout que l'institution est le producteur de ce service, et, en tant que garante de l'intelligibilité des ressources, l'énonciateur d'un savoir.

Par rapport aux stratégies liées à des institutions patrimoniales, nous avons constaté quelques différences, notamment en ce qui concerne le type d'images qui constitue les ressources ou la communauté qui encadre les images et les ressources sur ces images.

Dans un cas, la stratégie se fonde sur la *participation des utilisateurs* et sur le principe du *libre partage* des données, notamment en favorisant le modèle de l'encyclopédie sur Internet. Nous avons que ce sont les utilisateurs qui fournissent les ressources qui constituent la base d'images numériques, qui est toutefois organisée en

fonction de catégories et de thématiques générales et qui est surveillée par une communauté d'administrateur.

Dans l'autre cas, il s'agit de défendre l'idée d'une *coopération entre les institutions*, c'est-à-dire que la conversion numérique des ressources patrimoniales est l'occasion d'un effort de coopération entre institutions pour agglomérer et standardiser ces ressources, mais surtout pour défendre un patrimoine culturel à travers une identité européenne. Dans ce second cas, nous avons également remarqué la mise en place de discours de vulgarisation concernant des images. Par exemple, en intégrant des parcours de lecture qui se basent sur des textes écrits qui concernent des ressources. D'une manière ou d'une autre, ces discours sont produits par divers auteurs, mais ils s'intègrent plus généralement dans le discours du projet et des institutions qui participent au projet.

Ainsi, des projets comme *WikimediaCommons*, *WikiArt* ou *Europeana* mettent en évidence la même problématique que nous avons remarquée à propos des stratégies des institutions patrimoniales. Même si certaines caractéristiques des interfaces Web sont renouvelées (type et provenance des ressources ; importance du réseau hypertextuel dans la page-écran ; formes de présentation des résultats ; intégration de discours de vulgarisation), ces trois services Web favorisent surtout deux points de vue : celui de l'*utilisateur-explorateur*[425] et celui de l'*utilisateur-lecteur*.

6.3. Conclusions

Plusieurs institutions et projets envisagent donc la communication en ligne d'images comme la programmation d'un espace (l'interface) que l'utilisateur est amené à lire. Les services Web que nous avons présentés jusqu'à présent sont uniquement conçus comme un « inventaire » ou comme une « vitrine » des institutions ou du projet qui en sont à l'origine. Dans tous les cas, le producteur du dispositif de communication est le détenteur d'un *savoir* qui est transmis par le moyen du service, tandis que l'internaute (le destinataire) est caractérisé par une *volonté de savoir*. Pourtant, ce savoir n'est jamais disputé, il est présenté comme un « savoir brut » qui ne peut pas être polémiqué ; en effet, l'utilisateur a finalement très peu de prise sur ce savoir, puisque l'institution traite

[425] **Chap. 5**.

préalablement les significations des images et qu'elle est par là même la seule garante de l'intelligibilité des ressources qui consignent ce savoir.

En outre, même s'il s'agit d'archives et de collections d'images, ce savoir est en grande partie manifesté à travers des discours verbaux et des informations descriptives. L'expérience de ces services Web est ainsi complètement encadrée par un *style de visibilité* des ressources qui contraint à comprendre l'environnement informatique comme un *espace lisible*. Pour ce faire, chaque page-écran est constituée et organisée à l'aide de formes sémiotiques qui sont principalement héritées de l'écriture et de la lecture, cela même si le numérique se base sur un autre type de réflexivité[426] et qu'il offre certainement d'autres possibilités de traitement et de présentation d'un corpus d'images[427].

La première tendance stratégique et le premier renouvellement des stratégies de communication en ligne d'images débouchent par conséquent sur des dispositifs de communication en ligne qu'il est difficile d'envisager en dehors d'un encadrement verbal et d'une distribution cloisonnée des rôles énonciatifs et des parcours de lecture. Du point de vue de l'utilisateur, ces parcours prennent alternativement la forme d'un « parcours global » (qui le fait naviguer d'une page à une autre) ou d'un « parcours local » (au sein d'une page). En suivant les instructions d'usage mises en place par l'institution patrimoniale, l'utilisateur doit exploiter au mieux les seules fonctionnalités affichées dans l'interface qui est manifestée par la page-écran. Dans cette perspective, les bases d'images numériques sont des stocks d'images très utiles, mais qu'il faut nécessairement explorer à travers le prisme du langage verbal.

La différence entre la première tendance stratégique et ce premier renouvellement stratégique n'est donc pas forcément apparente. Pourtant, nous avons vu qu'un glissement dans la conception des interfaces Web est observable. En effet, la base d'images numériques n'est pas seulement conçue comme un simple inventaire des ressources ou comme une vitrine des activités de l'institution. Ce renouvellement transforme la page-écran en un espace rhétorique qui mobilise le numérique pour produire une expérience de lecture plus « dense ».

En effet, même si la stratégie d'Europeana est surtout d'agglomérer des archives d'images provenant de différentes institutions et fournir un instrument pour rechercher

[426] **Chap. 1**.

[427] Cf. **chap. 3 (sect. 3.1.4. & 3.2.4) & chap. 4**.

des ressources (pôle référentiel), ce projet veut aussi proposer à l'utilisateur d'autres formes d'expression pour encadrer la compréhension des ressources au sein d'une communauté discursive ; notamment en intégrant des discours de vulgarisation et des blogs pour valoriser les ressources. Ces formes d'expression ont avant tout une visée heuristique, car le but est de faire découvrir des ressources dans un cadre limité, à savoir celui d'une compréhension des images à travers une activité réflexive qui se fonde toujours sur un traitement documentaire et des parcours de lecture.

Pourtant, malgré cette volonté de renouvellement des interfaces Web à travers la coopération, nous pouvons faire ressortir les mêmes limites et contraintes que les stratégies déployées par les institutions locales. D'une part, l'utilisation du traitement informatique se limite dans la plupart des cas à des instruments de recherches par mots-clés ou des réductions de résultats en fonction des métadonnées descriptives[428]. D'autre part, à cette échelle, l'agglomération et la standardisation de métadonnées contraignent à sacrifier la précision des descriptions, ce qui risque à long terme d'augmenter le problème de la pertinence des images sélectionnées par l'ordinateur.

Dans le **chapitre suivant**, nous continuerons de considérer cette deuxième tendance stratégique afin de remarquer certains changements importants dans l'utilisation du support numérique pour concevoir des dispositifs de lecture. Selon nous, ces changements sont à l'origine du déplacement vers la troisième tendance qui concerne les stratégies actuelles et que nous considèrerons dans la **quatrième partie**. En effet, si le renouvellement a d'abord favorisé la conception de services qui s'apparentent à des « médiathèques en ligne » ou à des blogs qui peuvent contenir des « discours de vulgarisation », d'autres projets ont progressivement pris acte des caractéristiques et des possibilités spécifiques du numérique. Nous verrons que, plus récemment, certaines institutions patrimoniales investissent davantage les dimensions visuelle et interactive des interfaces Web, notamment en proposant des « expositions virtuelles »[429].

[428] **Chap. 3, 5 & 6**.

[429] À ce propos, voir aussi les discussions sur le musée virtuel dans le **chap. 2**, **sect. 2.2.**

CHAPITRE 7
Renouvellement des bases d'images numériques (2) : stratégies de l'interactivité

Dans les **chapitres précédents**, nous avons analysé plusieurs services Web dont le but est d'aménager un dispositif de communication en ligne. Les services que nous avons regroupés dans la première tendance stratégique reposent à peu près tous sur une même stratégie[430]. En effet, les bases d'images numériques des institutions patrimoniales sont conçues autour de la recherche par mots-clés et de filtres linguistiques, ou encore en fonction de thématiques générales. Les institutions favorisent une stratégie de communication où les interfaces Web fonctionnent comme une réserve de ressources structurée à partir d'un langage de contrôle des images et du visuel. En présentant quatre exemples de services Web, nous avons constaté deux fonctions principales du dispositif de lecture : l'interface Web fonctionne à la fois comme une *vitrine* de l'institution et comme un *catalogue* des ressources numériques disponibles. Comme nous l'avons vu, la communication en ligne d'images dans le contexte des pratiques d'archivage s'articule sur deux pôles de communication : celui du stockage et de la consignation de documents (pôle référentiel) et celui de la diffusion et la valorisation des activités de l'institution à l'aide de l'informatique (pôle médiatique).

[430] **Chap. 5**.

D'un autre côté, le contexte des humanités numériques encourage certainement les institutions à collaborer davantage avec divers spécialistes en informatique (transdisciplinarité) dans le but de renouveler les services Web qui rendent accessibles des archives et des collections d'images. Ce renouvellement s'appuie notamment sur des stratégies de coopération entre les institutions patrimoniales ou encore la participation active des utilisateurs dans la constitution d'une médiathèque où les ressources sont organisées en fonction d'une hiérarchie de catégories et de sous-catégories.

Jusqu'à présent, nous avons vu que ces stratégies de communication favorisent des conditions similaires de visibilité (une même économie) des images en permettant en effet à l'utilisateur de voir et de manipuler ces images dans un espace multimédia où la rationalité graphique et discursive est dominante.

Dans ce **septième chapitre**, nous montrerons que les conditions de visibilité des images se sont davantage renouvelées, car certaines institutions se sont repositionnées par rapport aux progrès plus récents dans le domaine de la communication en ligne. L'objectif de ces institutions est d'amorcer le passage des bases d'images, telles que nous les avons présentées, à des *musées virtuels d'images*[431]. Ce deuxième temps dans le renouvellement stratégique de la communication en ligne des larges ensembles d'images transforme progressivement les modalités d'énonciation du dispositif de lecture en ajoutant des possibilités d'interaction dans l'espace de la page-écran. Cette transformation prépare d'autres types de parcours pour les internautes, notamment à l'aide des *blogs de vulgarisation*[432], mais encore à travers des *expositions virtuelles*.

7.1. Des bases d'images numériques aux expositions virtuelles

L'exposition virtuelle est un « mode de visite » qui s'inspire de celui du musée. En tant que visiteur, l'utilisateur est « accompagné, guidé et informé » à travers une scénarisation qui est construite sur le principe de la page-écran (Després-Lonnet, 2009 : 26) ; le mode de construction de ces pages est, comme le suggère Marie Després-Lonnet (*ibid.* : 29), une « nouvelle forme d'écriture et de mise en textes [qui] est basée sur le substrat

[431] Voir Deloche (2001). Rappelons que, par rapport au musée traditionnel, le but des musées virtuels est de monter (de rendre « visibles ») des substituts qui ressemblent aux supports originaux ; cf. **chap. 2, sect. 2.2.**

[432] Voir aussi **chap. 6, sect. 6.2.3.**

technique et textuel, la préexistence d'un texte, structuré au sein d'une base de données, dont certains extraits seront choisis en fonction de proximités terminologiques dans les contenus de certains champs ».

> Les résultats des requêtes sont insérés dans des gabarits préformatés, pour être affichés à l'écran. Il y a là une forme d'industrialisation de la production des pages qui mène à un affichage indifférencié, quelles que soient les particularités des éléments qui les composent. Si ce mode d'affichage peut répondre aux attentes d'un spécialiste qui pense consulter un registre, il semble en revanche peu cohérent avec le projet de construire des parcours de découverte (Després-Lonnet, 2009 : 29).

L'auteur ajoute qu'il « ne faut cependant pas négliger les effets de sens produits par l'affichage, au sein d'un même espace graphique, de reproductions d'œuvres que l'idée même de parcours suggère implicitement d'associer » (*ibid.* : 30). L'espace graphique de la page manifeste une ou plusieurs mises en scène impliquant différents « co-énonciateurs » et invitant ceux-ci à participer à une pratique de production et de communication de sens au sein d'un ensemble d'activités sociales et culturelles influencé par des manières de faire qui sont fixées par les pratiques d'usage (voir aussi Pignier & Drouillat, 2008).

Dans la construction des expositions virtuelles, certaines modalités de production du texte et la manière de composer des ensembles signifiants évoquent sans doute les modalités des expositions *in situ* :

> L'imaginaire de la visite d'exposition est sollicité lorsque le lecteur/visiteur suit l'un des parcours, qu'il répond à l'invitation de découvrir un peintre ou une époque, ou qu'on lui propose de mettre en lien un ensemble d'œuvres sur la base d'une thématique ou d'une technique que l'on aura préalablement pris soin de commenter (Després-Lonnet, 2009 : 30)[433].

L'utilisateur consulte alors « des reproductions d'objets estampilles "de musée" et à des commentaires qui peuvent rappeler les panneaux qui accompagnent les œuvres dans les salles d'exposition » ; Després-Lonnet souligne en outre la « possibilité donnée d'agrandir les images suggère également un mouvement vers les objets, dont on a le sentiment de se rapprocher pour en découvrir les détails » (*ibid.* : 30). Dans cette

[433] Sur les relations entre expositions et les espaces multimédias interactifs, voir aussi Davallon (1999) ou encore Château & Darras (2000).

perspective, l'auteure (*ibid.* : 31) fait ressortir trois « modèles » qui influencent la page-écran : (1) le modèle de la conservation du patrimoine qui aboutit sur « la production normée de textes catalographiques » que l'auteur nomme *textes premiers* ; (2) le modèle de l'ingénierie informatique qui fournit « les cadres formels et graphiques propres à porter ces écrits à l'écran, dans la logique du projet documentaire, mais également en fonction de modèles (et de modules) techniques sous-jacents » ; (3) le modèle de la médiation culturelle qui réécrit (en partie) la logique du projet documentaire « pour tenter de présenter le catalogue à travers divers parcours » (*textes seconds*) qui font référence à l'exposition. Les internautes consultent ainsi « un gigantesque hyperdocument » composé de notices descriptives et de reproductions des milliers de ressources documentaires.

> [Ce] n'est pas à une exposition (même virtuelle) dans le sens muséal du terme [que l'utilisateur] est convié, mais plutôt à la découverte du contenu des dossiers documentaires et du gigantesque travail d'érudition nécessaire à la constitution d'une base destinée à collecter l'ensemble des données d'inventaire du patrimoine conservé […]. Il n'est plus question de mettre en scène quelques œuvres soigneusement sélectionnées et organisées selon une thématique particulière, mais d'envisager l'entièreté des collections des musées […] et d'ordonner les documents qui les décrivent à partir d'un langage de description normé et selon une stricte logique de classement, nécessaire dans toute opération d'inventaire de cette envergure (Després-Lonnet, 2009 : 32).

La communication permet de rapporter cet « hyperdocument » à un dispositif de lecture, un espace multimédia permettant des lectures locales d'un ensemble global d'objets et d'informations. Comme le remarque justement l'auteure, cette constatation permet de comprendre « l'intérêt, mais aussi la redoutable complexité d'un projet qui viserait à concevoir un dispositif propre à rendre intellectuellement appréhendable un tel objet. Ce dispositif ne devrait effectivement pas se contenter de mimer à l'écran les pratiques de médiation culturelle en cours dans les espaces physiques du musée ou de mobiliser les outils et pratiques d'écriture des conservateurs chargés de l'inventaire » (*ibid.* : 32). L'auteure renvoie alors aux propositions de Bruno Bachimont qui, comme nous l'avons vu[434], propose de prendre acte du passage de la « raison graphique » à la « raison computationnelle », en soulignant que ce passage met en lumière de « nouvelles

[434] **Chap. 1, sect. 1.4.**

formations du sens » qui, même si elles fondées sur le calcul et un automatisme aveugle, devraient permettre à l'internaute d'inscrire son propre parcours dans les interfaces Web.

Nous avons principalement analysé des services Web qui sont développés dans le cadre d'institutions patrimoniales[435]. Ces services visent principalement trois objectifs : (1) structurer et organiser un ensemble d'images numériques dans une base d'images numériques ; (2) concevoir une interface-utilisateur permettant la recherche par mots-clés ; (3) mettre en valeur ces images et les missions ou encore les activités des institutions à l'aide de discours verbaux manifestés par des textes. Dans la même perspective, nous pouvons évoquer trois autres services Web mis en place par la Bibliothèque nationale de France (BnF)[436], la bibliothèque de l'Institut National de l'Histoire de l'Art (INHA)[437] ou encore Paris Musées[438]. Même si nous pouvons observer la persistance, dans ces trois services, de la même double fonction du dispositif de lecture (inventaire/vitrine), les paragraphes précédents indiquent qu'une métamorphose des conditions de visibilité et de manipulations des larges ensembles d'images est en cours et qu'elle modifie la communication en ligne proposée par ces institutions.

Le service Web des collections de Paris Musées (fig. 56 & 57) s'ouvre sur une page d'accueil où nous retrouvons les mêmes caractéristiques que nous avons observées dans les bases d'images numériques des institutions patrimoniales (champ de recherche en fonction de mots-clés et de filtres ; menu renvoyant vers différentes fonctions du site ; rubriques informatiques ; expositions virtuelles ; illustrations ; hyperliens, diagramme). En outre, un diagramme représente la répartition des œuvres soit en fonction du type d'objet (ou de support) soit en fonction du nombre d'œuvres par musées. Cet outil révèle que plus de la moitié provient d'un seul musée (le musée Carnavalet – Histoire de Paris) et que la plupart des objets concernent des estampes, des dessins et des photographies.

Paris Musées propose un instrument pour rechercher des ressources, ou encore 17 *parcours thématiques* et 4 *expositions virtuelles*.

---

[435] Cf. **chap. 5 & 6**.

[436] https://www.bnf.fr/fr/banque-dimages et http://expositions.bnf.fr.

[437] https://bibliotheque-numerique.inha.fr/collection-toutes.

[438] https://www.parismuseescollections.paris.fr/fr.

Les parcours thématiques sont en relation avec des activités institutionnelles (par exemple le Festival de l'Histoire de l'Art à Fontainebleau ; l'exposition « Éléments : air, terre, feu, eau » à Montreuil) ou avec des sujets particuliers (histoire de Paris ; littérature ; peinture ; œuvre d'un artiste ; trésors de bibliothèques). La spécificité de ces parcours est qu'ils offrent des points de vue ciblés sur des corpus réduits d'œuvres (de 11 à 89 œuvres), ce qui permet d'éviter une recherche dans presque 400 000 œuvres réparties dans divers musées de la ville de Paris.

Les thématiques des expositions sont les suivantes : (1) « L'esprit Art nouveau » ; (2) le peintre « Alexandre Nozal (1852-1929) » ; (3) « Notre-Dame de Paris en plus de 100 œuvres » ; (4) « L'ombre et la lumière ». Par exemple, l'une de ces expositions virtuelles concerne les représentations de la cathédrale Notre-Dame de Paris dans 126 œuvres qui sont réparties dans cinq musées (fig. 58 & 59). L'exposition est composée de « plusieurs étapes » qui révèlent la structure de la scénarisation qui est prévue par l'institution (à peu près comme les chapitres dans les *stories* d'Europeana ; cf. fig. 60). Chaque étape est accompagnée d'un discours de vulgarisation.

Dans le même sens, le service Web de l'INHA propose des « expositions virtuelles » (fig. 61-62) qui ont pour but de « proposer un focus sur certains documents » ou de mettre en évidence les nouvelles numérisations de l'institution. Les expositions prennent encore une fois la forme d'un discours verbal avec de nombreux hyperliens et des illustrations tirées des ressources de l'institution.

En parallèle, le service offre un instrument de recherche qui permet à l'utilisateur d'explorer les ressources en fonction de différentes catégories qui correspondent au type de ressources (manuscrits, dessins, photographies, etc. ; cf. fig. 63) ; la quantité de ressources disponibles est inscrite en vis-à-vis de chaque catégorie. De plus, une rubrique « actualités » rassemble plusieurs liens hypertextes qui renvoient vers des articles concernant les ressources et les activités de l'institution.

Comme dans les deux exemples précédents, la banque d'image de la BnF propose, d'une part, un outil de recherche en fonction de mots-clés ou de différentes catégories thématiques (fig. 64), et d'autre part des « expositions virtuelles » (fig. 65). Même si les expositions proposent différents types de parcours de lecture, d'une manière générale, il s'agit toujours d'un parcours de lecture qui est divisé en plusieurs étapes (cf. fig. 68 &

70) ; la présence d'images a surtout une fonction illustrative (cf. fig. 66 & 67 & 69), voire explicative lorsqu'il s'agit par exemple d'une capsule vidéo (cf. fig. 71).

La deuxième tendance stratégique est caractérisée par la place plus importante accordée à la *vulgarisation* dans la communication en ligne, afin de « faire connaître » et de « faire voir » via Internet. Ces trois exemples indiquent la volonté affirmée des institutions patrimoniales de maintenir une visée rhétorique qui construit une relation d'énonciation à partir d'opérations de lecture. L'espace de la page-écran est toujours un espace écrit et dont le but est d'être *lisible* parce qu'il est destiné à un *utilisateur-lecteur*.

Comme nous l'avons déjà constaté, les stratégies institutionnelles sont tenues de s'accommoder de changements récurrents dans la « médiasphère informatique » (ou « hypersphère »)[439]. Dans le **chapitre précédent**, nous avons déjà présenté un projet qui insiste sur la participation des internautes pour constituer une médiathèque en ligne. Dans la continuité de cet exemple, nous verrons que la volonté de renouvellement des stratégies est observable à travers d'autres projets qui favorisent surtout de nouvelles possibilités de communication en ligne par le moyen de l'*interactivité* et de la *visualisation de médias*[440]. Les sections suivantes permettront auparavant de remarquer que, progressivement, les institutions patrimoniales accordent plus d'importance à la place du visuel et à la multiplication des points de vue dans la conception de leurs services.

[439] **Chap. 2, sect. 2.2.2.** Voir aussi les discussions sur WikimediaCommons dans le **chap. 6, sect. 6.1.**

[440] Dans la suite du présent travail, nous analyserons d'autres projets qui se fondent sur les progrès plus récents en informatique pour proposer des solutions innovantes aux institutions pour réagir à ces changements ; cf. **quatrième partie**.

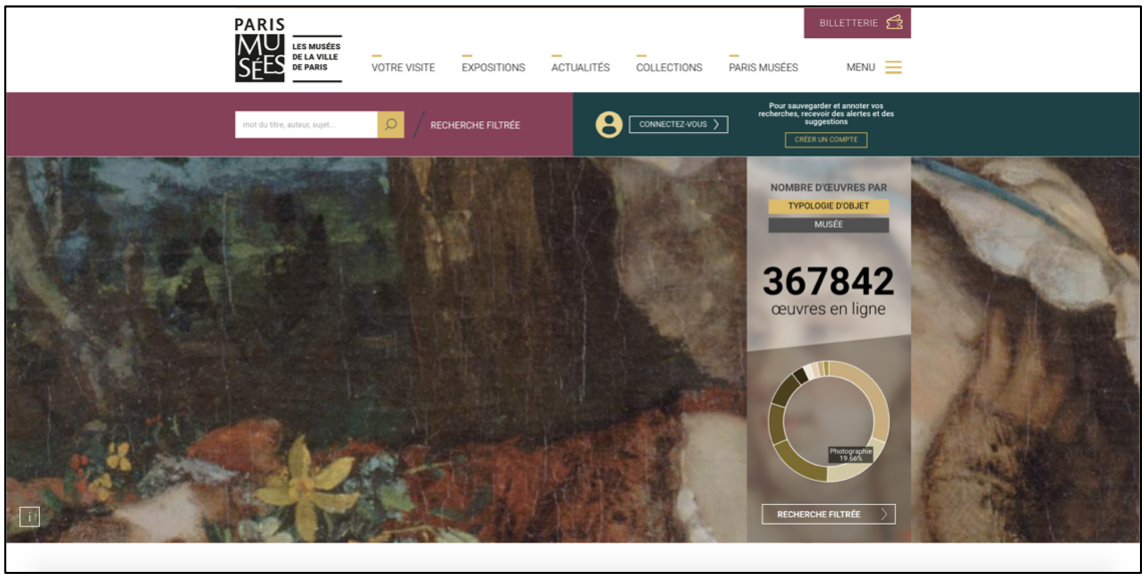


Figure 56 : Page d'accueil de *Paris Musées* ; remarquons l'intégration, à droite, d'un diagramme qui représente la répartition quantitative des ressources disponibles en ligne.

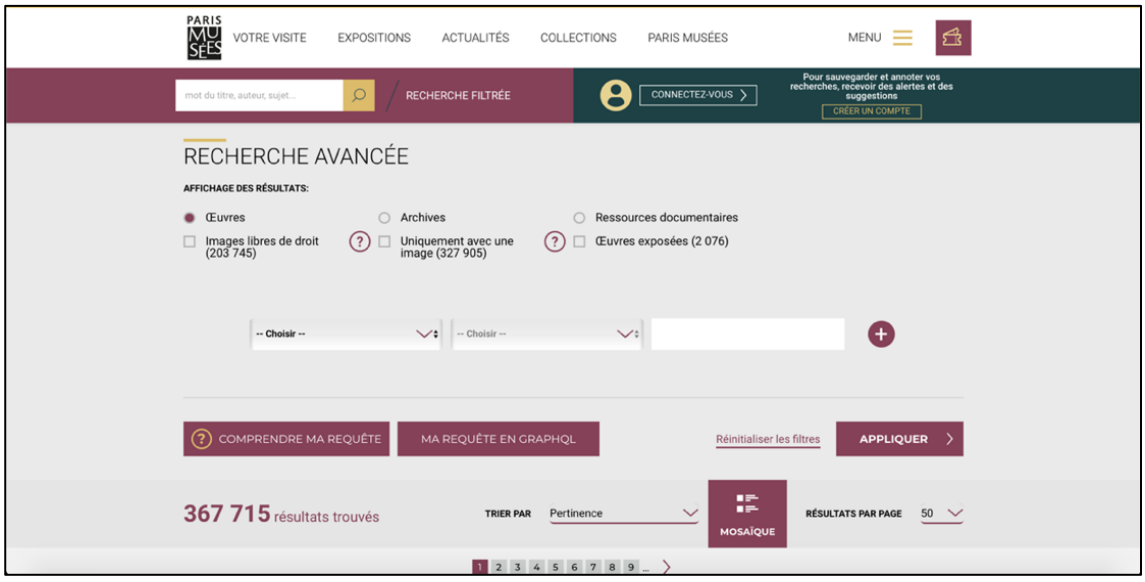


Figure 57 : Moteur de recherche avancée de *Paris Musées*.

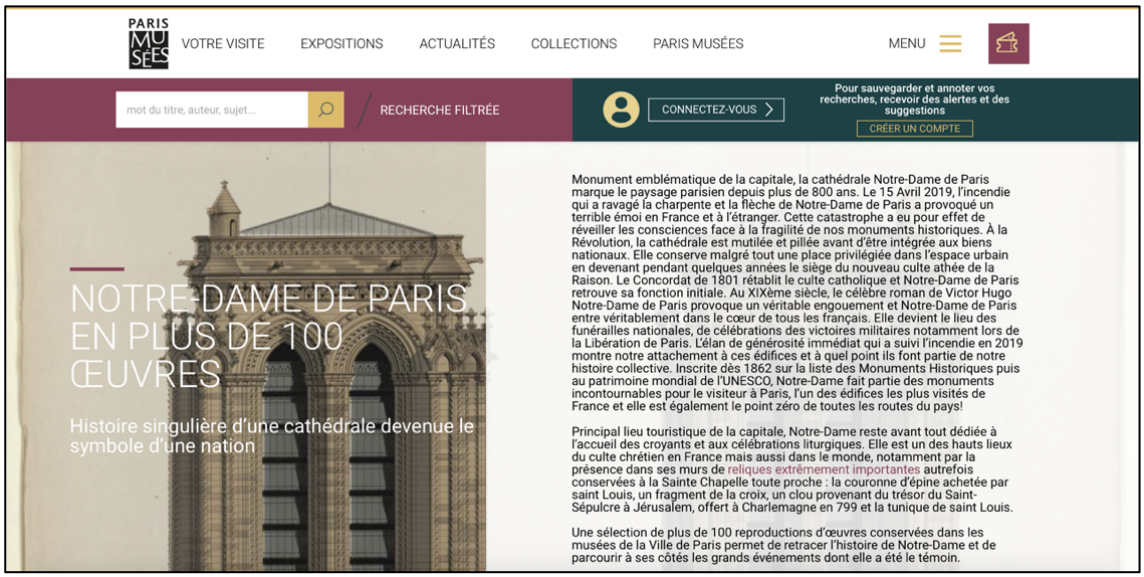


Figure 58 : Exemple d'une exposition virtuelle proposé sur le site *Paris musées.* À droite, le texte de présentation de l'exposition.



Figure 59 : Même exemple de parcours thématique proposé sur le site *Paris musées.* Cette partie reprend les musées de Paris concernés par cette exposition, quelques mots-clés qui caractérisent le sujet, la période, le thème, le domaine et les personnalités relatives à cette exposition.

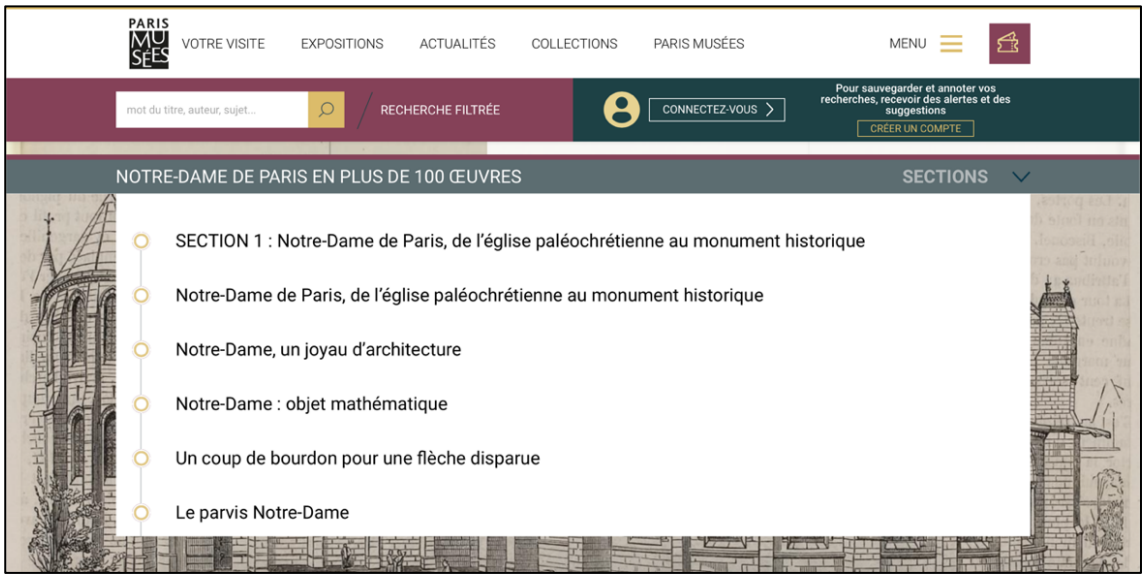


Figure 60 : L'exposition virtuelle (*Paris musées*) est composée de plusieurs étapes (scénarisation).

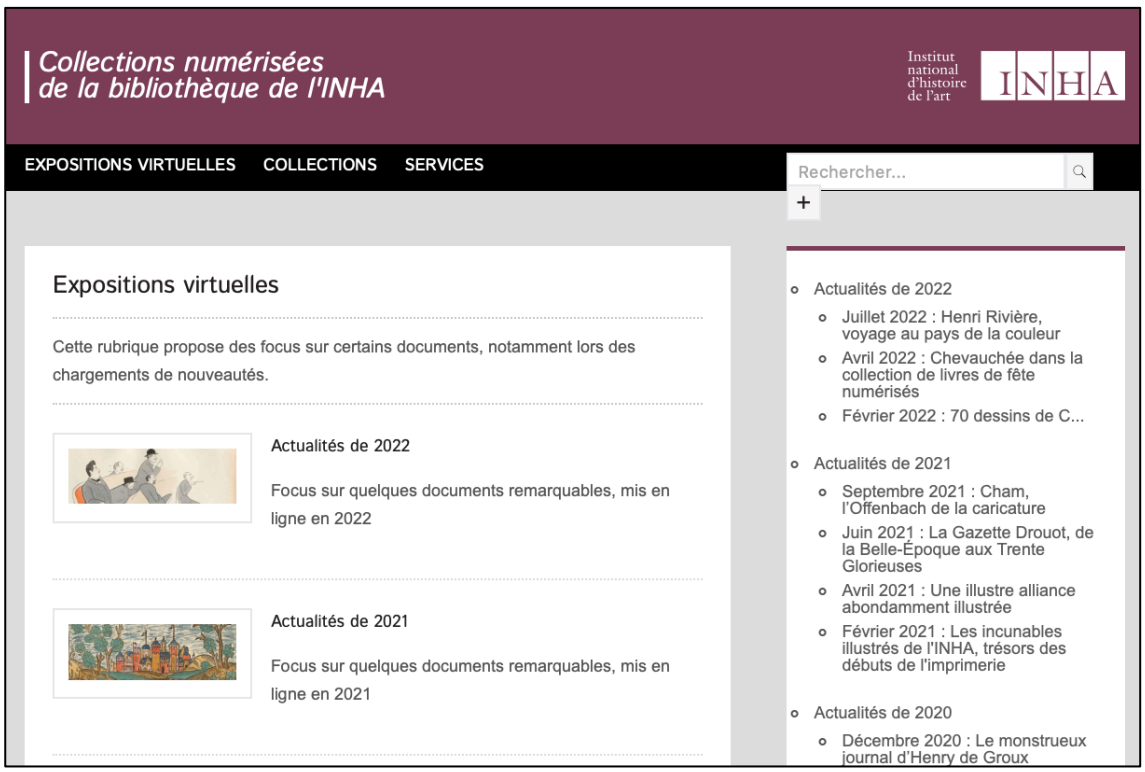


Figure 61 : Page d'accueil des collections numérisées de la bibliothèque de l'*Institut national d'histoire de l'art*.

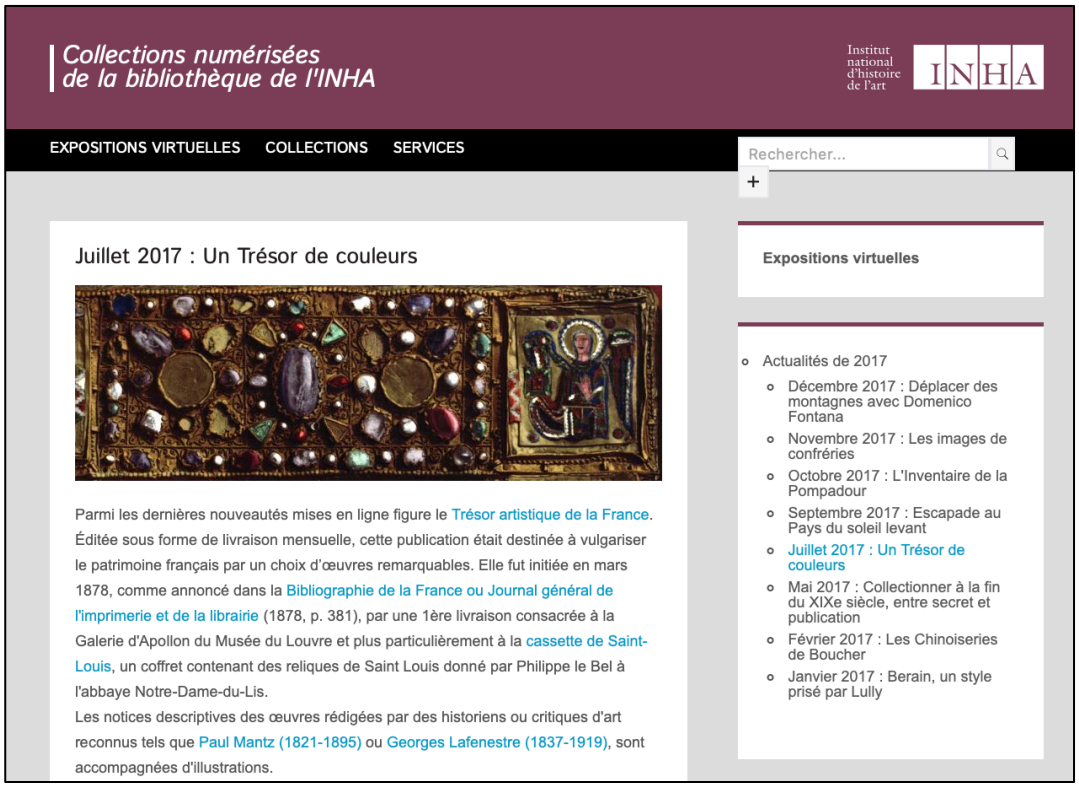


Figure 62 : Exemple d'une exposition virtuelle (*Institut national d'histoire de l'art*).

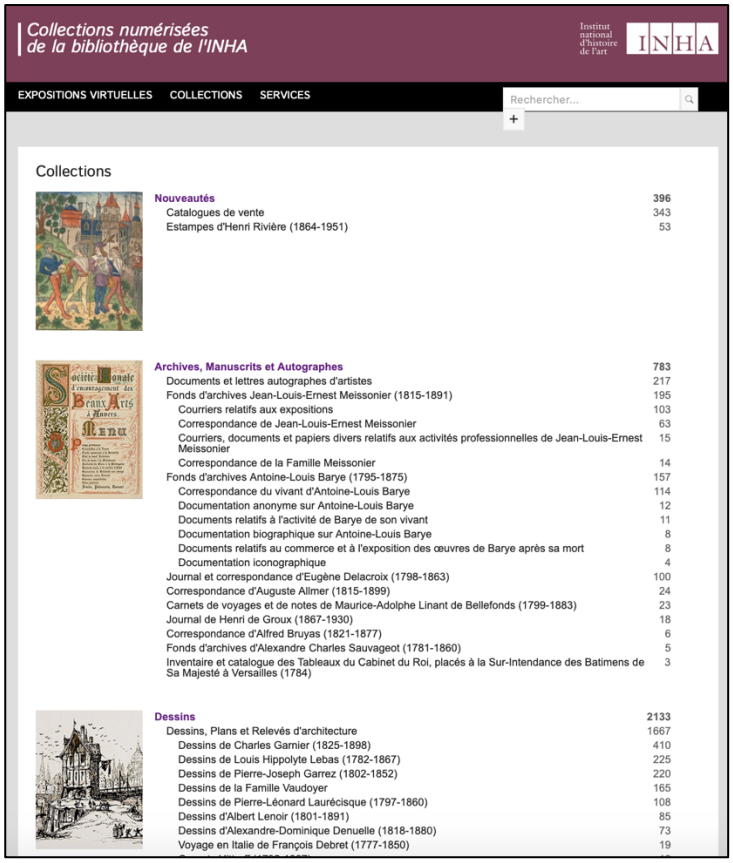


Figure 63 : Page de présentation des collections en ligne de l'*Institut national d'histoire de l'art*.

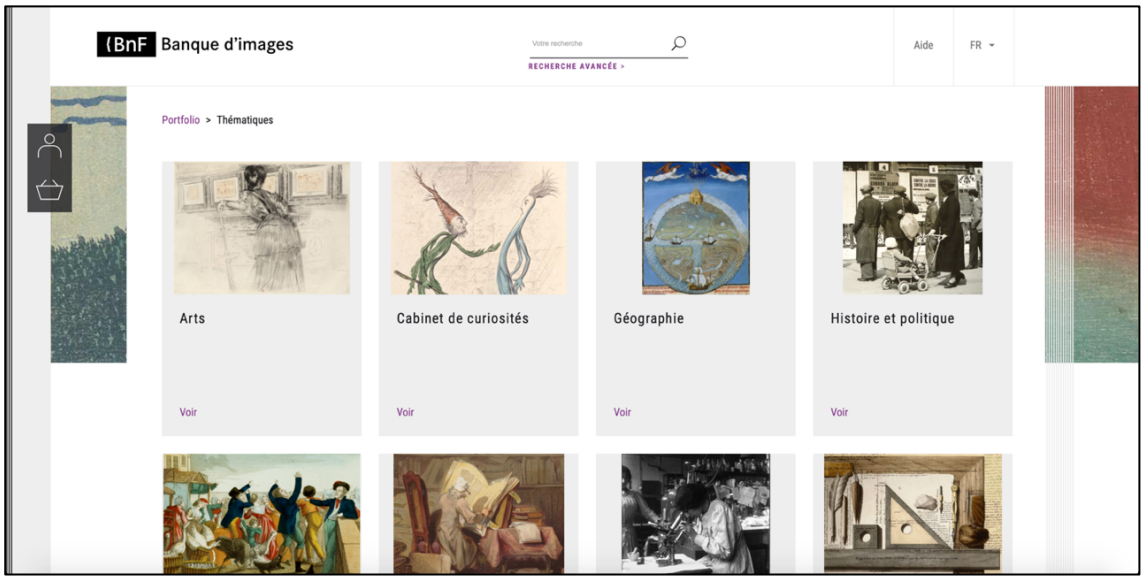


Figure 64 : Banque d'images de la *Bibliothèque nationale de France.*



Figure 65 : Page d'accueil des expositions virtuelles proposées par le service Web de la *Bibliothèque nationale de France.*

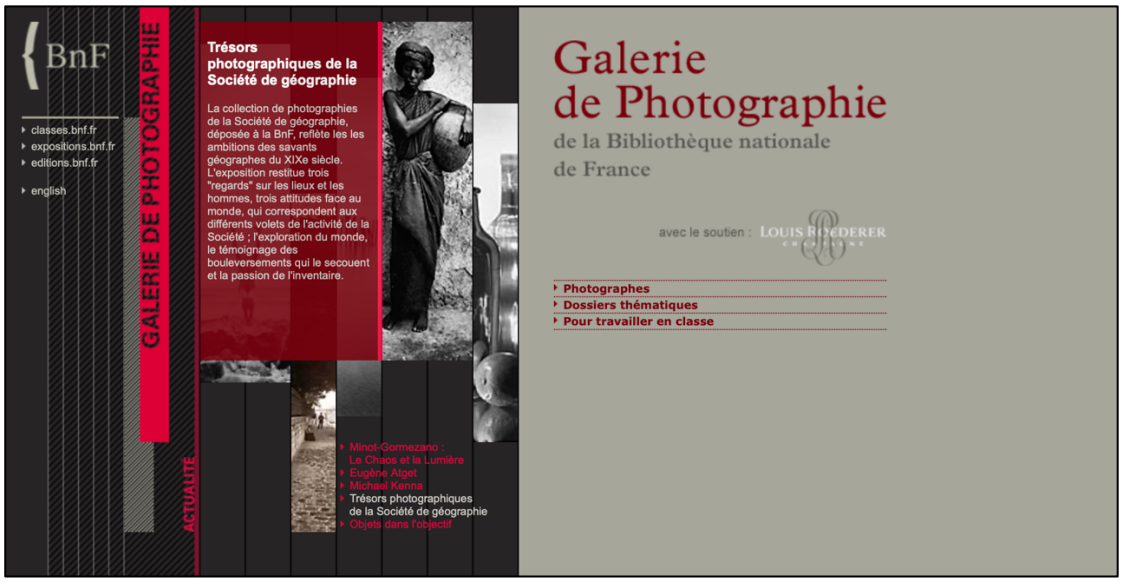


Figure 66 : Page d'accueil de la *Galerie de Photographie de la Bibliothèque nationale de France*.



Figure 67 : Exemple d'exposition virtuelle proposée par la *Bibliothèque nationale de France*.

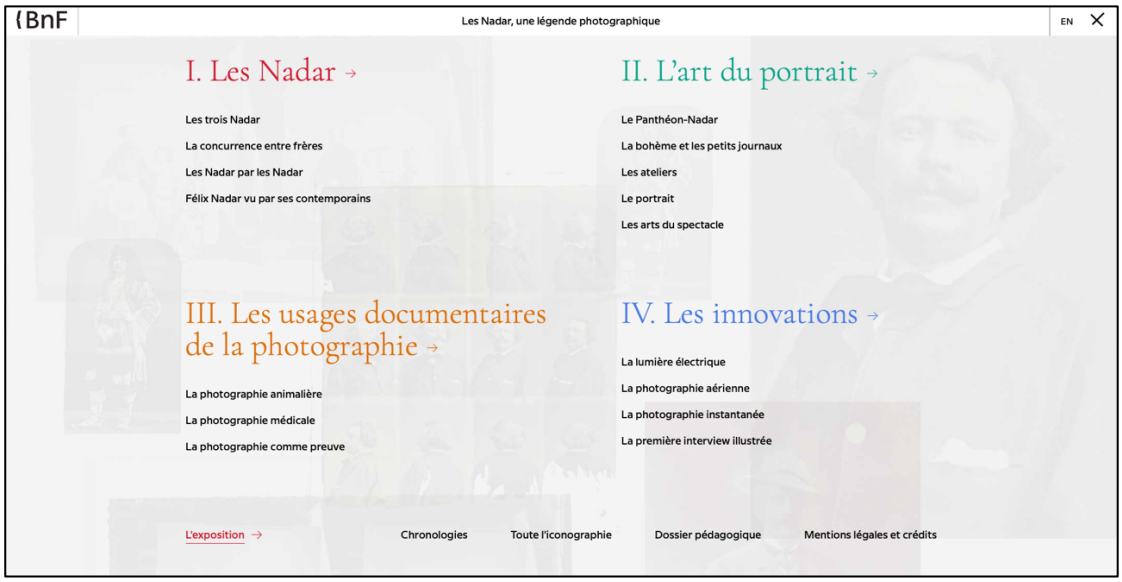


Figure 68 : *Idem* que figure 67.



Figure 69 : Autre exemple d'exposition virtuelle (*Paysages français*) de la *Bibliothèque nationale de France*.

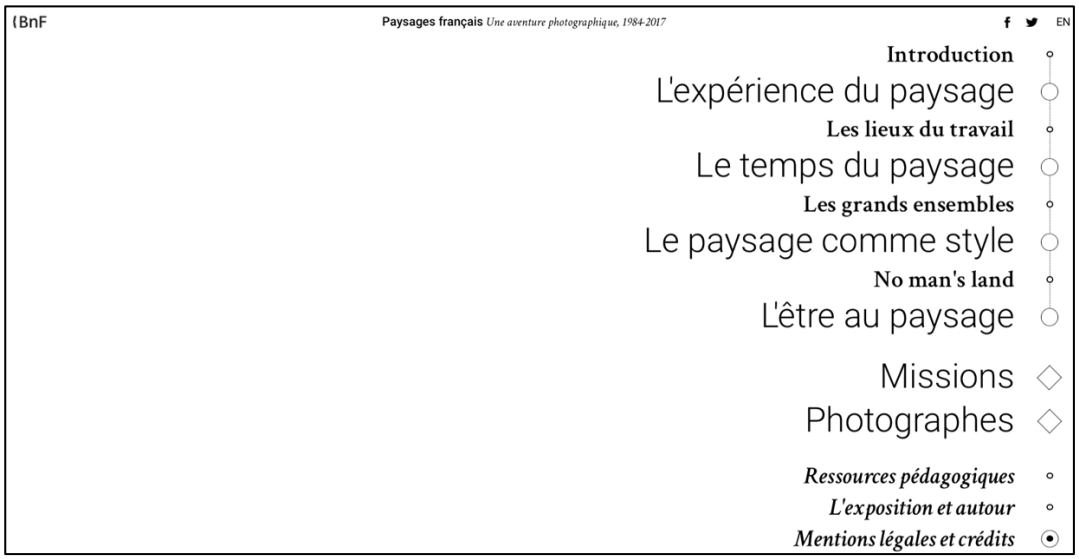


Figure 70 : *Idem* que la fig. 69.



Figure 71 : *Idem* que la fig. 69 ; dans ce cas, l'introduction est manifestée par le moyen d'une capsule vidéo.

7.2. Renouvellement d'un service Web à travers les applications interactives

Comme nous l'avons remarqué, la communication en ligne a d'abord été utilisée pour proposer un service de recherche dans un ensemble de ressources numériques (utilisateur-explorateur), puis pour agglomérer des ressources en vue de les intégrer dans des blogs ou des expositions virtuelles (utilisateur-lecteur).

La stratégie du projet[441] suivant a pour but d'améliorer l'ergonomie des interfaces en proposant plus de fonctionnalités qui se basent sur des applications. Nous verrons que ce projet accorde une place plus importante à la participation des utilisateurs dans le fonctionnement de l'interface. Cet exemple mobilise les possibilités de médiation de la page-écran en vue d'une rhétorique qui multiplie les points de vue sur des ressources. Nous constaterons que ce dispositif de lecture s'éloigne d'une conception qui est en grande partie basée sur une rhétorique verbale, au profit d'une rhétorique visuelle qui assume davantage la *plasticité de la page-écran*.

Comme nous l'avons déjà remarqué[442], cette plasticité sert de support perceptible pour un plan de la représentation, en même temps qu'elle en délimite le cadre. Au risque d'anticiper un peu, soulignons déjà que l'énonciation qui est en jeu dans la communication en ligne des images ne se résume pas seulement à des effets d'objectivité du discours (énonciation énoncée) ou à une distribution figée de rôles énonciatifs. Le processus d'énonciation s'étend aussi jusque dans cette plasticité de l'interface Web[443].

[441] Les informations des sections suivantes proviennent de Gorgles (2013).

[442] **Chap. 5, sect. 5.3.**

[443] Cf. **infra & chap. 8**.

7.2.1. La stratégie du *Rijksmuseum*

Dans le cadre de la rénovation du *Rijksmuseum* d'Amsterdam (de 2003 à 2013), la fermeture du bâtiment principal a permis de numériser une partie importante des collections artistiques du musée[444]. En parallèle, cette institution a décidé de renouveler sa présence sur Internet en suivant deux perspectives. D'une part, cette institution a conçu un nouveau site Web. D'autre part, une application participative (Rijksstudio) a été intégrée dans ce nouveau site ; cette application permet d'explorer des copies numériques des images du Rijksmuseum (à peu près 125 000 numérisations de haute qualité), mais en insistant davantage sur la spécificité du support et de l'interface Web qui est utilisé pour manipuler les collections.

Nous distinguons trois objectifs de nouveau site Web du Rijksmuseum :

- Une utilisation simple et intuitive pour tout type d'utilisateur (priorité principale).
- Attirer un public plus large que celui des experts et des professionnels, en particulier un public susceptible de partager les ressources de l'institution via les réseaux sociaux.
- L'objectif à long terme de l'institution est de devenir un acteur important dans le champ de la diffusion informatique des collections artistiques.

D'un autre côté, Rijksstudio est une application *open source* qui fonctionne sur plusieurs types d'appareils (ordinateur, smartphone, tablette). Le but est de « connecter » les internautes avec l'histoire de l'art en suivant cinq valeurs fondamentales : (1) la simplicité ; (2) la particularité ; (3) l'authenticité ; (4) la qualité ; (5) l'innovation. Pour mettre en connexion les collections numériques avec un large du public, ce renouvellement se base sur une stratégie de communication adaptée à l'ère du Web et du numérique (*e-strategy*) ; à savoir une stratégie qui a pour but de faciliter l'accès aux collections numériques du musée en distinguant, d'une part, le domaine et le contenu qui sont propres au musée traditionnel, et d'autre part, le contenu et les modalités de présence propres au musée virtuel sur Internet.

[444] Auparavant, le site Web du Rijksmuseum comprenait seulement une collection d'images en faible définition et une base de données d'information sur les images. Le but de l'ancien site Web était de rendre accessibles des reproductions numériques des œuvres du musée sur Internet, mais cette version du site proposait seulement un ensemble de données fragmentées et, du point de vue de l'institution, l'interface Web ne permettait pas une navigation simple et intuitive.

Comme nous l'avons suggéré, si le Rijksstudio a pour but d'attirer un large public (des néophytes ou professionnels), c'est parce que ce projet vise aussi à attirer le public des réseaux sociaux. En effet, des internautes échangent tous les jours de grandes quantités d'images numériques sur les plateformes et des réseaux sociaux (par exemple : *Instagram*, *Pinterest*, *Flickr*, etc.) dont le but n'est pas de valoriser le patrimoine. Ces réseaux impliquent une forte participation des utilisateurs qui produisent et échangent des images. Étant donné que ce type de pratiques numériques (voir notamment les contributions dans Ablali & Bertin, 2020) se sont accrues durant ces dernières années, l'idée est de considérer ce public comme un ensemble d'utilisateurs potentiels, car ceux-ci ont l'habitude de regarder et de partager des images numériques avec leurs amis ou leurs *followers*.

Nous avons vu que la conception d'un service Web de communication en ligne nécessite de s'appuyer sur un protocole et des architectures informatiques[445]. Afin de développer une stratégie qui prend en compte ce contexte, les développeurs du Rijksstudio se sont notamment inspirés de *Scrum*. Ce service est spécialisé dans le développement de produits informatiques complexes et qui propose un cadre de référence théorique et des protocoles standardisés à suivre afin de concevoir des applications Web[446]. L'équipe du Rijksmuseum a notamment collaboré avec une entreprise spécialisée dans la conception de logiciels (*Q42*)[447]. D'un autre côté, l'ensemble des données numériques sont stockées sur une plateforme qui héberge des applications logicielles (*Google App Engine*)[448]. Tandis que la collection d'images numériques est gérée grâce à une bibliothèque d'image (*Mediabin*)[449].

---

[445] **Chap. 2**.

[446] Voir https://www.scrum.org/about. Dans les grandes lignes, il s'agit de renouveler complètement le design du site Web du musée, mais aussi de proposer une interface en adéquation avec les autres pratiques actuellement utilisées dans la conception de sites Web. En suivant la logique de *Scrum*, l'équipe du musée a par conséquent collaboré avec des webdesigners et des informaticiens pour concevoir un site Web plus ergonomique pour explorer la collection d'images numériques du musée.

[447] https://www.q42.nl/en

[448] https://cloud.google.com/appengine/docs/

[449] Voir : https://www.opentext.com/products-and-solutions/products/customer-experience-management/digital-asset-management/opentext-mediabin. *OpenText MediaBin* est une solution de gestion de ressources numériques, qui propose des services qui permettent de gérer efficacement la distribution et la publication de milliers de ressources en relation avec des clients et des activités potentiels. *MediaBin* permet entre autres d'extraire automatiquement les concepts clés de contenus audiovisuels et de fichiers images, en analysant et en catégorisant les ressources selon étiquetage automatisé. Il reconnaît des traits

Si nous avons montré que les institutions patrimoniales produisent souvent des bases d'images ou des explorations virtuelles (le numérique est alors utilisé pour réaliser des opérations programmées, tandis que l'interface fonctionne comme un dispositif de lecture qui réduit l'expérience de l'utilisateur à des opérations de lecture), dans le cas du Rijksmuseum, les choix stratégiques s'approprient un peu plus les possibilités de visualisation et de manipulation qu'offre l'espace de la page-écran.

7.2.2. De la vulgarisation à l'observation : multiplication des points de vue dans l'interface

Nous pouvons d'emblée observer les résultats de ces choix stratégiques dans la page d'accueil du nouveau site Web du *Rijksmuseum* (fig. 72). Comme dans certains projets que nous avons présentés dans la section précédente des reproductions numériques des œuvres remplissent, dans la plupart des pages, quasiment l'espace entier de la page, tandis que les informations verbales sont moins nombreuses et concernent surtout des hyperliens pour naviguer sur le site Web. Cette nouvelle version du site Web propose toujours un service de recherche (*search the collection* ; cf. fig. 73) semblable aux autres bases d'images numériques que nous avons présentées ; l'utilisateur réalise une demande de recherche à l'aide de mots-clés, et en retour le système sélectionne des images en fonction de ces mots-clés et les présentent dans l'interface. Les résultats sont alors présentés sous la forme de mosaïque, et elles sont réparties en fonction de trois catégories principales : *works of art*, *sets of works*, *stories*. Ces catégories visent à bien différencier les reproductions numériques des œuvres (*works of art*), des propositions de sélection d'images qui sont valorisées par le musée (*sets of works*) et des discours de vulgarisation sur ces œuvres (*stories*). Cette partie du service rassemble à peu près toutes les caractéristiques des deux premières tendances stratégiques. Pourtant, ces fonctionnalités ne sont pas au centre de ce service, qui en proposent d'autres par ailleurs (cf. **infra**).

En outre, nous avons suggéré que le *Rijksstudio* est conçu en suivant une volonté d'utiliser le potentiel interactif de l'interface Web pour manipuler les collections artistiques du musée en fonction des intérêts personnels de l'utilisateur. Un filtre permet

caractéristiques, tels que la couleur, et peut fournir des tableaux de ressources très détaillées. Cet outil permet également de combiner les ressources avec d'autres contenus, par exemple, au sein d'une entreprise particulière, ou encore sur les réseaux sociaux et la recherche Web. Pour ce faire, les ressources du musée sont diffusées à l'aide de *Leaflet* (cf. https://leafletjs.com/reference-1.4.0.html) dont le but est d'améliorer la diffusion de contenus sur Internet en intégrant différentes fonctionnalités dans l'interface Web.

de réduire les propositions d'œuvres qui sont exposées (à peu près 8000 images, réparties dans 20 000 sélections). Ce que nous voulons mettre en évidence, c'est que la stratégie qui guide ce projet prend en compte que les utilisateurs sont habitués à d'autres pratiques sur Internet, pratiques qui mobilisent la participation active de l'utilisateur, autrement dit, son « intégration » dans le dispositif. Pour ce faire, cette partie du service se base sur des *sets*, c'est-à-dire des sélections d'images (des corpus). Une partie de ces sélections sont réalisées par l'institution patrimoniale (fig. 74 : quelques exemples de sélections réalisées par l'institution). Néanmoins, l'objectif est de favoriser la manipulation des collections d'images en proposant à l'utilisateur un instrument pour composer ses propres corpus à partir des images qu'il découvre sur le site Web (fig. 75 : exemple de sélection d'images d'un utilisateur du service ; dans ce cas-ci, l'utilisatrice a rassemblé différentes images d'animaux).

La stratégie de ce projet met en évidence un élément crucial, car elle indique comment les institutions se positionnent par rapport à d'autres pratiques avec lesquelles ces institutions s'estiment en compétition directe. Si, dans un premier temps, nous avons vu que les institutions s'attribuent le rôle de détenteur d'un savoir à communiquer, par là même, ces institutions ne prévoient pas un autre rôle pour l'utilisateur que celui de spectateur motivé par une volonté de savoir. Dans un deuxième temps, les institutions renouvellent surtout leurs stratégies afin de transformer le rôle de l'utilisateur, qui n'est plus uniquement caractérisé par une « volonté de savoir », mais plus généralement par une *volonté de participer*, mais d'une manière différente que *WikimediaCommons*. Ce renouvellement est le résultat de changements plus généraux dans les pratiques numériques, notamment la concurrence d'autres services Web dont le but est la communication en ligne d'images, mais dans d'autres contextes et pour d'autres raisons.

L'ensemble des ressources visuelles (*works of art* ; plus ou moins 750 000 images numériques) et des sélections d'images (plus ou moins 150 000 *sets*) qui sont disponibles dans ce service ne se limite pas à inventaire ou à une vitrine. À travers un dispositif de lecture participatif, ces ressources sont le fruit d'une rencontre, certes virtuelle, entre une institution cherchant le renouvellement et des internautes habitués à une implication plus importante. La question n'est pas de savoir si ce service est aussi populaire que *Instagram*. Il s'agit surtout de souligner que cette institution veut intégrer les « points de vue multiples » des internautes. Cette multiplicité des points de vue est différente de celle qui est manifestée par les « listes de résultats » pour au moins une raison : dans les listes, seul

l'utilisateur aperçoit le point de vue qu'il génère par sa recherche ; dans le cas du Rijksstudio, la multiplicité des points est intégrée *dans* l'interface Web *de façon permanente.*

Nous présenterons rapidement trois fonctionnalités du service Web qui emploient davantage le potentiel participatif et interactif de la page-écran.

(1) Les *Stories* proposent différents « parcours thématiques » qui sont répartis en fonction de différentes catégories générales (*news* ; *most popular* ; *masterpieces* ; *themes* ; *Dutch Masters* ; *Behind the Scenes* ; *Research* ; *Exhibition*). À travers ces catégories, des sous-catégories spécifiques renvoient chacune vers d'autres pages Web qui proposent des applications Web dont le but est justement de « de voir, de lire et d'entendre des histoires révélatrices du *Rijksmuseum* » (cf. fig. 76) sur des œuvres, mais également de découvrir les activités de conservation et de recherche scientifique du musée. Au contraire des services Web précédents, il faut remarquer que ces *Stories* ne reposent pas sur des discours verbaux, mais plutôt sur des vidéos explicatives ou d'autres applications interactives.

Par exemple, la catégorie « *masterpieces* » comprend plusieurs sous-catégories (*Rembrandt* ; *Vermeer* ; *Gallery of Honor* ; *Before 1600* ; *1600-1650* ; etc.). Dans cette perspective, chaque sous-catégorie comprend plusieurs séquences vidéos, dont la durée varie de quelques minutes (p. ex. : la reproduction numérique de l'œuvre est rapidement commentée par une voix off ; éventuellement, le cadre de cette animation vidéo agrandit l'image et se déplace dans la reproduction en fonction des commentaires ; cf. fig. 77) à plusieurs minutes, voire une vingtaine de minutes (p. ex. des vidéos explicatives sur le style d'une peinture et comment le reproduire ; des documentaires courts sur des thématiques précises). Ainsi, l'écrit est remplacé par l'audiovisuel, même si le contenu des *Stories* s'apparente clairement à un discours de vulgarisation. L'objectif reste la même (« faire connaître » et « faire voir »), sauf que la forme d'expression pour réaliser cet objectif change.

(2) Le service Web comprend également l'application *Then and now* (cf. fig. 78 & 79) qui compare des peintures de la ville d'Amsterdam à des photographies actuelles selon le même point de vue. Dans ce cas-ci, l'interface propose un parcours à partir de points d'intérêt numérotés sur l'image. Chaque point d'intérêt active le dispositif de comparaison entre peinture et photographie. Certes, cette mise en scène apporte quelques

informations verbales sur des détails de l'image, notamment par le moyen de légendes qui sont intégrées dans l'interface. Mais nous observons que le but est surtout de confronter certaines zones d'une peinture à des photographies. L'espace de la page-écran s'efface au profit des détails de l'image, en tant que surface à scruter ; détails auxquels des photographies font écho, à travers un changement de support et de temporalité.

(3) Les *Rijksmuseum Walks* illustrent également l'effort de l'institution pour proposer des parcours d'exploration qui sortent du canevas et des normes de l'exploration des services que nous avons présentés jusqu'à présent. La fonctionnalité de cette application interactive est présentée par une courte description :

> Say Rijksmuseum and you say Amsterdam. Amsterdam as the home of Rembrandt, but also of Marten and Oopjen. Amsterdam as the centre of 17th century trade and the gate to Asia. Walk with us through this historic city and learn more about the history of our home town.

Trois ballades virtuelles sont proposées à l'utilisateur (*Rembrandt Walk*, *Asia Walk*, *Marten and Oopjen Walk*) par le moyen d'une image (cf. fig. 80). Chaque image est un hyperlien qui renvoie sur une application interactive : un dessin schématique de la ville d'Amsterdam au XVII^{e} siècle (cf. fig. 81) ; cette « carte dessinée » met en évidence le parcours de Rembrandt dans cette ville, avec différentes étapes illustrées par des ressources du *Rijksmuseum* (cf. fig. 82). Dans la plupart des points d'intérêt, ceux-ci sont accompagnés d'une légende et d'une image (peinture, dessin, photographie) conservée par l'institution. Comme dans l'exemple précédent, l'interface propose un parcours en fonction de plusieurs points d'intérêt, mais dans ce cas, ces points sont intégrés sur un dessin schématique d'Amsterdam.



Figure 72 : Page d'accueil du *Rijksmuseum.*

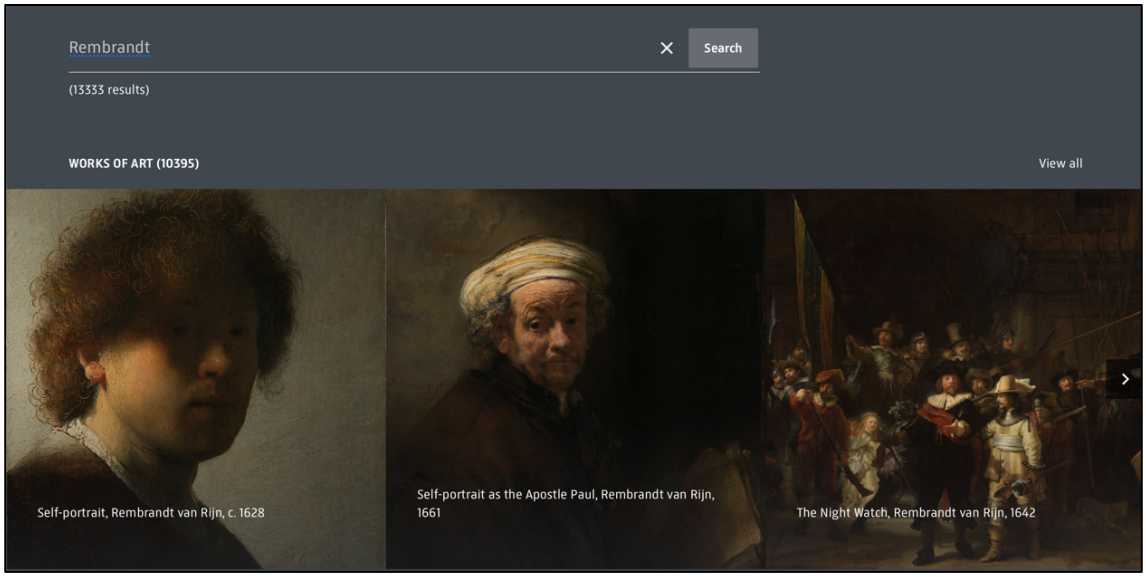


Figure 73 : Moteur de recherche (*search the collection*) du *Rijksmuseum.*

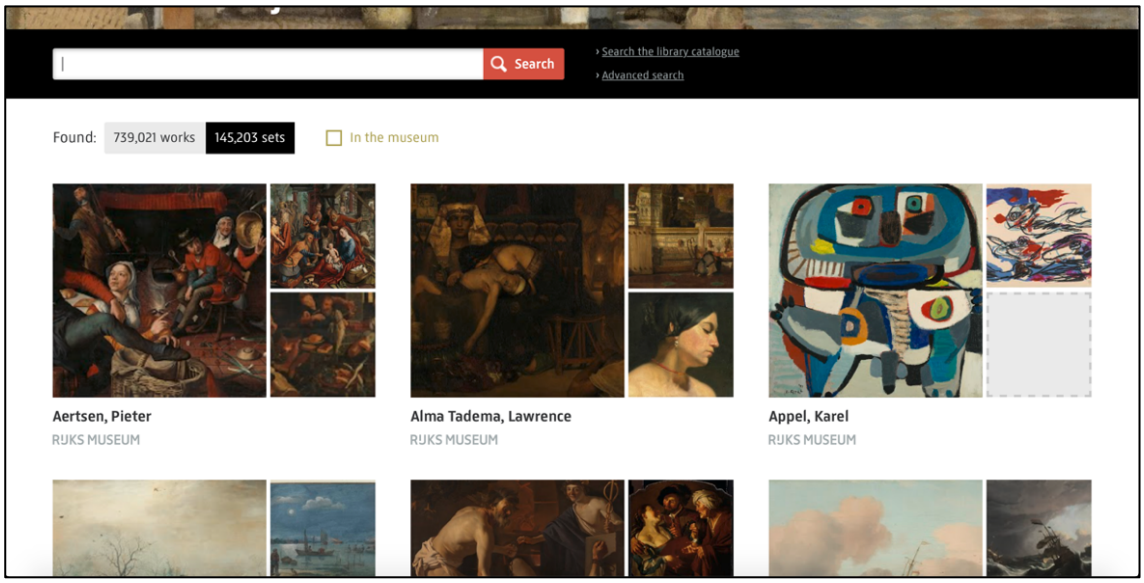


Figure 74 : Sélection d'images sur le *Rijksstudio* (sélection réalisée par le *Rijksmuseum*).

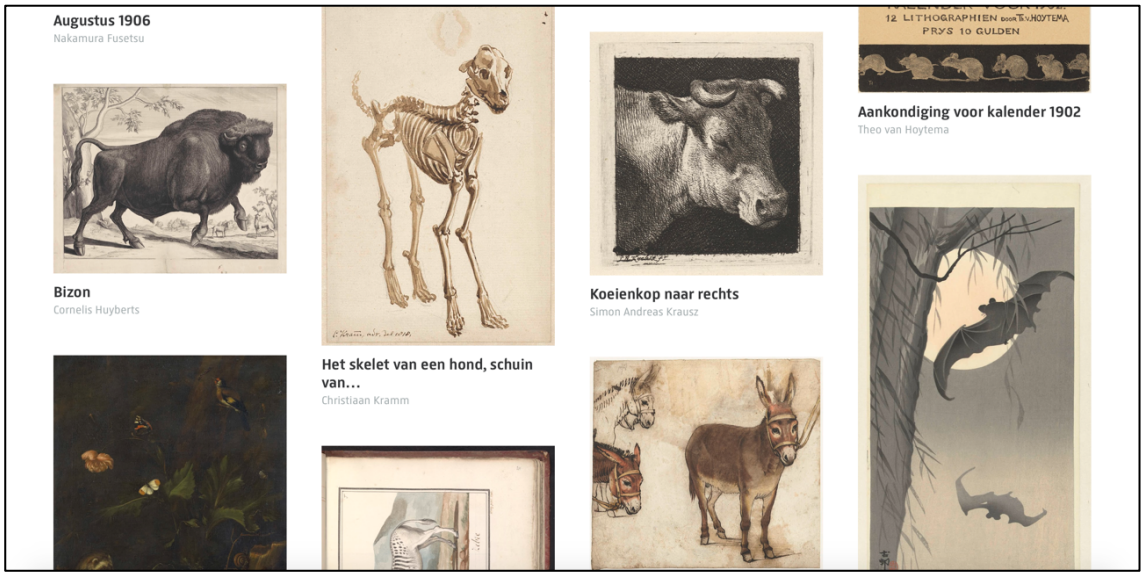


Figure 75 : Sélection d'images sur le *Rijksstudio* (sélection réalisée par une utilisatrice).

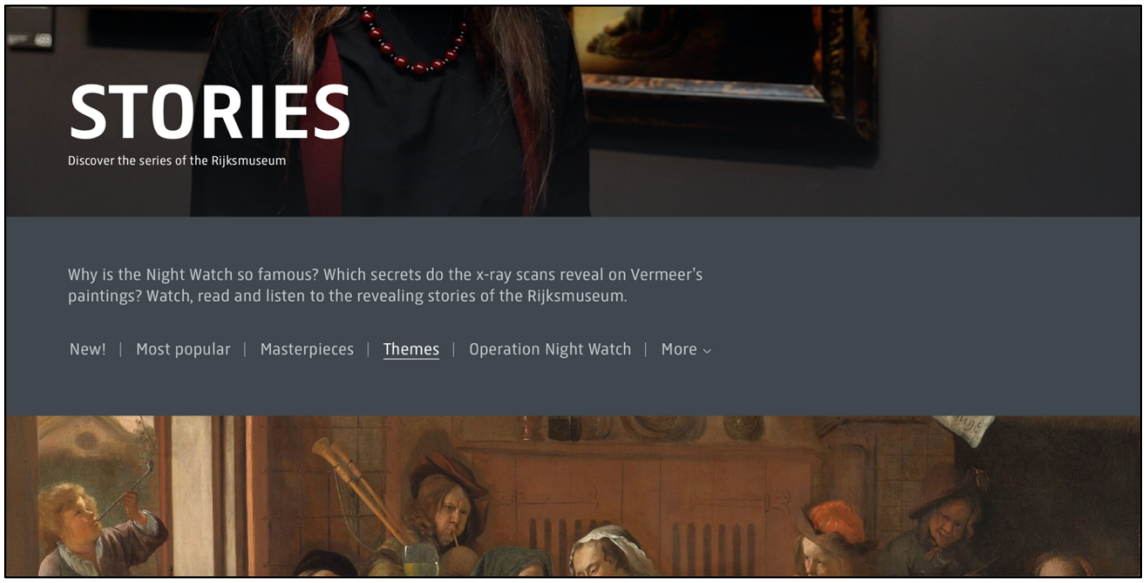


Figure 76 : Page d'accueil des *stories* (*Rijksmuseum*).

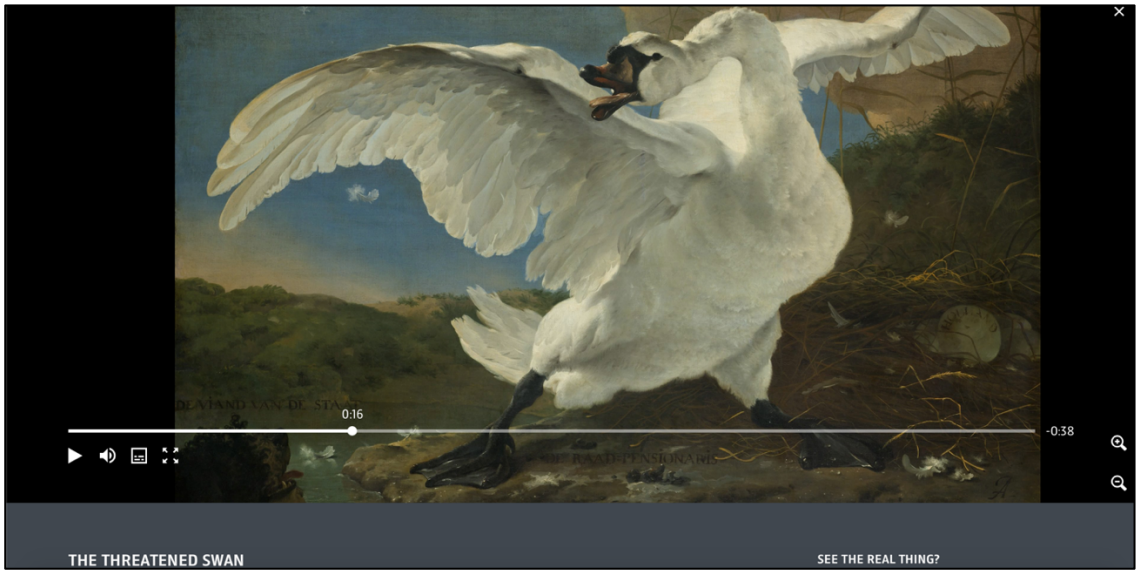


Figure 77 : *Story* sur le Cygne menacé de Jan Asselyn, ca. 1640-1650, sous la forme d'une vidéo explicative (*Rijksmuseum*).

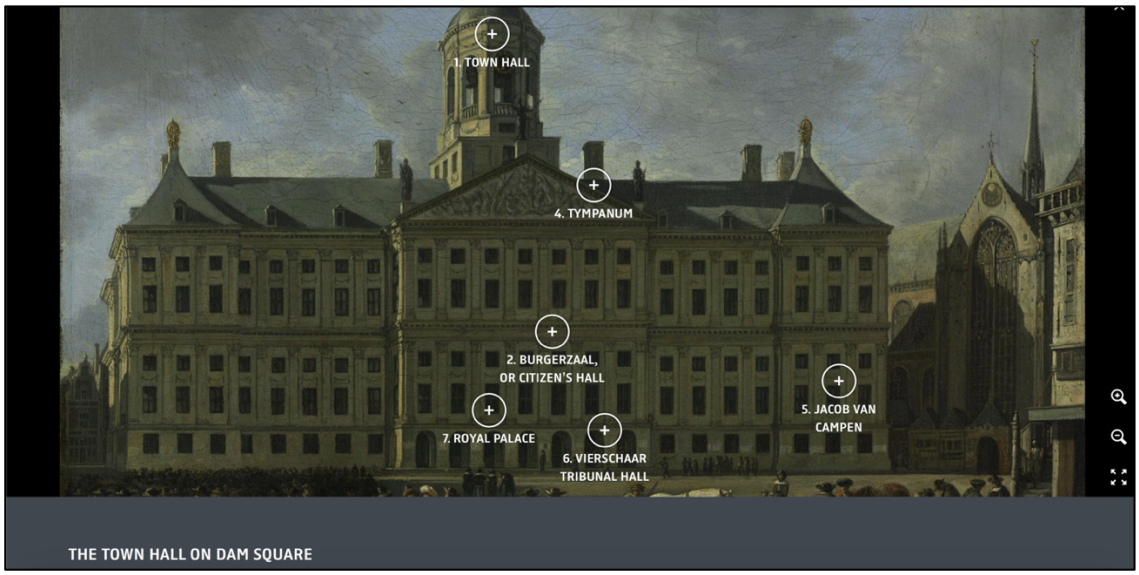


Figure 78 : Application *then and now* (*Rijksmuseum*).

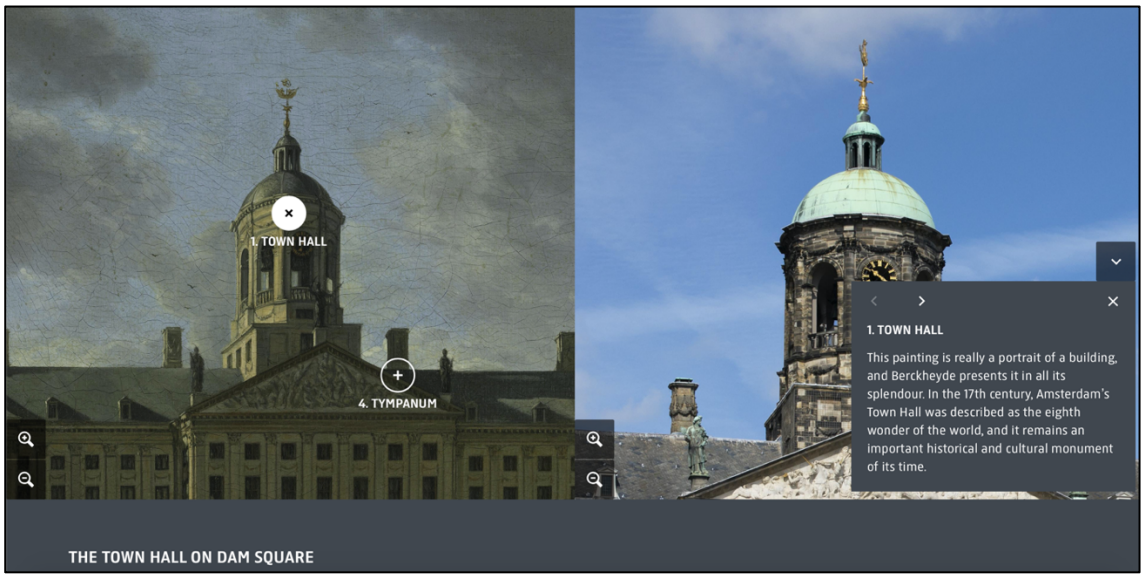


Figure 79 : Application *then and now* (*Rijksmuseum*).



Figure 80 : Page d'accueil des *Rijksmuseum walks*.

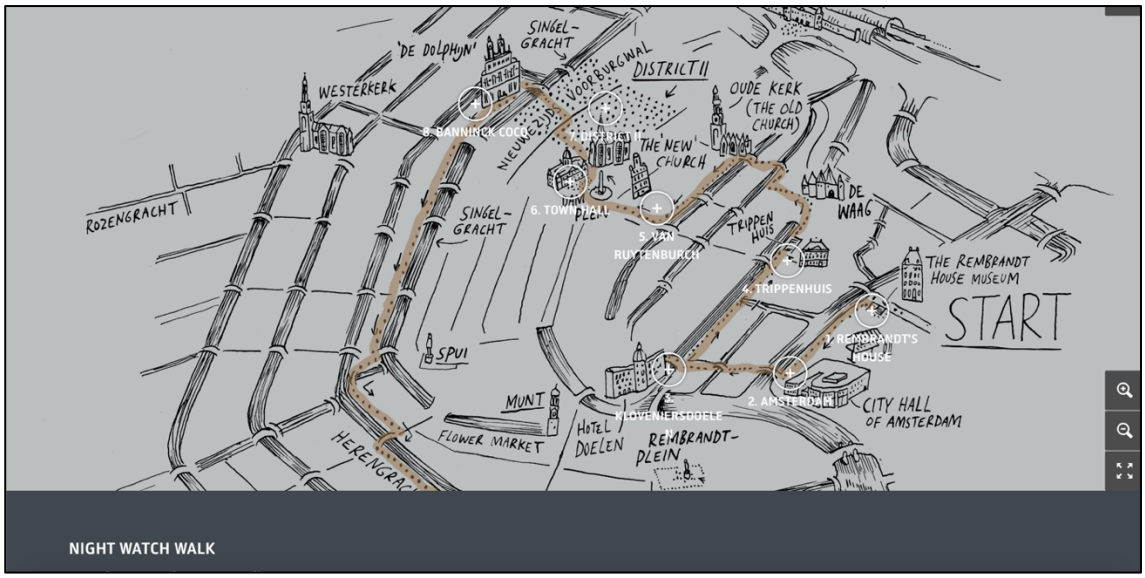


Figure 81 : Exemple de ballade virtuelle (*Rijksmuseum walks*).

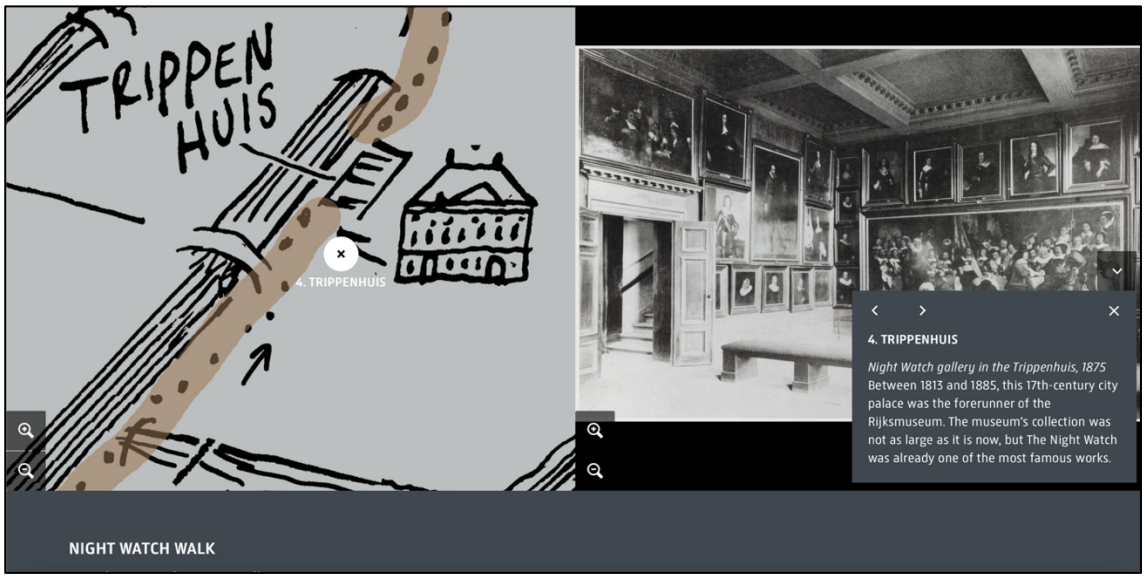


Figure 82 : Exemple de « ballade virtuelle » (*Rijksmuseum walks*).

Dans tous les projets précédents, nous avons constaté la place centrale accordée aux inscriptions graphiques dans l'énonciation de la page-écran, parce que celle-ci est conçue comme un espace qui doit être lisible. Au contraire, nous constatons que la stratégie du Rijksmuseum est d'essayer de changer cette conception. Dans l'espace de la page, les informations textuelles sont reléguées à un plan secondaire, tandis que les reproductions numériques des œuvres sont mises en avant-plan dans l'interface ou encore dans les applications (cf. **supra**). Ce choix est étroitement lié aux principes mis en avant par la mission de l'institution : rendre les œuvres visibles à l'aide de l'informatique. En outre, la rénovation du site Web du Rijksmuseum montre qu'une institution patrimoniale propose un service visant une autre relation d'énonciation. Cette volonté de renouvellement révèle l'évolution des choix stratégiques des institutions patrimoniales qui cherchent à adapter leurs activités sur Internet en développant des applications qui nécessitent la participation de l'utilisateur.

En effet, l'un des buts du projet est de mettre en valeur des reproductions en haute qualité des œuvres et de proposer un instrument de manipulation de ces reproductions d'images. L'utilisateur peut donc observer les moindres détails de la « surface dématérialisée » de ces images, tout en bénéficiant d'une contextualisation qui fournit

toutes les informations de l'institution sur ces images. Même si l'institution est à l'origine de toutes ces informations produites dans le cadre d'une communauté discursive, notamment à travers les métadonnées descriptives proposées par les institutions, nous constatons que ce dernier projet se focalise davantage sur l'exploitation de la *dimension visuelle* de la relation d'énonciation qui est mise en scène par l'interface, au contraire d'autres projets qui se basent plutôt sur une *dimension graphique*. Dans le cas du Rijksmuseum, l'énonciation visuelle occupe une place beaucoup plus importante dans la communication qui est en jeu dans l'interface, quoique celle-ci mobilise encore des structures graphiques et des éléments verbaux pour favoriser la lisibilité de la page-écran.

Si les premières stratégies des institutions se sont concentrées sur la mise en place de bases d'images numériques ou des expositions virtuelles, d'autres institutions proposent des services Web où nous observons une métamorphose de l'interface, dans le sens où l'espace de la page-écran permet à l'utilisateur d'emprunter d'autres voies en multipliant les points de vue pour découvrir des images. Le deuxième renouvellement stratégique exploite davantage la *spatialisation de la page-écran* pour proposer des possibilités de scénarisation des ressources qui s'éloignent progressivement de l'*utilisateur-explorateur* ou de l'*utilisateur-lecteur*, au profit de la construction d'un autre type de point de vue : celui de l'*utilisateur-observateur*.

Cette spatialisation est ce qui différencie l'interface proposée par le Rijksmuseum des autres interfaces Web que nous avons présentées. C'est notamment pourquoi ce projet a été développé en vue d'un usage sur tablette tactile afin de prendre acte de l'évolution des pratiques numériques. L'interface manifeste surtout de larges boutons à l'aide desquels l'utilisateur peut interagir avec ces doigts, ce qui modifie radicalement l'expérience que l'on fait face à une œuvre d'art, que l'on peut dès lors « toucher ». Dans cette perspective, l'interface Web ne vise pas seulement à présenter une sélection automatique d'images, mais elle cherche surtout à établir une relation d'interaction avec l'utilisateur par le moyen d'un autre *style de visibilité*.

7.3. Autres exemples de communication en ligne dans une perspective interactive

Nous pouvons rapidement présenter d'autres projets qui s'appuient sur la même idée, à savoir renouveler l'interface Web afin de proposer une mise en scène interactive où l'utilisateur participe à la construction de l'énonciation numérique en cours de déploiement. Dans la section suivante, nous présenterons encore trois exemples de projets qui utilisent le traitement informatique pour aider la recherche en sciences humaines et sociales dans une perspective interactive. Nous remarquerons que cette volonté de renouvellement des services Web favorise d'autres types d'utilisation de la communication en ligne pour interagir avec des images ou des ressources audiovisuelles.

7.3.1. *12 Sunset : Exploring Ed Ruscha's Archive*

En plus de proposer un instrument de recherche (*Getty Search Gateway*), l'institution *Getty* a également participé à la conception du projet *12 Sunset : Exploring Ed Ruscha's Archive.* Ce projet a conçu un dispositif interactif pour explorer des archives photographiques de l'institution ; à l'origine, ces photographies servaient de support créatif pour l'artiste Ed Ruscha qui a photographié pendant plusieurs décennies l'architecture et de nombreuses rues de Los Angeles.

Ces archives d'images sont conservées au sein du *Streets of Los Angeles Archive*, qui comprend plus de 60 000 photographies représentant l'évolution du *Sunset Boulevard* entre 1974 et 2010). Toutes les images ont été numérisées et couplées à une interface Web. D'une manière similaire aux ballades virtuelles du Rijksmuseum, ce projet aboutit sur un dispositif de lecture qui représente ces reproductions numériques à l'aide d'un outil de cartographie en ligne (cf. fig. 83 & 84). Grâce à ce dispositif, l'utilisateur explore différentes portions du *Sunset Boulevard* à travers une juxtaposition de deux archives d'images ; cette juxtaposition offre deux points de vue en fonction d'un positionnement affiché sur la carte.

L'espace de la page-écran est divisé en quatre parties (de haut en bas) : (1) un menu principal (avec quatre hyperliens : *Drive*, *Sort*, *About*, *Share*) ; (2) une reproduction d'une archive ; (3) la carte ; (4) une seconde reproduction d'archive, mais inversée. Dans cet espace, la position de l'utilisateur sur la carte est représentée par le moyen d'un « simulacre visuel » : une voiture rouge qui peut être déplacée en utilisant le clavier de

l'ordinateur (les touches directionnelles déplacent la voiture vers la gauche ou vers la droite). Lorsque ce simulacre est déplacé, le point de vue sur le boulevard change et, par conséquent, les reproductions d'images changent. En outre, étant donné que cette collection d'images rassemble des images produites durant différentes années, les touches directionnelles haut et bas permettent de modifier le point de vue en fonction de l'année de production.

Les reproductions numériques des photographies qui construisent ce point de vue virtuel sur le boulevard sont également des liens hypertextes renvoyant sur une mise en page qui fait ressortir des *visual tags*. Lorsqu'une photographie est sélectionnée, celle-ci est présentée dans une autre page comprenant la reproduction numérique, le lieu de boulevard qui est représenté par l'image, différents visual tags, ainsi que les éléments textuels que le traitement informatique a reconnus dans l'image (fig. 85). De plus, l'interface propose aussi une fonctionnalité pour sélectionner des images au hasard et les présenter sous la forme de mosaïques d'images, avec la possibilité d'appliquer des filtres linguistiques (fig. 86).

Par rapport aux autres scénarisations des ressources que nous avons présentées, celle-ci intègre le positionnement de l'internaute dans le dispositif de lecture. Cette position représente et simule, dans l'acte d'énonciation, un double point de vue : (1) un point de vue spatial et (2) un point de vue temporel. Dans ce service, en plus du principe de structure modulaire[450], c'est surtout la dimension interactive apportée par l'application Web qui contrôle l'*aspectualité* du procès en cours en fonction de deux paramètres (lieu et temps) ; cet usage du numérique est donc très différent des choix stratégiques qui s'appuient sur le langage verbal pour « contrôler » le point de vue de l'utilisateur sur le résultat de ce processus d'énonciation. Concernant ce processus, nous remarquons que l'instance d'énonciation est manifestée dans la page-écran (le logo *Getty*, en bas à gauche), de même que la place prévue pour le spectateur (la voiture). D'un côté, cette instance projette un « non-ici » (débrayage spatial) et un « non-maintenant » (débrayage temporel) dans l'espace figuratif de la page (plan et cadre de la représentation). De l'autre, les gestes du spectateur déterminent les déplacements (le parcours du simulacre) dans cette espace à travers une simulation. Cette simulation se

[450] **Chap. 4, sect. 4.1.2.**

base à la fois sur cet effet de représentation et sur ces opérations de débrayage qui sont produits par le dispositif.

Comme dans le cas du Rijksmuseum, nous pouvons remarquer un changement en faveur de l'intégration d'applications interactives. En effet, cette intégration a pour but de simuler une interaction avec les ressources en proposant un « point de vue » qui est représenté dans le processus d'énonciation. Dans ce dispositif, où l'énonciation vise et se mélange à la simulation, l'exploration est étroitement liée aux gestes de l'utilisateur. D'un côté, une manipulation simulée se substitue en partie aux gestes des usagers. De l'autre, cette manipulation, en tant que *prothèse technologique*, offre de nouveaux points de vue qui engagent d'autres types de gestes énonciatifs.

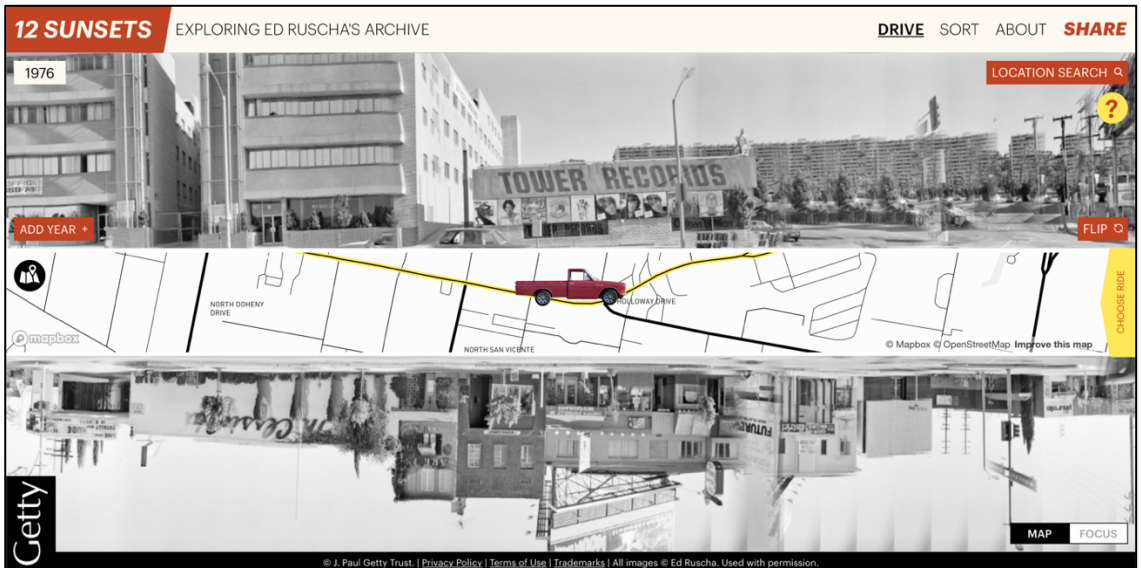


Figure 83 : Interface Web du projet *12 Sunset* qui présente un double point de vue sur Sunset Boulevard en fonction de deux paramètres : la position sur la carte et l'année (ici 1976).

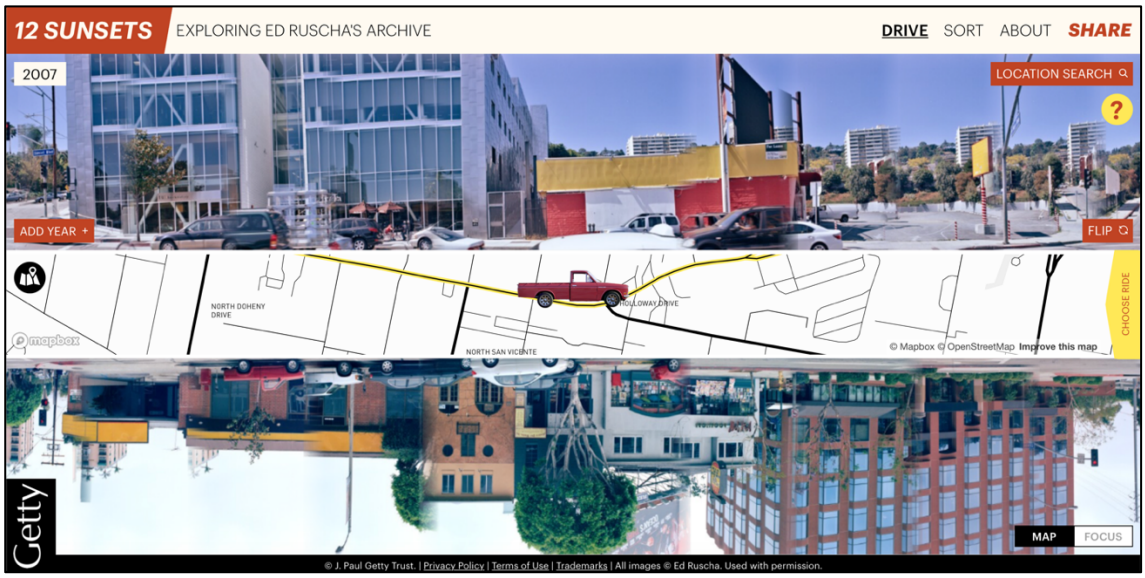


Figure 84 : Interface Web du projet *12 Sunset* qui présente un double point de vue sur Sunset Boulevard en fonction de deux paramètres : la position sur la carte et l'année (ici 2007)



Figure 85 : Présentation d'une ressource unique dans le projet *12 Sunset*.



Figure 86 : Présentation d'une sélection aléatoire d'images ; en haut, un filtre permet de réduire les résultats en fonction d'une période temporelle, de tags visuels ou d'éléments textuels dans l'image.

Dans un travail antérieur (cf. Honba Honba, 2019), nous avons montré que le design des interfaces peut participer à une discrétisation du dispositif et à un effacement des « gestes somatiques » au profit de la simulation et de la manipulation virtuelle de ressources numériques. Si certains dispositifs ont tendance à privilégier l'immersion de l'utilisateur, parfois au détriment de l'expérience corporelle, ces dispositifs proposent aussi de nouvelles manipulations qui peuvent trouver un sens au sein des pratiques numériques. En prenant en compte la dimension de l'expérience et du corps dans les pratiques de communication, nous avons défendu qu'il s'agit d'élucider davantage le statut de ces « prothèses technologiques » dans les processus de médiation du sens, notamment en considérant l'influence du support numérique et des interfaces-utilisateurs sur les pratiques communicationnelles parmi lesquelles la technologie numérique est intégrée et utilisée. L'étude de la dimension communicationnelle de la simulation informatique mérite sans aucun doute d'être approfondie, en priorité la relation entre les corps et les gestes de l'usager et la représentation de ce corps et de cette gestualité dans l'énonciation de cette simulation.

Nous pouvons d'ailleurs remarquer deux autres initiatives intéressantes qui vont dans ce sens.

7.3.2. *VISUALL-tek : Ikonik Analysis Toolkit*

Développé dans le cadre du projet *ANR VISUALL* (2013-2017) et de la plateforme « Sciences et cultures du visuel » au sein de l'Institut de Recherches Historiques du Septentrion (CNRS/Lille 3), VISUALL-tek est une bibliothèque d'outils informatiques qui permettent de questionner des images d'une manière collective, c'est-à-dire à travers une expérience de groupe[451]. Celle-ci est le résultat de la collaboration entre le sociologue Mathias Blanc (coordinateur du projet) et les ingénieurs informaticiens Cécile Picard-Limpens (montage et suivi du projet scientifique) et Julien Wylleman (développeur et concepteur des applications informatiques). Cette collaboration a abouti sur trois applications : *Ikonokat*, *V-Atlas* et *Coddeix*[452].

Du point de vue théorique, le projet *ANR VISUALL* vise à identifier plusieurs débats actuels qui sont relatifs aux images dans les traditions scientifiques allemande et anglophone, et surtout à étudier leurs cadres d'analyses pour en assurer la diffusion et la discussion en France. Dans nos pratiques socioculturelles, la place de plus en plus importante accordée aux images, en tant que moyen de significations, conduit à questionner à nouveaux frais le cadre épistémologique, ainsi que les méthodes d'analyse des images qui sont parfois inadéquates dans le contexte et les pratiques actuels[453]. Le but est de rassembler les différents points de vue dans le but d'établir des références de base qui permettent d'étudier les objets visuels. Le projet *ANR VISUALL* s'inscrit donc dans une perspective pluridisciplinaire, qui entrelace l'histoire de l'art, la sociologie et les sciences informatiques, afin de considérer d'une manière innovante la réception des œuvres par les publics.

L'application *Ikonokat* (*Ikonik Analysis Toolkit*) est outil d'analyse qui met en évidence les caractéristiques visuelles d'œuvres picturales sans pour autant passer par le

---

[451] http://www.agence-nationale-recherche.fr/Projet-ANR-13-PDOC-0005 ; https://visuall.hypotheses.org. En outre, les informations sur ce projet sont issues d'une conférence de Mathias Blanc, dispensée à l'Université de Liège (12 novembre 2018).

[452] http://www.visuall-tek.org ; dans ce travail, nous ne considérons pas l'application *Coddeix*, qui est une application dont le but est d'aider à analyser les interactions sociales dans des enregistrements vidéo.

[453] Sur ce sujet, voir les discussions dans le **chap. 3**.

biais du langage. En effet, dans le cadre d'une visite dans un musée de peinture, le spectateur est équipé d'une tablette tactile qui représente successivement les différentes œuvres du musée reproduites par le moyen de copies numériques. Le spectateur a ainsi la possibilité d'admirer les œuvres « originales », et surtout de dessiner des traits relatifs aux éléments qu'il considère comme caractéristiques ou importants dans chaque tableau (p. ex. : entourer des personnages, des objets ou des zones dans l'image, tracer des droites ou des courbes qui relient ou opposent des éléments). L'utilisation du « tracé » sur une image permet en quelque sorte de proposer un « questionnaire visuel », plutôt que de chercher les mots adéquats pour décrire son expérience personnelle vis-à-vis d'une œuvre ; le spectateur indique et montre directement sur la surface de l'image les « zones visuelles » qui ont retenu son attention, sans pour autant être contraint par une barrière linguistique, ou encore l'âge du spectateur.

L'application *Ikonokat* nécessite un dispositif qui recueille les différents tracés afin de constituer un ensemble de données. Pour ce faire, l'équipe du projet a développé un outil informatique dont le but est de récolter les tracés des spectateurs, puis de leur appliquer un traitement numérique pour obtenir des statistiques iconométriques concernant la consécution des tracés. Il est ainsi possible d'ébaucher autant de parcours de lecture de l'image qu'il y a de participants, ce qui offre de nouvelles possibilités d'échantillonnage. En fournissant ainsi des données précises et individuelles sur les observations des spectateurs, sans pour autant utiliser un questionnaire, l'avantage principal de cette application est d'offrir un instrument aux chercheurs qui s'intéressent à la réception esthétique des œuvres d'art dans les musées.

D'autre part, l'application *V-Atlas* (*Video Analysis TooL Assisted*) propose d'identifier, de classer, de catégoriser des motifs iconographiques et narratifs, relevés individuellement par l'utilisateur dans un ensemble d'images fixes ou animées. L'application s'inspire des principes de « montages » de *l'Atlas mnémosyne* d'Aby Warburg[454] en l'adaptant aux technologies contemporaines du Web et du numérique. Rappelons-le, Warburg regroupait des images d'œuvres ou d'artefacts sur des planches pour constituer cet *Atlas*, ce qui favorise, selon certains historiens de l'art, une analyse anthropologique et comparative des œuvres, en tant que trace qui renvoie au passé, mais

[454] **Chap. 3, sect. 3.1.2. & sect. 3.4.**

aussi au présent (par exemple, à travers sa reproduction dans une sphère culturelle donnée).

Dans le même sens, *V-Atlas* permet d'intégrer tout type de médias (images fixes, mais aussi en mouvement) sur les planches d'un « atlas numérique », afin d'évaluer empiriquement des rapprochements ou relations entre différents objets visuels, tout en fournissant un moyen de consulter en même temps des images fixes et animées. Le but du dispositif est d'accompagner les modalités interprétatives de tout un chacun face à des corpus d'images animées ou fixes. Pour ce faire, l'application met en place une interface logicielle qui sert de support à un processus de visualisation d'images, que l'usager peut explorer et organiser, seul ou en groupe, selon des motifs iconographiques, rythmiques ou narratifs relevés dans le corpus. Chaque action de l'utilisateur est enregistrée de telle sorte que cet outil offre la possibilité d'analyser différents parcours dans les corpus d'images qui ont été organisées par les usagers durant la visualisation.

En résumé, ce projet vise donc à proposer de nouveaux outils d'analyse des images, à partir de programmes informatiques qui permettent d'enregistrer et d'analyser des manières dont le public caractérise visuellement des œuvres d'art. Dans ce cas, l'instrumentalisation informatique est utilisée afin de rassembler et de mettre automatiquement en évidence des caractérisations d'image réalisées par des humains ; mais des caractérisations qui sont accompagnées par des dispositifs techniques, qui permettent de traiter les données et de les mettre en relation au sein d'une représentation, pour pouvoir les analyser dans une perspective sociologique.

7.3.3. *Rekall*

Pour terminer cette section, nous pouvons rapidement évoquer le projet *Rekall* de Clarisse Bardiot (cf. Bardiot, 2009, 2015a, 2015b & 2015c ; Bardiot *et al.*, 2014). *Rekall* est un environnement informatique en libre partage utilisé pour documenter et analyser les processus de création et faciliter la reprise des œuvres dans le domaine des arts du spectacle (théâtre, danse, cirque, marionnette, installations, etc.). Il s'agit d'un ensemble d'outils qui permettent de documenter les spectacles et les installations plastiques, en prenant en compte le processus de création, la réception et les différentes formes qui participent à la création artistique. Le but est d'offrir à l'utilisateur une vision synthétique des types, de la quantité et de l'organisation des documents relatifs à une œuvre, tout en rendant compte des technologies utilisées pour mettre en scène le spectacle.

Rekall a été conçu pour remédier à deux problèmes relatifs aux arts du spectacle : (1) les artistes, les régisseurs rencontrent parfois des difficultés lorsqu'il s'agit de reprendre un spectacle, une installation dont les technologies sont devenues obsolètes ; (2) les chercheurs en arts de la scène sont actuellement confrontés à des ensembles de documents très nombreux et de types très différents, qui nécessitent des outils pour les analyser. *Rekall* s'utilise comme une application Web, mais une application que ne nécessite pas de connexion Internet, car l'analyse des documents est directement réalisée sur le disque dur de l'ordinateur ou d'un disque externe de l'utilisateur. L'application permet d'accumuler un nombre infini de documents (textes, images, vidéos, etc.) en lien avec une œuvre, et de les articuler selon plusieurs « strates temporelles » : (1) le processus de création ; (2) la représentation elle-même (voire ses différentes versions) ; (3) la réception (ou « l'après-spectacle »). Elle se base sur les métadonnées présentes dans chacun de ces documents pour en extraire des informations caractéristiques (auteur, date de création, lieu de création, mot-clé, etc.), qui sont ensuite utilisées par les outils d'analyse et de représentation de l'information, qui mettent en évidence des relations et des informations insoupçonnées.

7.4. Conclusions

Dans l'ensemble, cette **troisième partie** montre une évolution dans les usages du support numérique pour diffuser en ligne des archives et des collections d'images.

Nous avons mis en évidence deux tendances stratégiques. La première tendance a pour but de représenter le patrimoine culturel d'une région (un pays) par le moyen de substituts photographiques. Ces « substituts » sont d'abord convertis en documents numériques, et par la suite, ces documents sont transmis via Internet et par des écrans d'ordinateur. Par la suite, nous avons vu qu'une deuxième tendance favorise la coopération entre différentes institutions, mais aussi la coopération entre les internautes pour agglomérer des ressources visuelles. Cette deuxième tendance stratégique a notamment permis de regrouper différentes ressources visuelles dans le but de standardiser les archives d'images et les métadonnées de plusieurs institutions, même si ces ressources peuvent être rassemblées pour différentes raisons (par exemple : défendre un sentiment d'identité européenne ou proposer une médiathèque où des ressources en libre partage sont fournies par les internautes).

Néanmoins, comme nous l'avons remarqué, ce changement n'évacue pas complètement le problème de la classification des images qui, nous l'avons vu, se base encore sur un système informatique qui traite et classe avant tout des images selon leurs métadonnées descriptives. Même si cette deuxième tendance renouvelle les fonctionnalités des interfaces Web, elle se fonde également sur une compréhension et une classification qui s'appuient entièrement sur une *traduction des images à travers un langage*.

Si nous avons constaté que la communication en ligne maintient certaines contraintes issues de la première tendance stratégique (les bases d'images numériques), d'autres initiatives intègrent de plus en plus des fonctionnalités qui correspondent mieux aux spécificités du numérique. Il s'agit d'une étape considérable dans l'évolution des stratégies et, plus généralement, de l'économie des larges ensembles d'images sur Internet.

Dans ce **septième chapitre**, nous avons considéré davantage cette deuxième tendance stratégique en remarquant qu'il s'agit de plus en plus de mettre l'accent sur la dimension interactive de la navigation Web. Nous avons vu que des « expositions virtuelles » débouchent sur des interfaces plus ergonomiques et qui proposent d'autres fonctionnalités pour explorer les ressources numériques. La mise en scène du site Web est alors conçue comme un dispositif de communication où les parcours qui sont proposés sont conçus comme un « espace à observer », « à scruter », autrement dit : un espace qui donne *à voir* plutôt qu'*à lire*. Ces applications interactives sont des exemples de conceptions qui s'articulent sur la structure modulaire des nouveaux médias, mais dans une perspective réflexive. Dans les exemples analysés, cette perspective est remarquable à travers un changement dans l'énonciation de l'interface : de plus en plus de « points de vue » sont proposés à l'observateur en fonction des images et non pas uniquement en fonction d'un métalangage. En réaction à l'augmentation des pratiques numériques et des réseaux sociaux, le réseau Internet n'est plus seulement un lieu de stockage ou le moyen d'une publicité des larges ensembles d'images. Les institutions renouvellent leurs dispositifs de communication en ligne dans le but de s'aligner sur d'autres pratiques qui risquent de faire concurrence aux services Web de ces institutions. Progressivement, le réseau Internet devient un endroit de coopération, de participation et d'interaction, et surtout un « espace omnidirectionnel » qui propose une multiplicité de parcours et de points de vue possibles.

Ce point est important pour comprendre la *troisième tendance stratégique*, car celle-ci a principalement pour objectif d'améliorer les possibilités d'interaction avec un large ensemble d'images numériques en intégrant des dispositifs de visualisation dans les interfaces Web. Comme nous l'avons vu, les techniques de visualisation présentent un point de vue global sur un ensemble de données. En effet, c'est actuellement tout un domaine de recherche qui se regroupe autour de cette problématique de l'interaction et de la visualisation avec une importante quantité de données numériques. En outre, nous avons suggéré que les humanités numériques favorisent un contexte d'échange entre les sciences humaines et les techniques informatiques. L'un des objectifs des humanités numériques est sans doute de mettre en place un champ de recherche pluridisciplinaire concernant l'impact du numérique sur les pratiques scientifiques.

Dans la **quatrième partie**, nous présenterons cette troisième tendance qui concerne surtout des projets s'inscrivant dans le domaine d'analyse des *cultural analytics* et des *big visual data*. D'un côté, dans la continuité de la deuxième tendance stratégique, nous verrons que l'idée est de proposer des services Web qui favorisent la participation active de l'utilisateur à l'aide d'applications Web. De l'autre, nous constaterons comment la troisième tendance essaye progressivement à s'éloigner de la classification d'images en fonction de métadonnées descriptives, au profit d'une classification qui se base sur les caractéristiques visuelles des images[455]. Encore une fois, il s'agit d'utiliser la communication en ligne pour diffuser un large ensemble d'images numériques, mais dans ce cas-ci nous remarquerons un déplacement important en ce qui concerne les usages du traitement informatique d'images et la manière de présenter les résultats de ce traitement dans une interface. Nous verrons que la programmation d'algorithmes et les techniques de visualisation favorisent un renouvellement important des conditions de visibilité et de manipulation des larges ensembles d'images.

[455] Comme nous l'avons vu, la réflexivité du visuel peut être considérée en comparant et en confrontant des images pour mettre en évidence des continuités et des ruptures formelles dans un corpus d'images ; cf. **chap. 3**.

QUATRIÈME PARTIE

ANALYSE DU CORPUS (2) : CULTURE VISUELLE ET *BIG VISUAL DATA*

CHAPITRE 8
Troisième tendance stratégique : communication en ligne et visualisation d'images

Dans la **troisième partie**, nous avons vu comment le contexte des humanités numériques favorise la collaboration entre les sciences humaines, les institutions patrimoniales et des spécialistes en informatique[456]. Dans la plupart des cas, la conception de services de communication en ligne d'images aboutit principalement sur deux dispositifs de communication en ligne : (1) les *bases d'images numériques* et (2) les *expositions virtuelles* comprenant des *applications Web*. Le passage d'un dispositif à l'autre indique une volonté de renouvellement des services Web qui rendent accessibles des archives et des collections d'images[457]. Pour terminer le présent travail, cette **quatrième partie** présente une dernière tendance qui continue ce renouvellement en modifiant les conditions de visibilité et de manipulation des larges ensembles d'images à l'aide des progrès plus récents dans le domaine de l'analyse informatique d'images.

Dans ce **huitième chapitre**, nous verrons comment les stratégies de visualisation transforment les dispositifs de communication en ligne et par conséquent l'expérience de l'utilisateur des larges ensembles d'images numériques sur Internet. Nous présenterons d'abord quatre projets très différents. Si chaque projet propose différentes solutions pour

[456] **Chap. 2**.

[457] Voir **chap. 5 (sect. 5.4), chap. 6 (sect. 6.3.) & chap. 7 (sect. 7.4.)**.

utiliser l'analyse informatique et pour l'appliquer sur de larges ensembles de données culturelles, ces projets partagent chaque fois le même double objectif :

- Montrer les performances de l'informatique dans le traitement, la classification et la visualisation d'un large ensemble d'images.
- Diffuser en ligne les résultats de ces performances par le moyen d'un service Web.

Nous procéderons en deux temps, parce que deux étapes principales peuvent être distinguées en ce qui concerne l'usage des techniques de visualisation. Ces projets révèlent en effet un déplacement de techniques de visualisation qui représentent « abstraitement » les données traitées à des techniques de visualisation qui représentent, « dans la visualisation », des copies de médias (photographies, peintures, dessins).

Dans un premier temps, nous considèrerons des services Web où la visualisation fait principalement ressortir des tendances statistiques relatives aux données et aux métadonnées qui sont traitées par un système informatique. Un premier projet propose une interface Web interactive pour aider la recherche en histoire de l'art. Tandis que le second projet s'appuie sur la *data visualization* pour créer une interface Web qui met en relation plusieurs œuvres issues d'une collection d'images photographiques.

Dans un second temps[458], nous analyserons deux autres projets qui ont proposé des dispositifs de visualisation d'images dans un espace qui s'apparente à un système de coordonnées où les images sont automatiquement classées. Dans ce cas, la visualisation comprend des miniatures d'images qui sont organisées en fonction de leurs similarités visuelles et présentées sous la forme d'une juxtaposition, c'est-à-dire dans un espace où les images qui sont similaires sont proches dans le système de coordonnées, tandis que les images non similaires sont éloignées.

Nous verrons par la suite[459] que de plus en plus de projets s'orientent désormais vers ce second type de visualisation afin d'offrir aux utilisateurs un nouveau type de moteur de recherche pour explorer un large ensemble d'images.

[458] Cf. **infra, sect. 8.2.**

[459] Dans le **chap. 9**, nous présenterons en détail un projet qui est clairement développé pour proposer un instrument destiné aux historiens de l'art et autres chercheurs qui étudient des images.

8.1. La *data visualization* au service des recherches en sciences humaines

Nous l'avons vu, la visualisation des données offre un point de vue sur un ensemble de données en les considérant du point de vue statistique[460]. Ce point de vue peut être considéré comme un exemple de *distant reading* (« lecture à distance ») qui se distingue de la *close reading* (« lecture rapprochée »), c'est-à-dire une analyse qualitative telle qu'elle est souvent réalisée dans des travaux en sciences humaines (cf. Moretti, 2013).

Par exemple, si nous considérons un corpus de peintures, une analyse qualitative « classique » ne permet pas d'observer en un seul coup d'œil l'ensemble d'un corpus d'image ni d'observer les variations, les transformations, l'évolution du style artistique dans cet ensemble. Au contraire, les visualisations en « lecture à distance » présentent la totalité de la production d'un artiste d'un seul coup, au sein d'un « diagramme » où toutes les images sont connexes et situées en fonction de certains paramètres d'analyse.

Ce diagramme constitue donc une nouvelle représentation qui présente, tout en le simulant, un positionnement « à distance » sur un ensemble de données. La question est toutefois de savoir comment ces visualisations donnent à voir des ressources numériques, et surtout de savoir si ces techniques de visualisation apportent des outils utiles pour les approches en sciences humaines qui étudient la culture visuelle[461].

Pour ce faire, les sections suivantes présentent deux services Web qui utilisent la visualisation pour faire ressortir des relations et des dynamiques globales dans un ensemble de données relatives à des pratiques artistiques. Le point en commun de ces deux services est que chacun concerne des archives ou des collections photographiques.

[460] En effet, les techniques de visualisation de données sont souvent utilisées pour représenter abstraitement et quantitativement des relations au sein d'un ensemble de données à l'aide d'un système graphique ; cf. les discussions sur la *data visualization* dans le **chap. 4, sect. 4.2.1.**

[461] Comme dans le cas des applications interactives dans le **chap. 7, sect. 7.3.**

8.1.1. Conception d'un service Web pour la recherche en histoire de l'art : *E.A.T. Datascape*

E.A.T. Datascape (Datascape) est un projet qui associe les recherches en histoire de l'art avec des outils développés dans le cadre des humanités numériques, afin de mettre en lumière l'entreprise de l'organisation *Experiment in Art and Technology* (EAT)[462]. L'objectif du projet est d'interroger les apports des interfaces et du design numériques en ce qui concerne les recherches théoriques et les méthodes en histoire de l'art. Le but est notamment de mettre en évidence de quelle manière la visualisation de données permet d'observer ce qu'il n'est pas possible de voir avec les instruments traditionnels de l'histoire de l'art, mais aussi, de considérer comme ces dispositifs de visualisation peuvent, en retour, modifier les théories et les méthodes de l'histoire de l'art.

L'une des particularités de l'EAT est que cette organisation concerne des dispositifs qui utilisent les nouvelles technologies de l'époque dans le cadre d'une collaboration entre des artistes et des ingénieurs pour réaliser des projets qui sont plus ou moins artistiques. Si l'EAT et les activités de ses membres ont été longtemps mises de côté, d'aucuns remarquent aujourd'hui de plus en plus d'expositions et de conférences à son sujet (cf. Leclercq, 2017). Il s'agit donc d'un cas d'étude idéal pour les recherches en histoire de l'art qui ne s'intéressent pas tant aux relations générales entre l'art et la société dans une perspective sociologique, mais qui s'intéressent plutôt aux pratiques artistiques comme produit d'une collectivité.

Afin d'examiner cette entreprise particulière à travers le prisme de la technologie du numérique et des dispositifs de visualisation de données, l'équipe du projet a développé un service Web[463] qui rassemble les archives de l'organisation pour permettre à des historiens de l'art de mieux comprendre la complexité des données de l'EAT. Le but est de mettre en place un « instrument de recherche » qui rassemble toutes les archives de

[462] À l'origine, E.A.T. a été créée en 1966 (par Robert Rauschenberg, Robert Whitman, Billy Klüver et Fred Waldhauer) pour faciliter la collaboration entre artistes et ingénieurs dans la réalisation de dispositifs artistiques, mais aussi de projets qui s'éloignent des pratiques artistiques à proprement parler. L'objectif est de rendre disponibles la documentation et les archives (conférences, publications, assistance technique, etc.) qui entourent la coopération entre les différents participants d'un projet artistique contemporain, mais aussi de mettre en évidence que c'est justement la coopération qui permet de parvenir à des résultats que les acteurs n'auraient pas pu réaliser séparément.

[463] http://jiminy.medialab.sciences-po.fr/eat_datascape.

l'EAT et, surtout, un « instrument » qui combine ces archives à un processus informatique permettant d'*explorer visuellement* les ressources extraites des archives de l'EAT.

L'équipe affirme que ce projet met en évidence qu'il ne peut y avoir d'histoire sociale de projet complexe sans recourir à des dispositions adéquates qui permettent de conserver à la fois les « données » et les « traces » de sa réalisation (Leclercq, *op. cit.*). L'idée est de restituer ces données et ces traces selon différents points de vue qui fournissent des indices pour comprendre la complexité de l'EAT, tout en évitant les explications simplistes soit selon les rôles de l'artiste ou des techniciens soit par le contexte.

Réciproquement, l'histoire sociale de l'art doit reconsidérer son cadre théorique et ses méthodes lorsque ce type d'instrument est utilisé. Comme nous l'avons remarqué[464], si les chercheurs utilisent des instruments d'analyse qui s'éloignent des outils qui sont habituellement mobilisés par les historiens de l'art, il faut encore prendre en compte l'impact de ces instruments sur les pratiques d'analyse et d'interprétation issues des sciences humaines. Pour les concepteurs de ce projet, la diversité et la richesse des archives d'EAT en font un cas d'étude idéal pour les humanités numériques (*ibid.*). L'élaboration du Datascape montre qu'il est possible d'étudier différemment l'histoire de projets artistiques ou interdisciplinaires qui impliquent plusieurs acteurs aux compétences variées, en rappelant toute l'importance des personnes physiques et morales impliquées dans l'élaboration de ce type de projet. Les différentes fonctionnalités du Datascape visent deux objectifs :

- Visualiser ce qu'il n'est pas possible de voir autrement dans un sujet d'histoire sociale si particulier.
- Considérer la question difficile de la collaboration entre différents acteurs au sein de même ou de diverses activités.

Ainsi, il s'agit non seulement de varier les modèles de visualisations, mais aussi de proposer, au sein d'une même visualisation, différentes perspectives (ou « points de vue ») concernant un même acteur ou une même activité. Pour ce faire, le projet comprend trois outils principaux pour explorer les ressources documentaires qui se rapportent à des *acteurs* et à des *activités* en lien avec des dispositifs artistiques ou projets réalisés dans le contexte de l'EAT. En effet, le service Web comprend trois pages qui

[464] **Chap. 3, sect. 3.1.4.**

proposent chacune une mise en scène qui n'organise pas tant un parcours de lecture encadré par une institution patrimoniale, mais qui propose plutôt différents *points de vue distants* sur les données :

- La page *actor* (« acteur »).
- La page *activities* (« activités »).
- La page *overview* (« vue d'ensemble »).

La page « acteur » rend visuellement compte de la diversité des profils et des rôles relatifs à un projet, notamment à travers des informations descriptives sur un acteur de l'EAT : nom, rôle, mais aussi les activités auxquelles l'acteur a participé. Par exemple, la page « acteur » (Julie Martin) comprend une ligne du temps qui représente la participation de cet acteur à des expositions (en jaune) et à d'autres activités (en gris) depuis 1966 (fig. 87). À titre d'exemple, la page « acteur » comprend une « ligne du temps » (*timeline*) qui représente l'implication de tel ou tel acteur au sein de l'EAT, mais en indiquant seulement les phases des activités auxquelles un acteur est associé. Ce dispositif permet de visualiser d'un seul coup d'œil à quel moment l'acteur a plus ou moins participé à des projets, à des expositions, et avec quels autres acteurs. En outre, l'interface permet de changer l'échelle de la visualisation, afin de présenter à la fois une vue distante et une vue rapprochée.

La page « activité » (cf. fig. 88) permet de penser le processus et le résultat dans l'ensemble, de sorte que les multiples relations entre les informations et les visualisations au sein de la page rendent compte de la complexité et de la durée de vie de la réalisation. Cette page affiche en premier lieu l'identité d'un acteur ou d'une institution (musée, galerie, fondation, etc.) ainsi qu'un qualificatif professionnel (artiste, ingénieur, etc.) et une série de mots-clés relatifs à la description des activités dans lesquelles cet acteur ou cette institution sont impliqués. Cette page porte sur la réalisation en commun d'une activité, avec des informations sur l'activité décrite, son nom, son type (œuvre, projet, conférence, festival, etc.). Ainsi, la page « activités » fait ressortir le nombre de personnes mobilisées et nécessaires à la création d'une activité et à son maintien et, ce faisant, elle évite de présenter le processus créatif comme un acte individuel et solitaire, pour s'intéresser à la façon dont agissent collectivement artistes, ingénieurs et technologies au sein d'un projet. Il faut donc identifier une « activité » (une œuvre d'art, un projet, une

publication ou un service) qui est considérée comme un ensemble de « phases »[465] ou d'« actions »[466]. En outre, une *timeline* présente la « durée de vie » de cette activité (cette « ligne du temps » distingue plusieurs phases, de la conception/production jusqu'à la manifestation) et en dessous de laquelle se trouve la liste des noms des acteurs impliqués dans l'activité.

La page « vue d'ensemble » (cf. fig. 89), quant à elle, est le résultat de l'agrégation des données qui ont été identifiées pendant l'analyse des ressources relatives aux pages « activités » et « acteur ». La visualisation et l'exploration de ces données agrégées permettent alors de dégager des tendances, selon plusieurs perspectives qui redistribuent les données. Dans la partie supérieure de la page, un graphique représente le nombre d'activités et de personnes qui sont impliquées dans une activité en fonction de l'année. Au centre de la page, des nuages de tags concernent des catégories utiles pour la recherche : la liste de ces activités (*activities*) ainsi que leur découpage en phases (*type of activity*), les acteurs (*actors*), leurs profils (*profiles*). Les différentes pratiques artistiques (*art*) et technologiques (*techno*) peuvent être triés soit par ordre alphabétique, soit en fonction du nombre d'acteurs impliqués. Dans la partie droite, une colonne présente la liste des personnes les plus fréquemment enregistrées dans les archives EAT.

[465] Chacune de ces phases est le résultat d'une analyse et d'un découpage des ressources en annotations, ce qui permet d'identifier, dans certains cas, à la fois les acteurs en jeu, de baliser le commencement et la fin d'une phase, de préciser le ou les lieux de l'action.

[466] Par exemple : une action de conception, de production, d'exposition ou de restauration.

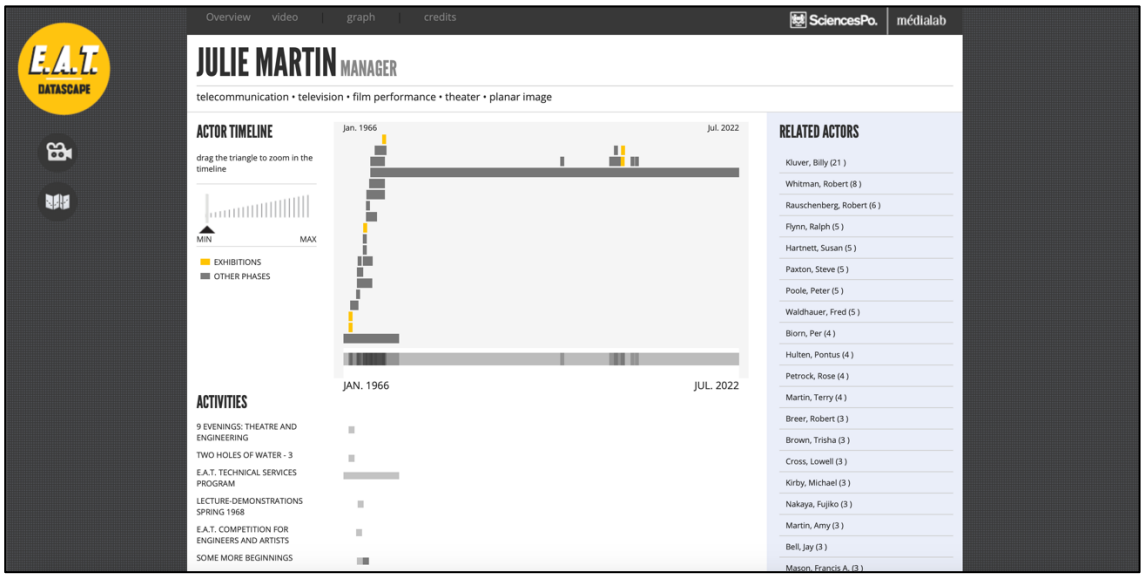


Figure 87 : Page « acteur » ; à droite, une liste d'acteurs (*related actors*) avec lesquels un acteur particulier a été le plus souvent associé, ce qui permet de rendre compte de la coprésence de certains acteurs par rapport à une même activité.

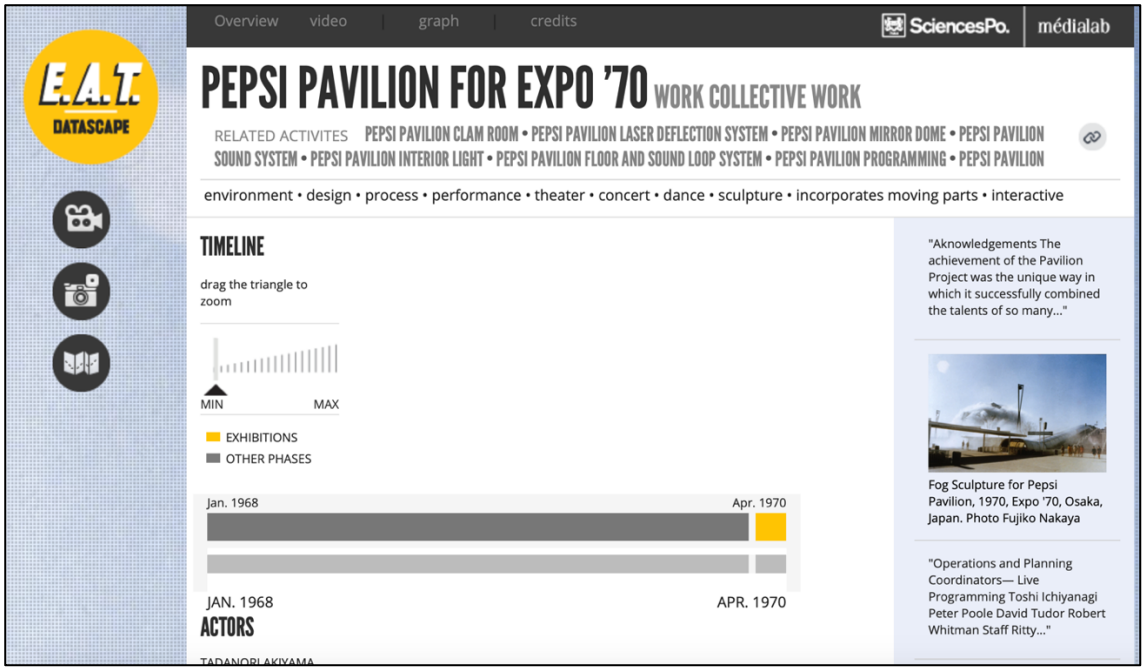


Figure 88 : Exemple de page « activité » (*E.A.T. Datascape*).

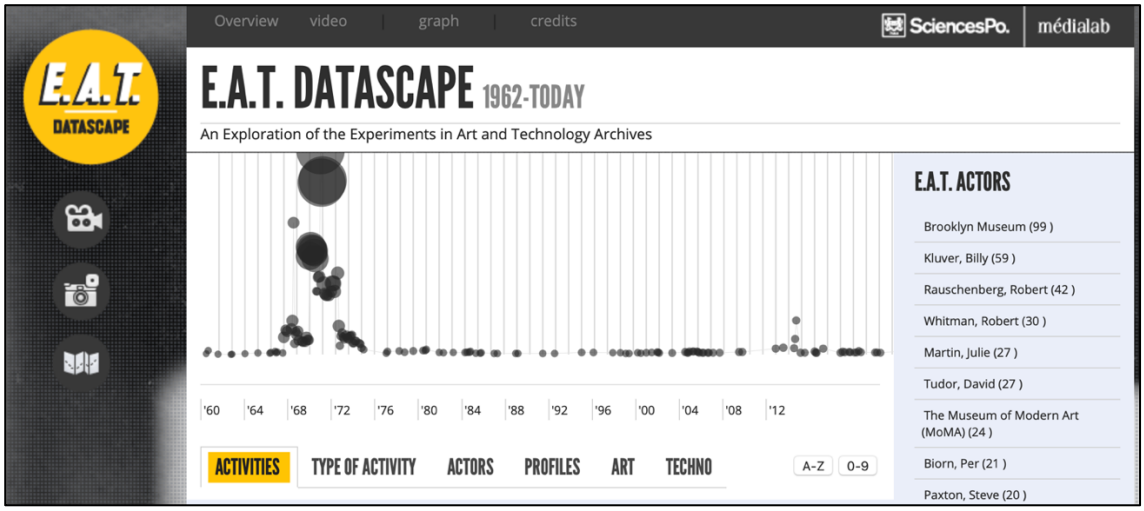


Figure 89 : Vue d'ensemble (*E.A.T. Datascape*) ; dans ce cas, il s'agit d'un diagramme qui représente la quantité d'activité en fonction de la durée (années).

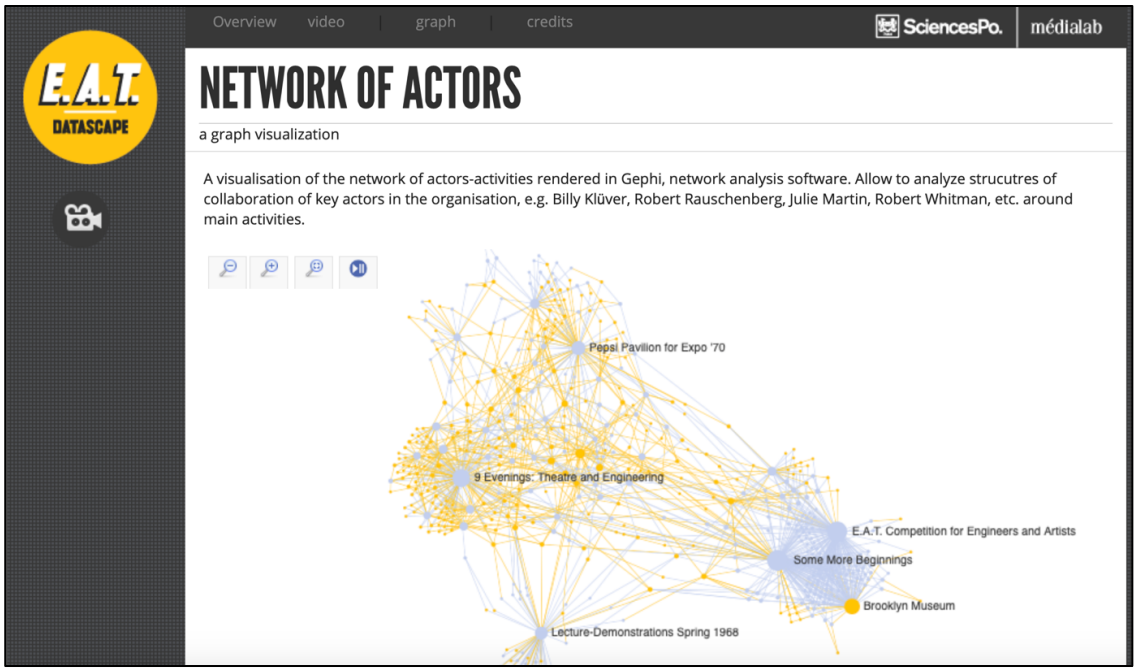


Figure 90 : *Graph visualization* (*E.A.T. Datascape*) ; cette visualisation en réseau distingue plusieurs événements (les gros points avec un titre) et met en évidence les relations entre les acteurs et ces événements. L'utilisateur peut agrandir la visualisation afin d'observer et de suivre ces relations.

Ces trois modèles de visualisation concernent les mêmes informations, de telle sorte que l'interface *multiplie les points de vue* sur les ressources relatives aux acteurs qui ont participé aux activités de l'EAT. D'un côté, les pages « Acteur » et « Activités » permettent une approche plus « particularisante » en présentant en détail les collaborations entre artistes, ingénieurs, etc., et d'autres informations. De l'autre, la page « Vues d'ensemble » restructure et reconfigure toutes ces informations afin de mettre en scène un point de vue plus « distant » sur ces ressources, car nous avons suggéré que l'objectif principal de ce projet est de proposer un service Web qui représente visuellement un réseau d'acteurs et d'activités.

En plus des pages « acteur », « activités » et « vue d'ensemble », le renouvellement à travers une stratégie de visualisation des ressources est observable dans l'application *graph visualization*. Grâce à une visualisation en réseau, cette application représente l'ensemble des relations entre les ressources disponibles sur le site Web ; l'application est dynamique, car elle peut être représentée soit en fonction d'un *point de vue global* soit selon un *point de vue local* (cf. fig. 90). Dans ce graphique, la présentation des résultats d'une recherche sous la forme de listes, de grilles ou de mosaïques est remplacée par des modèles de visualisation plus complexes qui ne correspondent pas à un positionnement de l'utilisateur à travers une « scénarisation » et des « discours de vulgarisation ». Le but de ces modèles est de mettre en place une relation d'énonciation entre l'utilisateur et une « lecture à distance » permettant de construire une multiplicité de points de vue sur des données.

Dans une certaine mesure, ce projet correspond à la visée rhétorique des interfaces de la deuxième tendance, celle qui correspond aux stratégies de l'interactivité[467]. De même, la troisième tendance indique que l'interface Web est conçue en vue d'une utilisation interactive, c'est-à-dire en favorisant la participation active de l'utilisateur pour attribuer du sens au dispositif et aux ressources. Sauf que nous observons que certaines fonctionnalités du Datascape poussent encore plus loin le renouvellement des interfaces Web en mobilisant des techniques issues de la *data visualization*. Dans ce cas-ci, l'objectif est différent. D'une part, le but est de montrer que le traitement informatique fait ressortir des tendances statistiques dans un ensemble de données (« tendances » qui

[467] Cf. les discussions sur les applications du site Web du *Rijksmuseum* et 12 *Sunset* dans le **chap. 7**.

seraient difficilement observables sans ces instruments computationnels). D'autre part, il s'agit de rendre visible *en un seul coup d'œil* un ensemble de données numériques sélectionnées automatiquement par un système informatique.

La multiplication des points de vue (et donc des parcours possibles) se fonde, dans ce cas-ci, sur une conception qui place l'automatisme et le calcul au centre du dispositif de lecture. Néanmoins, un autre problème persiste. En effet, nous avons souligné à plusieurs reprises l'importance du langage verbal dans la conception des dispositifs de communication en ligne[468]. D'une part, les services Web que nous avons présentés fonctionnent comme un dispositif pour valoriser les activités d'une institution ou d'un projet. D'autre part, ces services proposent systématiquement un moteur de recherche en fonction de mots-clés. Dans tous les cas, la conception des interfaces débouche sur un espace multimédia où la navigation est complètement encadrée par une composante discursive. Ainsi, lorsque les institutions patrimoniales produisent un dispositif de lecture, elles le conçoivent en visant une rhétorique unidirectionnelle (l'institution est le détenteur d'un savoir et l'utilisateur est le destinataire de ce savoir).

Par rapport aux bases d'images et aux expositions virtuelles, la principale caractéristique des stratégies de visualisation est de mobiliser le numérique pour faire ressortir *visuellement* des relations entre les données. La mise en scène de la page-écran ne construit pas un lecteur, elle construit un observateur qui est contraint d'interagir avec un dispositif où la composante discursive devient minoritaire. Ce point est important, puisque, comme nous l'avons suggéré, l'utilisation de ces techniques est de plus en plus observable dans les projets qui valorisent des archives et des collections d'images à l'aide du numérique.

Dans la même perspective, nous pouvons mettre le Datascape en relation avec un autre projet qui utilise la visualisation de données, mais dans ce cas pour mettre schématiquement en évidence des relations au sein d'une collection d'images numériques. Ce projet est un autre exemple qui illustre bien comment les techniques de visualisation ont d'abord été utilisées dans le domaine des archives d'images.

[468] Voir les analyses dans la **troisième partie.**

8.1.2. *Object : Photo*

En 2010, la *Fondation Andrew W. Mellon*[469] a accordé une subvention au *Museum of Modern Art* pour étudier la « collection Thomas Walther ». Dans ce contexte, le projet *Object : Photo* a favorisé la collaboration entre des chercheurs, des conservateurs et des restaurateurs qui ont étudié les 341 photographies de la collection[470]. Le but est de concevoir un service Web qui situe ces photographies par rapport à un contexte historique de production, mais surtout d'intégrer dans ce service des instruments qui rendent possibles des points de vue inédits sur ces images. Les résultats du projet sont en partie présentés sur un site Web qui présente la collection de photographies par le moyen de quatre modèles de visualisation interactifs pour explorer la collection d'images, mais aussi pour mettre en évidence des relations entre les photographies, les artistes, les lieux géographiques et les événements, ou encore les matériaux et les techniques qui ont été utilisées.

L'idée est d'utiliser la *data visualization* pour mettre en évidence ces relations par rapport à des lieux géographiques ou à des époques historiques. Ainsi, nous verrons que, dans les exemples suivants, la visualisation est employée pour exprimer visuellement un ensemble des données concernant chaque photographie ou photographe de la collection. Il en résulte des constellations abstraites qui représentent toutefois des relations concrètes entre des individus, des lieux, des événements et des matériaux ; relations qu'il est difficile de mettre en évidence dans un format linéaire ou dans des listes.

Dans la présentation du projet (disponible sur le site Web)[471], les conservatrices Maria Morris Hambourg, Lee Ann Daffner et Mitra Abbaspour remarquent que ces « visualisations mettent en évidence l'étendue géographique et matérielle de la collection ainsi que les réseaux entrelacés qui ont favorisé l'innovation créative de la photographie à cette époque » (nous traduisons de l'anglais). En outre, ces visualisations sont conçues dans le but de proposer à l'utilisateur « la possibilité de poser ses propres questions, de découvrir des relations historiques et matérielles susceptibles d'ouvrir d'autres voies d'investigation et de favoriser ainsi la prochaine génération d'études sur la photographie moderne ». Pour ce faire, ces modèles de visualisation sont « interactifs » dans le sens où

[469] https://mellon.org.

[470] Cette collection comprend des images rares de la photographie européenne des années 1920-1930, mais surtout elle comprend des images variées en termes de technique, de géographie, de genre et de matériaux.

[471] https://www.moma.org/interactives/objectphoto/the_project.html#intro.

les ressources qui sont représentées peuvent être filtrées pour révéler des intersections et des relations entre les photographies, les artistes et les matériaux et techniques qu'ils ont utilisés.

Deux modèles sont assez proches d'une forme de présentation de ressources numériques que nous avons déjà présentée[472] . En effet, le service Web POP propose une application qui utilise une carte afin de représenter la localisation des ressources visuelles accessibles dans les bases d'images numériques liées au service. Dans la même perspective, « *mapping photographs* » (fig. 91) est un dispositif de visualisation de données qui se base sur une carte interactive dont le but est de montrer la portée géographique de la collection Walther en fonction d'une époque déterminée (1900 à 1950). Les photographies sont indiquées par un cercle gris dont la taille reflète le nombre de photographies qui ont été réalisées à cet endroit. Ainsi, les cercles rassemblent les ressources numériques, tandis que des points noirs indiquent des « points de rencontre » qui reliaient les artistes de cette époque. En outre, ce dispositif permet d'obtenir plus d'informations sur ces points de rencontre, notamment en renvoyant directement l'utilisateur vers d'autres pages du site Web qui regroupent ces informations descriptives sur les photographes (éléments biographiques ; moments importants durant leurs carrières artistiques ; ouvrages et essais sur le photographe), mais qui regroupent aussi l'ensemble des œuvres du photographe disponible dans la collection.

D'un autre côté, le dispositif de visualisation « *mapping artists' lives* » propose également une carte interactive, mais dans ce second cas, la visualisation permet de cartographier les événements de la vie d'un seul artiste ou de comparer les événements de la vie de deux artistes (fig. 92). Ce dispositif permet de sélectionner deux artistes pour voir où ils ont vécu, travaillés, exposés, publié et voyagé, mais aussi pour voir où ces deux artistes se sont éventuellement croisés. Différents événements sont indiqués par un cercle de couleur, dont la taille indique le nombre d'événements qui ont eu lieu à cet endroit.

Les deux autres modèles observables sur ce site Web s'éloignent de ce principe de répartition des images ou des artistes de la collection sur une carte. D'un côté, le modèle de visualisation « *comparing photographs* » (fig. 93) propose une ligne du temps interactive qui permet de visualiser et de comparer des photographies en fonction des

[472] **Chap. 5**.

attributs matériels, des techniques de production, mais aussi en fonction des sujets ou des styles. Il est donc possible d'ajouter des « attributs » pour afficher des photographies qui partagent certaines « propriétés ». Le but est de mettre en évidence des matériaux, des techniques, des sujets et des styles dans la collection de photographies, de telle sorte que ce dispositif de visualisation se base sur la sélection automatique des ressources grâce à des filtres, mais le dispositif fonctionne aussi comme un « diagramme »[473]. Dans chaque ligne, les points rouges représentent les photographies de la collection qui partagent l'attribut choisi par l'utilisateur. Lorsque celui-ci clique sur un point rouge, l'interface affiche plus d'informations sur la photographie dans la barre latérale.

Certes, nous avons vu que les bases d'images numériques permettent aussi de filtrer les résultats d'une recherche, surtout pour réduire le nombre de résultats. Sauf que, dans ce cas-ci, il faut remarquer que c'est plutôt la manière de présenter les résultats qui change radicalement le processus d'énonciation qui est en jeu dans l'interface Web : les listes et les mosaïques d'images sélectionnées en fonction de mots-clés sont remplacées par des « représentations quantitatives » qui indiquent visuellement à l'utilisateur des « relations abstraites » à partir du traitement informatique des ressources. En effet, ce projet est conçu afin de favoriser des recherches scientifiques, du moins pour fournir un instrument pour corroborer des recherches en s'appuyant sur le *principe de comparaison*. La mise en scène énonciative de l'interface est conçue, d'une part, pour faciliter cette comparaison entre des images en fonction de leurs caractéristiques ou encore en fonction des années, et d'autre part pour faire ressortir de manière schématique la comparaison entre différentes ressources (cf. fig. 93).

En outre, nous pouvons déjà remarquer que l'un de ces modèles de visualisations ne se limite pas à des *comparaisons linéaires* qui font ressortir des relations de quantité sur une période déterminée, mais ils permettent également des « visualisations en réseaux ». Comme le *graph visualization* du Datascape, le modèle « *connecting artists* » (fig. 94) propose une visualisation qui représente des relations entre des artistes en fonction de différents points de rencontre (par exemple : expositions influentes, publications, écoles, centres industriels et culturels). Chaque point sur le périmètre du cercle peut être associé avec les artistes regroupés au centre du cercle. Lorsque plusieurs

[473] Cf. **chap. 4, sect. 4.2. & infra.**

points sont sélectionnés, le point représentant chaque artiste grandit pour révéler la portée de son influence par rapport à un groupe de lieux et d'événements.

Dans tous les cas, nous constatons que l'intégration des techniques de visualisation dans la communication en ligne a pour fonction de représenter visuellement des relations au sein d'une collection de photographies. Néanmoins, le projet Object : Photo n'a pas complètement dépassé l'influence de la dimension verbale dans la conception de l'interface, car ces relations entre les photographies se basent toujours sur un métalangage de description.

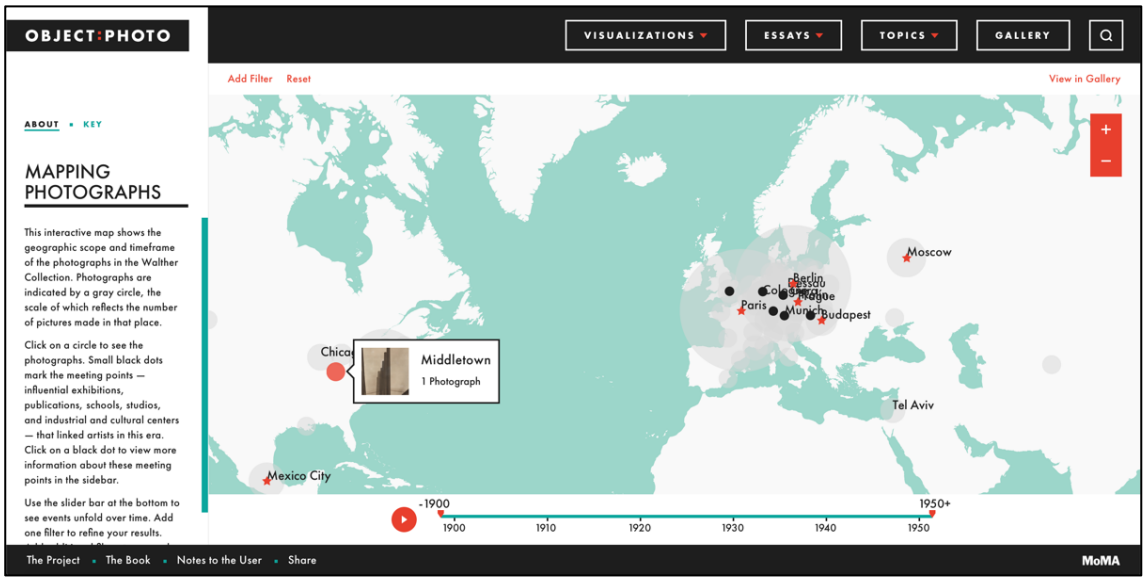


Figure 91 : Application *mapping photographs* ; à gauche, une rubrique informative, à droite, l'outil de cartographie en ligne couplé à un dispositif de visualisation, et en bas, la *timeline* (« ligne du temps »).

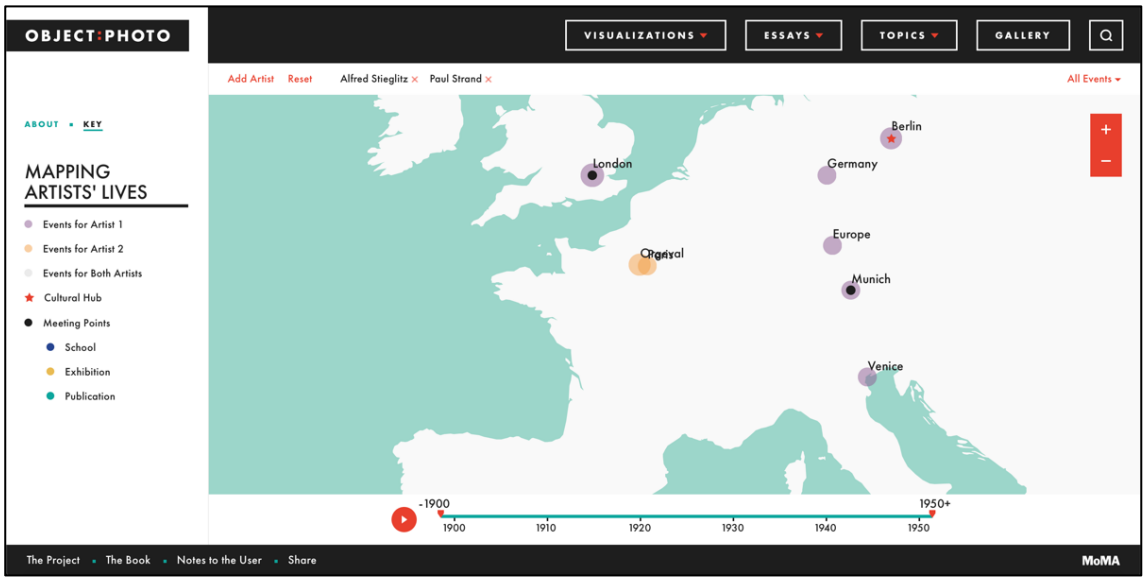


Figure 92 : Application *mapping artists' lives* (à gauche, une légende, à droite, le dispositif de visualisation, en bas, la ligne du temps).

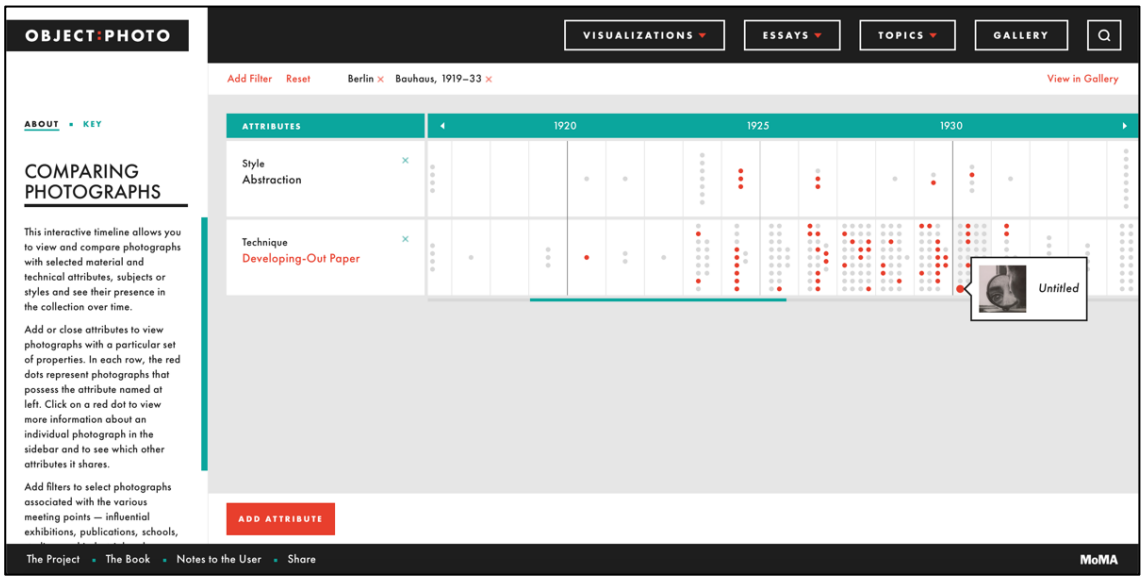


Figure 93 : Application *comparing photographs* (l'outil permet de comparer la quantité de photographies produites en fonction de différentes périodes temporelles).

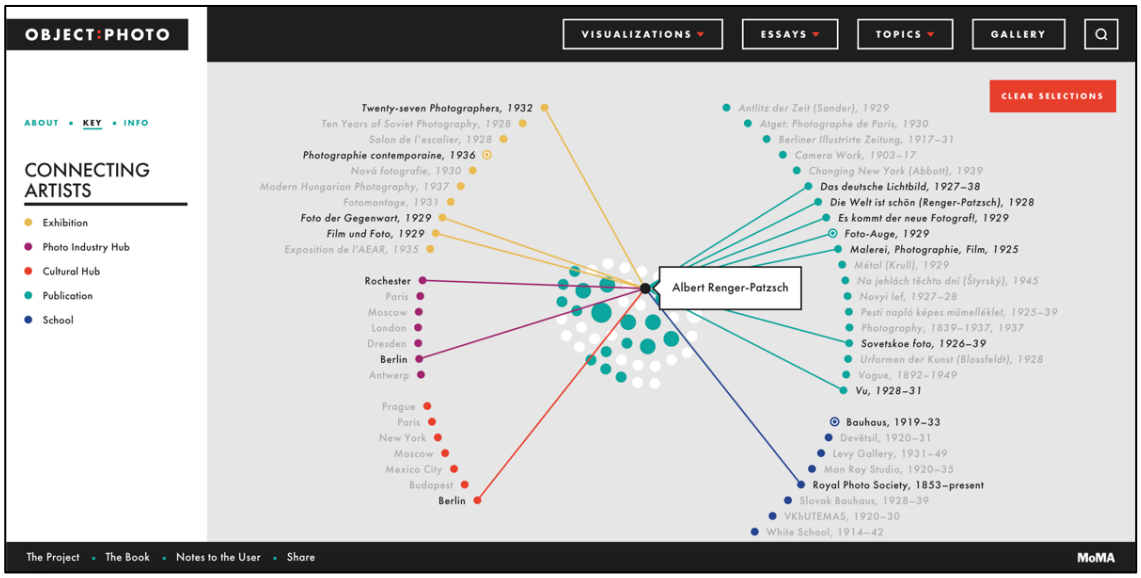


Figure 94 : Application *connecting artists* (cet outil permet de visualiser des relations entre un artiste et d'autres artistes, ou encore entre cet artiste et des lieux, des institutions, des expositions).

Somme toute, l'objectif commun d'*Object : Photo* et de l'*E.A.T. Datascape* est de proposer un service Web où l'utilisateur construit son propre parcours à l'aide des visualisations intégrées dans l'interface Web.

Le passage de la première tendance stratégique à la deuxième, puis de la deuxième à la troisième, montre clairement que le renouvellement des conditions de visibilité et de manipulation se réalise au niveau de l'énonciation des interfaces Web. Ce déplacement stratégique ressort lorsque nous considérons que l'interface est un moyen d'énonciation dont le support perceptible est la page-écran[474]. D'un côté, dans les bases d'images et les expositions virtuelles, les listes et les discours de vulgarisation sont conçus pour *exprimer graphiquement* des « lots de résultats » (des listes que nous pouvons envisager comme des « documents de synthèse » à explorer). De l'autre, les stratégies de visualisation sont conçues pour *exprimer visuellement* les relations dans un ensemble d'images, ce qui permet de présenter ces relations d'une manière nouvelle et non linéaire qui peut ouvrir la voie à une recherche scientifique ; même si la page-écran comprend systématiquement des descriptions verbales, elle s'organise en grande partie sur une mise en scène spatiale simulant un « point de vue distant » sur un ensemble d'images.

[474] Voir aussi les discussions dans le **chap. 7**.

Mieux, ces visualisations débouchent sur une autre forme de scénarisation de l'information, une scénarisation visuelle qui est susceptible de faire découvrir des relations insoupçonnées entre des ressources documentaires. À l'aide de la *plasticité* et de la *spatialité* de la page-écran, la visée de ces modèles est l'énonciation d'un savoir à travers une interface où le positionnement de l'observateur et la comparaison des ressources sont fondamentaux. Cette rhétorique de la visualisation offre à la fois un « point de vue global » (les relations qui sont vues d'un seul coup) et un « point de vue local » (lorsqu'une image particulière est sélectionnée) en fonction des relations qui unissent une *totalité* avec ses *parties*.

Cependant, nous pouvons remarquer que ces techniques de visualisation transforment les images pour qu'elles puissent être visualisées dans des modèles abstraits. Étant donné que la visualisation transforme seulement les images en données représentées par des éléments abstraits, nous verrons que ces projets se sont rapidement orientés vers d'autres dispositifs, car, dans les exemples précédents, l'observateur ne peut pas vérifier la pertinence des rapprochements des images au sein d'une même visualisation (encore une fois, les relations entre les images se fondent surtout sur un métalangage descriptif). Pour pallier ce problème, l'objectif des techniques plus récentes de visualisation (*media visualization*) est de ne pas sacrifier la visibilité des images au profit de leur visualisation. Dans le cas de la *media visualization*, il est plus approprié de parler de *big cultural objects collection* (« artefacts culturels massifs ») ou de *big visual data*, que de considérer qu'il s'agit de *big data* (« données massives »).

8.2. De l'observation à l'expérimentation : *computer vision* et *media visualization*

Les projets qui ont été présentés dans ce travail pourraient laisser penser que le traitement et la classification d'images numériques concernent seulement des images artistiques ou des archives d'images sur le patrimoine culturel. Nous avons toutefois remarqué que les données numériques sur la culture visuelle ne concernent pas seulement des images documentaires et des reproductions d'œuvres d'art.

Dans cette perspective, les sections suivantes montreront que le traitement informatique d'images est notamment appliqué sur d'autres types images[475]. En effet,

[475] Voir aussi **chap. 6, sect. 6.2.**

nous avons suggéré que les institutions patrimoniales sont confrontées à l'émergence d'autres pratiques numériques qui rassemblent et font circuler de nombreuses images, comme les réseaux sociaux, où des milliards d'images sont partagées par des centaines de millions de contributeurs.

De plus, comme nous l'avons suggéré[476], il s'agit aussi de mettre en place des instruments de classification en fonction des caractéristiques plastiques des images. Nous avons également suggéré que ce type d'instrument est utilisé pour mettre en évidence des relations de similarité entre différentes images, tout en manifestant ces images dans la visualisation. Nous présenterons deux projets qui ont permis de tester des modèles de traitement et de visualisation d'images qui proposent des instruments de ce type.

8.2.1. Montrer les performances de l'analyse informatique d'images (photographies, peintures)

Dans les sections suivantes, nous verrons que la *media visualization* a beaucoup impacté le potentiel rhétorique des interfaces qui donnent accès à de larges ensembles d'images. La mise en place de ces techniques de traitement informatique d'images est une étape importante dans l'évolution des méthodes pour étudier des images. En effet, c'est sans doute ce genre d'initiatives qui a popularisé les avantages de l'analyse d'informatique d'images en fonction de leurs caractéristiques visuelles. Nous commencerons par évoquer le travail de Nadav Hochman, Jay Chow et Lev Manovich. Ces derniers ont mis en place des visualisations d'une grande quantité de photographies numériques provenant d'*Instagram* (cf. Hochman & Manovich, 2014).

8.2.2. *Phototrails*

Phototrails est un projet de recherche issu du *Cultural Analytics Lab* qui utilise des techniques de visualisation de médias afin d'explorer le potentiel de ces techniques. À partir d'un échantillon de 2,3 millions de photographies qui proviennent de plusieurs villes du monde entier et qui ont été diffusées sur le réseau social *Instagram*, les visualisations de Phototrails ont pour but de montrer comment les informations (métadonnées) sur les photographies, mais aussi leurs caractéristiques visuelles peuvent révéler des tendances sociales, culturelles et politiques parmi l'activité des utilisateurs

[476] **Chap. 4, sect. 4.2.4.**

d'*Instagram*. En s'appuyant sur les progrès dans le domaine de la classification automatique d'images[477], ce projet a mis en place des techniques de visualisation qui présentent un grand nombre d'images au sein d'une représentation unique. À partir d'un corpus de plusieurs millions de photographies prises à New York, San Fransisco, Tokyo et Bangok, ces modèles de visualisation regroupent les images en fonction des couleurs ou de la luminosité.

L'objectif est de montrer qu'il est possible de faire ressortir des dynamiques temporelles et chromatiques à l'échelle planétaire dans des photographies générées et partagées par les utilisateurs. Ainsi, chaque image est présentée et positionnée par rapport à un ensemble d'images, de telle sorte que le dispositif de visualisation fait ressortir des relations structurelles entre les images, que les auteurs nomment des « signatures visuelles » de ces villes (cf. fig. 95).

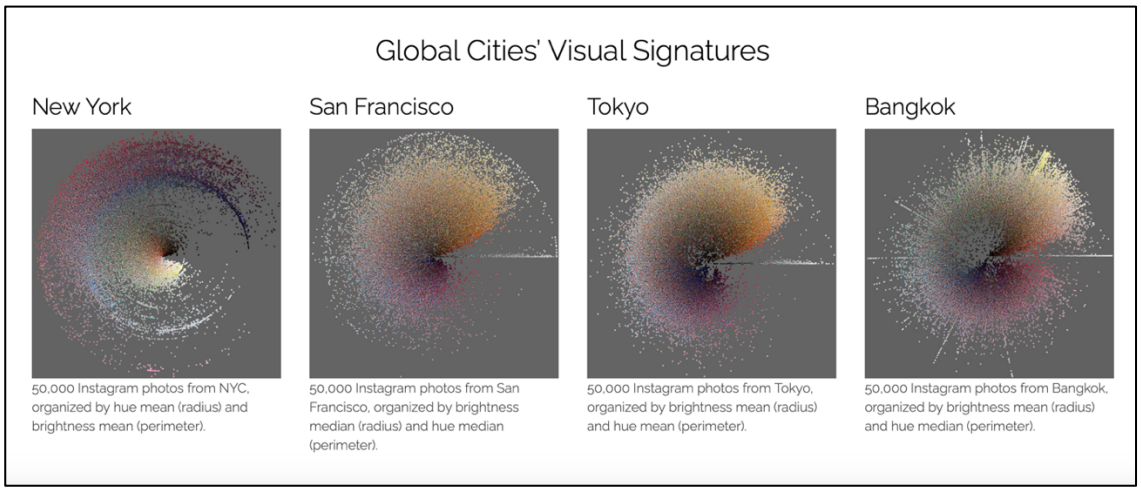


Figure 95 : Quatre exemples de « signatures visuelles » ; dans chaque cas, les images sont organisées en fonction de paramètres colorimétriques ou de luminosité, donc des caractéristiques plastiques des images.

S'il faut remarquer que cette démarche s'intègre dans la continuité d'autres dispositifs de lecture qui se fondent sur une analyse computationnelle, cette démarche apporte un aspect complémentaire à cette analyse : un *effort de synthèse visuelle*. En effet, ces visualisations sont le résultat d'un même traitement, d'une suite d'opérations qui visent à rassembler, à présenter une vue d'ensemble, à faire le point sur un ensemble d'images. Ces visualisations se rattachent au « raisonnement diagrammatique » considéré par Maria Giulia Dondero (2010)[478]. Si un diagramme est une représentation graphique

[477] **Chap. 4**.

[478] **Chap. 4, sect. 4.2.1.**

et schématique des composantes d'un objet complexe, la notion de diagramme concerne aussi « *le processus exploratoire de la pensée* sur un objet qui est rendu saisissable à travers la *construction*, l'*observation* et la *manipulation eidétique d'une visualisation* ».

En effet, l'objectif de la *media visualization* est de proposer un processus exploratoire dans un ensemble de contenus culturels en combinant l'analyse statistique, la *computer vision* et la *data vizualisation*. Si nous avons vu que, dans les modèles de visualisation en réseaux, les données sont présentées sous la forme d'un diagramme composé de *points* et de *lignes*, dans Phototrails, les « lignes » s'effacent au profit de « tensions invisibles » (le calcul de la similarité entre différentes images), tandis que les « points » sont remplacés par des images miniatures situées dans un espace cartésien où les images partageant les mêmes caractéristiques sont regroupées dans les mêmes zones.

Sur ce sujet, il faut évoquer ici les travaux déjà effectués par Dondero (2018b & 2019). L'auteure défend deux idées importantes : (1) la visualisation d'images implique un « geste technologique » ; (2) « le design de ces visualisations est en lui-même une opération métavisuelle ». En effet, les instruments relatifs à la *media visualisation* concernent à la fois un *savoir-faire* et un *savoir-penser*[479], et ils peuvent en outre avoir un objectif scientifique, car les visualisation visent une *analyse diachronique des images à partir des caractéristiques plastiques d'un grand nombre d'images.* Pour l'auteure, trois raisons vont dans ce sens :

- L'organisation des images dans ces visualisations se base sur des descripteurs visuels et non pas verbaux (caractéristiques visuelles).
- La répartition automatique des images est accessible d'un seul coup d'œil, tout en représentant visuellement des relations entre les images en fonction de leurs positions dans un système d'abscisses et d'ordonnées (analyse diachronique).
- Les visualisations peuvent être considérées comme des « images d'images ».

Plus précisément, Dondero distingue deux utilisations réflexives des techniques de visualisation issues du *Cultural Analytics Lab* : (1) les *montages d'images* et (2) les *diagrammes d'images*.

[479] Rappelons que Bruno Bachimont (2010 : 116 & 118) distingue le *savoir-faire* (concernant les instruments qui donnent à percevoir par le moyen de la programmation d'un geste technique) et le *savoir-penser* (qui concerne la programmation de la pensée à travers des contenus). Les instruments sont des inscriptions instrumentales prescrivant des actions dans le monde des choses, tandis que les contenus sont des inscriptions qui prescrivent des actions dans le monde des représentations (*ibid.* : 124).

Les « montages » utilisent des métadonnées ou des informations colorimétriques relatives aux images afin d'organiser les images dans une seule totalité. Soit la classification se base sur la date de production (métadonnées qui sont exclusivement contextuelles ; cf. fig. 97) soit la classification dépend des caractéristiques plastiques propres de l'image, notamment les couleurs (fig. 96). Par rapport à une vision locale (celles des producteurs et des observateurs des images), ce type de visualisations statistiques valorise des tendances qui sont plus facilement remarquables à travers une vision globale et distante. Même si ce dispositif permet également de revenir directement à une observation rapprochée des parties qui composent le montage, ce type de visualisation a surtout pour objectif de montrer les performances de l'informatique dans la classification d'une grande quantité de photographies ; la « lisibilité » de ce type de montage est limitée, parce qu'il ne s'agit pas d'un dispositif d'exploration au sens d'une recherche, mais plutôt d'une synthèse visuelle des *relations*, *contextuelles* ou *plastiques*, au sein d'un ensemble d'images. Dans ce cas, la page-écran est utilisée pour manifester une *image d'analyse d'images*.



Figure 96 : Exemple de montage d'images (Phototrails).



Figure 97 : Exemple de montage d'images (Phototrails ; vue rapprochée).

Les « diagrammes d'images » n'organisent pas la collection d'images en juxtaposant des images, mais utilisent des techniques de visualisation qui représentent les images dans un « espace abstrait » en fonction des caractéristiques plastiques qui ont été extraites des images elles-mêmes ; par exemple, les « diagrammes » d'images groupent les images d'un corpus selon les paramètres de brillance et de saturation (fig. 98 & 99 : vue rapprochée d'un diagramme). Si les montages ordonnancent les images d'après leurs métadonnées (p. ex. : l'auteur, la localisation, la date, etc.) ou leurs caractéristiques plastiques, les diagrammes d'images se basent uniquement sur l'analyse automatique des caractéristiques plastiques des images pour classer et regrouper les images dans un système de coordonnées à deux dimensions, de telle sorte que les images qui partagent les mêmes caractéristiques plastiques sont regroupées dans les mêmes zones du système de coordonnées. Ainsi, les positions des images apportent une caractérisation de leurs « qualités visuelles », tout en mettant les images en relation les unes avec les autres[480].

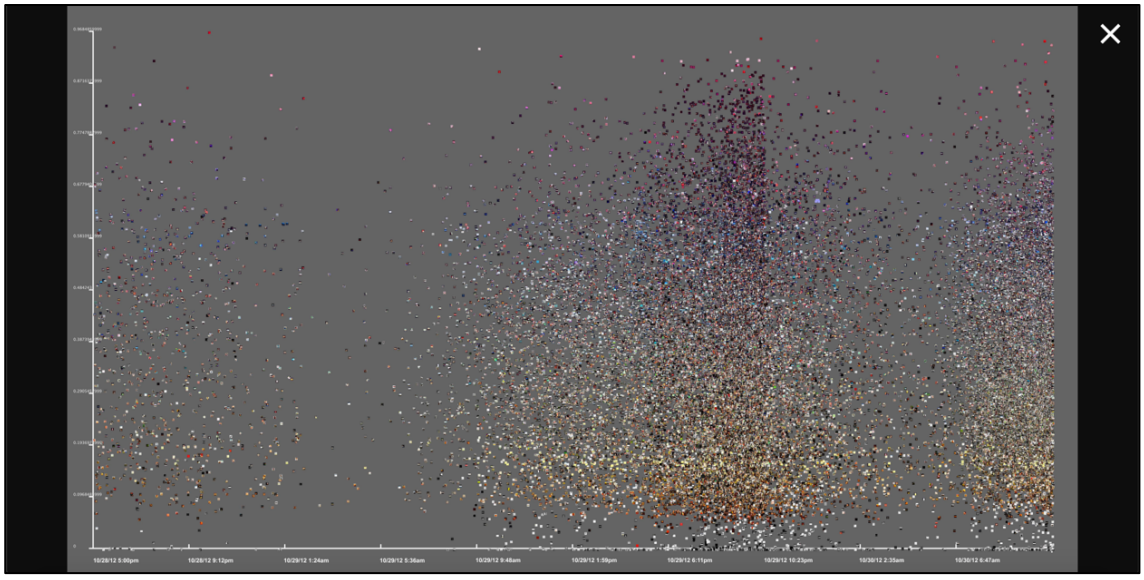


Figure 98 : Exemple de diagramme d'images correspondant à *scatter plot* ou « nuage de points », où les points sont remplacés par des reproductions miniatures des images(Phototrails).

[480] C'est déjà ce qu'évoquait Moles lorsque nous avons présenté les limites de classifications d'images en fonction de mots-clés : envisager la classification d'images en fonction de « matrices de similarité » ; cf. **chap. 1**. Nous avons vu que des approches en histoire de l'art et en sémiotique ont pris acte des possibilités que le support numérique offre pour réaliser ce type de tâches ; cf. **chap. 3**. D'une manière plus générale, l'évolution de certains travaux en informatique sur le traitement et la visualisation d'images indique aussi qu'il est désormais possible de concevoir des processus très performants pour aider les recherches sur le visuel et les images ; cf. **chap. 4**.



Figure 99 : Exemple de diagramme d'images (vue rapprochée).

Nous avons vu que la *media visualization* repose principalement sur deux étapes complémentaires. D'une part, il s'agit d'utiliser les outils de la *computer vision* pour mesurer automatiquement les caractéristiques visuelles des images (teintes ; contraste ; couleurs ; texture ; orientations des lignes et courbes ; présence de visages). D'autre part, il s'agit d'utiliser la *data visualization* pour créer des représentations visuelles qui rassemblent l'ensemble des images traitées, et surtout, qui les classent soit en fonction des métadonnées des images (artiste, date, lieu d'origine, etc.) soit en fonction des propriétés plastiques des images.

Comme l'a récemment remarqué Dondero (2018b, 2019b, 2020), nous pouvons dire que la visualisation automatique d'images en fonction de la similarité visuelle fonctionne comme des « méta-images »[481]. Par rapport aux outils d'indexation « classiques » qui ajoutent des caractéristiques externes aux images à travers un métalangage de description, l'un des avantages de la *media visualization* est de s'appuyer sur la détection des caractéristiques plastiques des images et donc d'être plus adéquats aux « caractéristiques internes » des images. De plus, ces dispositifs se caractérisent par le fait qu'ils expriment visuellement les relations dans un ensemble d'images numériques

[481] Du point de vue de l'auteure, le principe de *métavisuel* permet de distinguer différents types de réflexivité des images : la réflexivité d'une image par rapport à d'autres images (méta-énoncés), la réflexivité d'une image par rapport au support (métalangage), la réflexivité d'une image par rapport aux formes de la perception (métalangue) ou la réflexivité d'une image par rapport à son statut (métastatut) ; cf. **chap. 3, sect. 3.2.3.**

sans remplacer les images par des symboles ou en des figures géométriques (points, ronds, carrés, triangles).

Dans la section suivante, nous présentons rapidement un projet similaire, mais qui concerne la classification automatique d'images artistiques, en l'occurrence des peintures.

8.2.3. *Vy*

Vy est un service Web qui propose un diagramme d'images représentant l'ensemble de la collection de peintures du Musée national de Norvège. Grâce à une programmation et une interface Web, toutes les œuvres sont regroupées dans une seule visualisation en fonction de leurs similarités visuelles. Concernant l'interface Web, remarquons que la présentation est simple et se réduit à quelques phrases explicatives :

> Nous avons entraîné un algorithme pour identifier les similarités sur la base d'attributs tels que le style et le sujet. De plus, nous avons construit cette interface pour montrer comment elle regroupe les œuvres. Nous espérons que cette nouvelle façon de naviguer dans la collection permettra de découvrir plus facilement de nouvelles œuvres (nous traduisons de l'anglais)[482].

Ce projet développé par le Conseil des arts de Norvège est le résultat de la collaboration entre la Section de la gestion des collections numériques du musée, la société *Bengler* et Audun Mathias Øygard. Ce dernier a décrit sa participation à ce projet dans son blog personnel[483]. Ce dernier a été contacté par la société Bengler en 2014 dans le cadre d'un projet financé par le *Art Council of Norway* en collaboration avec le service *Digital Collection Management* du *Nasjonalmuseet*. La proposition initiale du projet était de mettre en place un instrument de reconnaissance des visages dans les œuvres, et l'équipe a progressivement décidé de considérer d'appliquer des techniques issues du *deep learning* à la collection d'images. Comme la plupart des institutions patrimoniales, le musée propose un site Web qui rend accessible une partie de la collection d'images du musée (à peu près 30 000 œuvres). Pour compléter ce premier site Web, l'équipe du projet Vy a développé un second service qui prend la forme d'une interface Web. Le but de Vy

[482] http://vy.nasjonalmuseet.no.

[483] Cf. https://www.auduno.com/2018/10/27/visualizing-an-art-collection/.

est de proposer aux utilisateurs un aperçu global de la collection d'images du musée par le moyen d'un dispositif de visualisation de médias (*media visualization*).

Pour ce faire, un processus de classification d'images qui se base sur des techniques relatives au *deep learning* a été mis en place pour extraire des caractéristiques visuelles dans les œuvres d'art, et ensuite, pour les classer automatiquement dans une visualisation en fonction des similarités visuelles entre les images. Les peintures ont été numérisées, puis un algorithme (t-SNE ou *Stochastic Neighbor Embedding*) les classe automatiquement en fonction des « motifs » (fig. 100) ou des « styles » (fig. 101) ; notons que ces deux notions cruciales (motif et style) ne sont pas définies, mais qu'elle repose entièrement sur le traitement informatique d'images. En effet, le processus informatique a permis de regrouper les peintures de paysage, mais aussi celles représentant des corps humains ou des portraits complets. Dans son blog, Øygard affirme que l'application d'un système de classification automatique à la collection du musée a permis de mettre en évidence des tendances dans la peinture norvégienne au cours d'une période spécifique (vers 1800, la peinture norvégienne représentait surtout des paysages ; vers 1900, les tableaux s'orientent vers le réalisme social ; vers 1950, les peintures sont davantage abstraites).

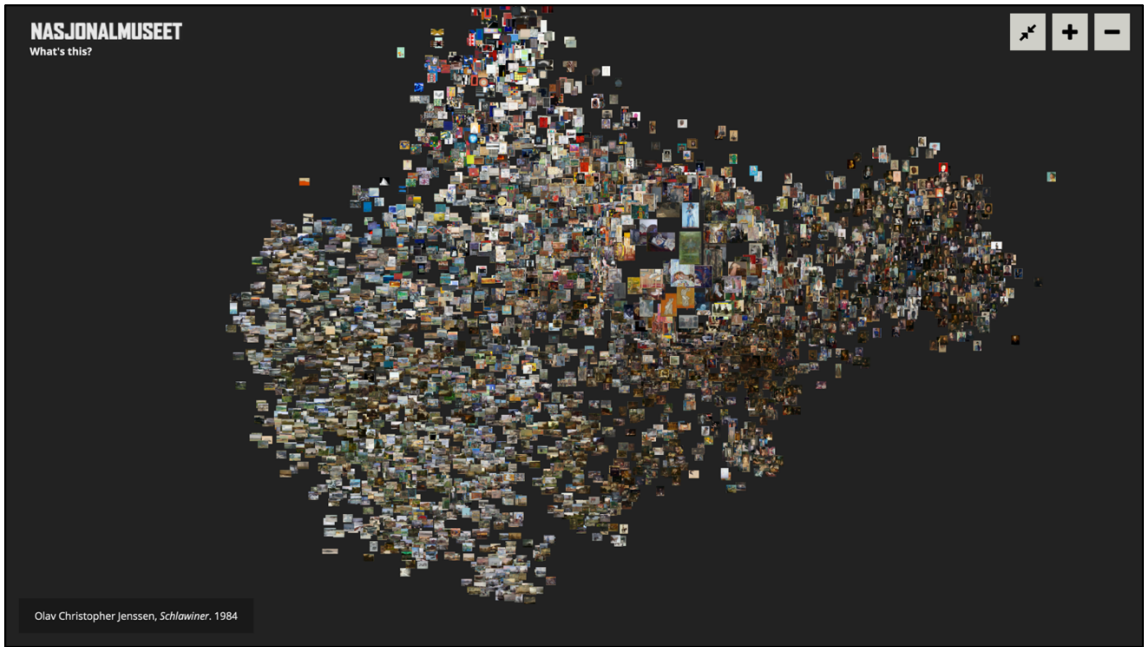


Figure 100 : Application *Vy* (toutes les peintures sont classées en fonction des « motifs » détectés dans les images).

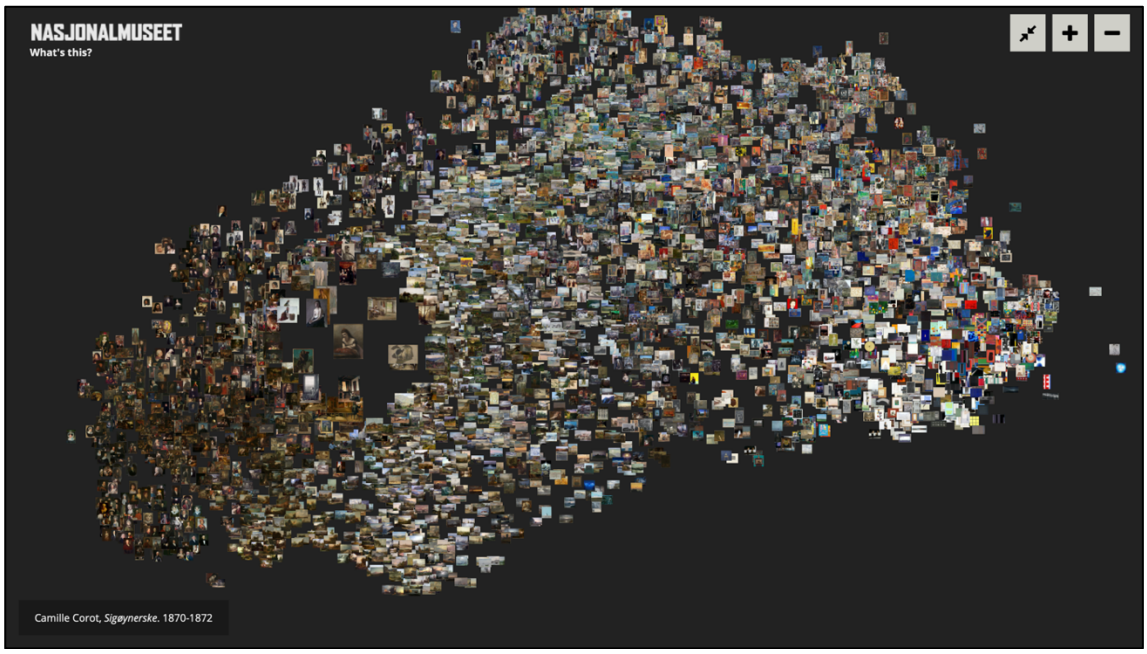


Figure 101 : Application *Vy* (toutes les peintures sont classées en fonction du « style » des images).

Même si cette troisième tendance s'approprie certaines caractéristiques des tendances précédentes, notamment la conception de l'interface comme un espace graphique et multimédia, elle cherche aussi à s'en distancier depuis que les institutions patrimoniales envisagent d'utiliser des techniques avancées de traitement informatique d'images[484]. En effet, la construction du point de vue de l'utilisateur se distingue nettement des deux premières tendances stratégiques, pour au moins deux raisons :

- Les résultats du traitement ne sont pas présentés sous la forme de listes ou de mosaïques d'images, mais sous la forme d'une visualisation d'images avec lequel l'utilisateur peut interagir dans l'espace de la page-écran.
- Dans certains cas, les images ne sont plus uniquement sélectionnées en fonction des descriptions verbales des images, mais aussi en fonction des caractéristiques visuelles des images.

L'exploration à travers une visualisation ne se limite plus à une organisation des images en fonction des métadonnées descriptives. Au contraire, le but est d'accorder une plus grande importance à l'analyse de la dimension plastique des images et à la présentation des résultats de cette analyse plastique. Ce point est crucial, car nous avons

[484] **Chap. 4**.

vu que les chercheurs en sciences humaines ne disposaient pas, jusqu'à très récemment, d'instruments capables d'analyse des larges ensembles d'images en fonction de leurs similarités visuelles. Nous avons aussi montré que différentes disciplines en sciences humaines prônent une ouverture à de nouvelles approches et méthodes d'analyse, et qu'elles défendent aussi un paradigme de la présence et de l'étude contrastée des images à partir du principe de formes visuelles[485]. En outre, nous avons souligné que la mise en présence et la comparaison de peintures sont une démarche fondamentale de l'analyse d'image en histoire de l'art, surtout lorsqu'il s'agit pour les historiens de chercher des relations dans la composition formelle de différentes images[486]. Ces techniques de visualisation de médias vont dans le même sens, sauf que, à cause de la conversion numérique des supports d'images, il est nécessaire de « sacrifier » la *dimension matérielle* en faveur de la seule *dimension plastique des images*.

Pourtant, dans le cas de Vy, que devons-nous comprendre par « classement en fonction des motifs » ou « classement en fonction du style » ? Il est peu probable que l'algorithme utilisé prenne en compte ces deux notions (motifs ; style) dans le même sens que les historiens de l'art. C'est là tout le problème : qu'est-ce que ces relations « disent » du corpus et des images qui le composent ? Même si les visualisations de Vy sont intéressantes, il n'est pas facile d'expliquer les rapprochements formels et plastiques entre les peintures en dehors de critères de sélection assez rudimentaires (teintes, brillance, présence de certaines couleurs). Du reste, ces critères sont rudimentaires parce que, comme nous l'avons vu, les performances de l'analyse informatique d'images sont limitées.

Nous l'avons vu, plusieurs travaux s'accordent sur le fait qu'il existe une multiplicité de méthodes, d'outils et d'instruments qui sont fonctionnels et efficaces pour étudier les images et le visuel. De plus, nous avons montré que les humanités numériques constituent le champ transdisciplinaire qui favorise ce rapprochement entre sciences humaines et technique computationnelle. Il ne s'agit donc pas de trancher entre les approches qui se base sur une « lecture rapprochée » (*close reading*) et ces nouveaux instruments qui offrent une « lecture distante » (*distant reading*). Il faut envisager de

[485] **Chap. 3, sect. 3.1.**

[486] **Chap. 3, sect. 3.4.**

combiner les deux approches pour favoriser des dispositifs de « lectures multi-échelle » (*multi-scale reading*).

8.3. Conclusions

Dans la deuxième tendance, nous avons montré l'importance de la coopération entre les institutions, ainsi qu'une volonté de renouveler les interfaces à travers des applications interactives. Dans le même sens, la troisième tendance vise toujours à concevoir des interfaces pour explorer de façon interactive de larges ensembles d'images, mais elle transforme aussi complètement le *style de visibilité* de ces images. Mieux, nous avons remarqué que la collaboration entre de nombreuses institutions patrimoniales et des programmes de recherches ambitieux change radicalement le cadre de réflexion et les méthodes pour comprendre le visuel.

Dans ce **chapitre**, nous avons présenté différentes stratégies de visualisation qui ont surtout pour but de modifier les modalités d'exploration et d'observation des larges ensembles d'images. À cet égard, nous avons souligné que le diagramme occupe une place centrale dans cette utilisation réflexive du numérique. La notion de diagramme peut aider à comprendre comment l'image et d'autres représentations visuelles sont susceptibles d'être un support de travail dans le but de faire émerger de nouvelles formes pour aider le chercheur. Dans le cas de la visualisation d'un large ensemble d'images numériques, les opérations computationnelles qui sont appliquées sur chaque image sont articulées avec une représentation globale des relations de similarité entre toutes les images. Même si c'est un programme informatique qui réalise l'extraction automatique des caractéristiques visuelles en vue de classer les images à la place d'un humain, les chercheurs en histoire de l'art et en sémiotique peuvent donc utiliser ces techniques à condition de ne pas croire que ces visualisations sont suffisantes pour comprendre un corpus d'images. Certes, l'analyse informatique d'image est un *instrument* qui amorce des changements importants dans notre manière d'analyser et de comprendre les images. Néanmoins, les chercheurs doivent encore s'approprier cet instrument s'ils veulent en tirer plus que des résultats statistiques.

C'est pourquoi le dialogue entre sciences humaines et sciences informatiques est nécessaire. En effet, le présent travail montre que les chercheurs en sciences humaines et les concepteurs de ces instruments d'analyse informatique sont voués à participer

ensemble à la mise en place d'un cadre théorique et méthodologique en vue de construire des pratiques d'usages appropriées à ces techniques de visualisation de médias. Autrement dit : *il faut intégrer ces techniques médiatrices dans une pratique rhétorique qui permettra de comprendre ce que ces visualisations nous « disent » sur la culture visuelle*.

Dans la **première partie**, nous avons remarqué que les historiens de l'art et autres spécialistes du visuel passent souvent beaucoup de temps dans des centres de documentation ou des collections artistiques qui ont progressivement acquis de nombreuses œuvres d'art. Ainsi, l'historien est contraint de se rendre sur place, c'est-à-dire dans des institutions patrimoniales, afin de consulter un grand nombre d'œuvres ou de reproductions photographiques d'œuvres.

Les deux premières tendances stratégiques ont surtout pour but de remédier à cette contrainte (cf. **troisième partie**). En effet, les historiens de l'art peuvent désormais accéder à des ressources visuelles à l'aide de base d'images numériques pour définir un corpus d'œuvres en fonction d'un même auteur ou d'un même genre pictural. L'évolution récente des stratégies indique aussi que la transformation de ces bases d'images en des expositions virtuelles aboutit actuellement sur des dispositifs de communication en ligne où la visualisation des images est centrale dans la construction du point de vue de l'utilisateur. Ces dispositifs construisent un point de vue qui permet de contraster un ensemble d'images dans un espace d'expérimentation qui rend visuellement compte des relations plastiques qui relient ces œuvres. La troisième tendance stratégique débouche ainsi sur un *utilisateur-expérimentateur*. En effet, ces « images d'analyse d'images » sont conçues comme un lieu d'expérimentation qui s'appuie sur une schématisation des propriétés plastiques et formelles des images. Chaque visualisation présente un point de vue à travers un « processus exploratoire schématisé » où le « raisonnement diagrammatique » est compris comme l'observation et la transformation des relations entre les parties et le tout d'un ensemble d'images.

CHAPITRE 9
De la communication en ligne d'images à l'analyse informatique des formes visuelles

Le présent travail a montré comment le traitement informatique a progressivement été appliqué pour organiser et présenter différents larges ensembles d'images.

Dans un premier temps, nous avons vu que les institutions patrimoniales ou d'autres initiatives proposent souvent des fonctionnalités d'exploration qui sont limitées, parce que la recherche se base uniquement sur des mots-clés[487].

Ensuite, nous avons vu que le renouvellement stratégique des bases d'images indique un intérêt croissant d'institutions patrimoniales en ce qui concerne l'amélioration des modalités d'accès aux reproductions numériques de leurs collections artistiques ou de leurs archives d'images, notamment en exploitant plus la spatialité et la dimension visuelle des interfaces Web[488].

Enfin, des exemples plus récents montrent que le traitement informatique d'images et les techniques de visualisation de données sont désormais utilisés par des institutions patrimoniales pour mettre en évidence des relations de similarité visuelle au

[487] **Chap. 5**.

[488] **Chap. 7**.

sein de larges ensembles d'images[489]. Cette classification automatique d'images nécessite toutefois une programmation complexe qui se base sur des langages informatiques et des compétences en programmation très spécifiques.

En effet, nous avons remarqué que des travaux envisagent de réorienter l'iconologie vers l'étude des continuités et des ruptures entre des compositions visuelles produites à différentes époques. À ce propos, nous avons déjà vu que des recherches en histoire de l'art et en sémiotique envisagent d'utiliser l'analyse informatique dans l'étude du visuel et des images[490]. Pourtant, comme nous l'avons remarqué, les méthodes informatiques d'analyse d'images (*image processing* ; *computer vision* ; *data visualization* ; *deep learning*) sont différentes des méthodes qui ont généralement été employées par l'histoire de l'art et les sciences du langage. Une machine ne prend pas en compte les modalités de présence des images ni son contexte de production ou d'usage, elle transforme une image en nombres et analyse uniquement la distribution des pixels dans une image.

Dans le **chapitre précédent**, nous avons remarqué que ces méthodes s'intègrent dans une tendance stratégique dont le but est de concevoir un dispositif de communication en ligne pour visualiser d'un seul coup d'œil un corpus d'archives ou une collection d'images. Cette nouvelle perspective met en avant des services informatiques susceptibles d'accompagner le chercheur en histoire de l'art dans ses analyses, et donc de participer à une pratique interprétative. Nous avons rapidement présenté un service qui utilise le traitement informatique pour détecter des similarités visuelles dans une collection de peintures en fonction de paramètres qui sont appris par la machine. Les résultats de ce traitement dépendent d'un « paramétrage » qui organise l'analyse en fonction d'un choix de paramètres d'analyse déterminant complètement ce que la machine doit et ce qu'elle ne doit pas prendre en compte. L'un des avantages de l'analyse informatique d'images est qu'elle permet d'envisager l'étude quantitative de ces images et surtout de visualiser le résultat de cette analyse à différentes échelles.

Dans ce **dernier chapitre**, nous proposons de présenter en détail un projet qui est en quelque sorte la synthèse de l'évolution des tendances stratégiques que nous avons mises en évidence. En outre, ce projet sera l'occasion de rapprocher ces techniques

[489] **Chap. 8**.

[490] **Chap. 3, sect. 3.1.4. & 3.2.4.**

informatiques des considérations sur l'analyse des formes visuelles que nous avons présentées dans la **deuxième partie**. Dans cette perspective, nous présenterons la conception et les résultats du projet *Replica*. Ce projet a pour objectif de développer un outil de navigation permettant de faire des recherches dans les bases d'images picturales en fonction des caractéristiques visuelles et non seulement d'après des informations textuelles sur les images. Pour ce faire, le projet utilise les avancées dans le traitement informatique d'images afin de mettre en évidence des similitudes formelles entre des images, plutôt que de faire ressortir uniquement des relations sémantiques entre des descriptions d'images.

9.1. *Media visualization* et analyse des formes visuelles

Nous savons que les peintres sont influencés par le travail des autres peintres. Selon les époques, certains chefs-d'œuvre picturaux participent d'un même thème spécifique et emploient des formes qui représentent les points culminants de certains styles visuels. Ces « points culminants » peuvent être considérés comme des motifs (*patterns*) dont la configuration visuelle est essentielle à l'étude de la diffusion de ces formes, ainsi qu'à la compréhension de la genèse d'une œuvre d'art particulière, de sa réception socioculturelle et de son influence à travers l'histoire de l'art.

Mis en place par le *Digital Humanities Laboratory* (DHLAB) de l'École Polytechnique Fédérale de Lausanne, le projet *Replica* concerne deux objectifs principaux (cf. Seguin *et al.*, 2016) :

- La numérisation des peintures et des photographies de la Fondation Cini à Venise, qui comprend plus ou moins deux millions d'images produites entre le XIV^e^ et le début XX^e^ siècles.
- La création de descripteurs et d'un moteur de recherche permettant d'effectuer des recherches dans la base d'images, selon le principe de la « reconnaissance de formes » (*pattern recognition*) qui sont détectées automatiquement dans des peintures d'après une programmation informatique.

La mise en place du projet est présentée en détail dans la thèse de doctorat de Benoît Seguin intitulée *Making large art historical photo archives searchable* (Seguin, 2018). Dans les sections suivantes, nous présenterons les étapes principales qui ont permis de réaliser ce projet qui, selon nous, est un bon exemple pour présenter les

possibilités qu'offrent actuellement les techniques d'analyse informatique d'images (cf. **chap. 4**).

9.1.1. *Replica Project*

Le point de départ de l'ouvrage de Seguin se base sur les mêmes constats que nous avons présentés dans la **première partie** (cf. **chap. 1, sect. 1.1.**). En effet, l'auteur remarque que de nombreux chercheurs en histoire de l'art ont l'habitude de manipuler des reproductions photographiques d'œuvres d'art, notamment pour les comparer, sans devoir se déplacer dans les lieux où se situent ces œuvres (cf. Seguin, 2018 : 4-5). La photographie est considérée comme un dispositif qui a profondément changé les pratiques des historiens de l'art, ainsi que d'autres disciplines scientifiques et pratiques culturelles. Cette reproduction en images du patrimoine débouche sur des archives photographiques qui nécessitent de se rendre dans l'institution (« sur place ») qui les gère et les conserve, lorsque celles-ci sont seulement disponibles en tant qu'objets physiques. Ces archives constituent de par le monde de vastes ensembles de ressources documentaires, parfois rares, à propos des œuvres d'art, bien que ces collections aient parfois été acquises et rassemblées durant une longue période.

Malgré l'effort récent de numérisation de ces collections, Seguin remarque que les institutions se focalisent en priorité sur des reproductions haute définition des œuvres les plus importantes de leurs collections. Il semble par conséquent difficile de rendre compte du large potentiel de ces archives d'images. Pourtant, en plus de comparer la documentation sur leurs contextes de production, rassembler et rendre accessibles ces archives d'images, comme un ensemble d'images à explorer, pourrait faire ressortir des relations formelles entre des œuvres. Comme le précise d'emblée et avec enthousiasme l'auteur : « la richesse des archives photographiques offre des possibilités incroyables, à condition d'être capable d'utiliser efficacement la quantité de données qu'elles représentent » (*ibid.* : 4 ; nous traduisons de l'anglais). Néanmoins, nous avons constaté que l'état actuel de la plupart des archives photographiques ne permet pas d'utiliser et d'exploiter efficacement l'information qu'elles représentent[491].

À ce propos, l'auteur suggère trois problématiques principales :

[491] Cf. **troisième partie.**

- De nombreuses collections ne sont pas encore numérisées. De plus, il s'agit de ressources complexes qui comprennent des informations textuelles et visuelles.
- La fragmentation des données réparties dans diverses archives, puisque chaque institution dispose de ses propres données et informations sur ces données, ainsi que la fragmentation des données au sein d'une même archive, puisque celle-ci rassemble des agrégats de documents qui proviennent de différents individus ou groupes.
- Même si les images sont numérisées, la recherche dans les collections se base souvent sur des métadonnées textuelles, ce qui suppose que chaque image est correctement décrite et indexée dans l'ensemble, même si cette description et cette indexation sont fondées sur des « jugements subjectifs ».

Pour pallier ces difficultés et problèmes, l'hypothèse de l'auteur est que l'analyse informatique permet d'extraire des informations visuelles et textuelles et d'indexer de larges ensembles d'images en fonction de cette extraction. En effet, nous avons vu que des travaux en *computer vision* proposent actuellement des instruments de classification d'images qui se basent sur des techniques automatisées d'apprentissage, et surtout que ces instruments réalisent de bonnes performances lorsqu'il s'agit de traiter automatiquement un large ensemble d'images numériques.

Ainsi, la numérisation intensive et l'indexation d'images numériques posent des problèmes que le projet Replica tente de résoudre à l'aide des avancées en *computer vision*. Selon Seguin, « la tâche de la recherche d'images dans ce domaine est sous-évaluée » (*ibid.* : 5), parce que la recherche par le moyen d'éléments formels des images est problématique à cause de nombreuses variations de supports (dont la qualité est elle aussi variable), mais aussi des variations de textures d'une image à l'autre. Selon Seguin, il est possible de développer des instruments de recherche plus performants pour explorer des bases d'images numériques. En se basant sur les acquis récents dans le domaine du traitement informatique d'images, il s'agit de défendre l'idée que les techniques de l'analyse informatique sont désormais assez développées pour classer un large ensemble d'images en fonction d'une similarité visuelle entre les images. En simplifiant, le but de Seguin est de répondre à deux questions fondamentales :

- Comment rendre accessibles les ressources d'une archive photographique ?
- Comment aider les historiens de l'art à explorer, à naviguer à travers de larges ensembles d'images ?

Pour répondre à ces deux questions, l'auteur a développé, d'une part, une *méthode* et un *pipeline* pour traiter les données, et d'autre part, un *instrument* pour explorer un grand nombre d'archives d'images en fonction des *caractéristiques plastiques* et de la *similarité visuelles* de ces images. En outre, l'auteur veut diffuser cet instrument sur Internet pour que de nombreux usagers puissent interagir avec les reproductions numériques à travers un dispositif de visualisation d'images. Dans la même perspective que le dernier service présenté dans le **chapitre précédent**, il est attendu que l'interface Web s'appuie sur une stratégie de visualisation qui représente des archives images en fonction de relations formelles et visuelles entre ces images.

9.1.2. Première étape : définir un cadre de référence

La première étape est de se procurer une base d'images où les images sont regroupées selon les critères « classiques » de l'histoire de l'art (description verbale ; contexte). Pour Seguin, la diffusion des caractéristiques visuelles est remarquable dans la peinture de la Renaissance, entre le XIVe et le XIXe siècles. Seguin choisi la *Web Gallery of Art* (WGA) qui propose certes une « petite » collection de peintures (à peu près 28 000), mais dont la plupart des œuvres correspondent à la période et à l'époque envisagée.

La première étape consiste également à fournir un cadre de référence théorique sur les techniques de traitement informatique qui sont utilisées dans ce projet (cf. Seguin, 2018 : 9-18). L'un des principaux objectifs de la *computer vision* est de permettre d'analyser des images à l'aide d'ordinateurs. Par exemple, il s'agit de localiser ou de classer des images selon les objets qu'elles représentent, de reconnaître des visages, mais aussi des lettres, des symboles, etc. : autrement dit, considérer *généralement* les images numériques comme des matrices de pixels. Ces matrices sont généralement converties en une représentation de caractéristiques (*feature representation*), c'est-à-dire un vecteur de moindre dimension que le nombre de pixels dans une image, mais une quantification qui permet toutefois une bonne représentation de l'information visuelle présente dans l'image. (Ce sont ces *feature reprensations* qui sont utilisées en « traitement automatique d'images », à la place des images numériques ; l'auteur recense plusieurs caractéristiques de ce type de traitement :

- *GIST descriptor*, qui divise une image en fonction d'une grille régulière, et calcule un histogramme pour chaque cellule puis rassemble toutes les valeurs au sein d'un seul vecteur représentant globalement l'image.

- *HoG descriptor*, qui représente une image selon des gradients dans les histogrammes.
- *SIFT algorithm*, proposant un processus complet, d'abord, pour détecter des points d'intérêts stables, auxquels correspondent des échelles et des orientations, ensuite, pour calculer un « descripteur local » pour chaque point d'intérêt.)

D'un autre côté, l'auteur remarque que le *deep learning* est devenu l'approche dominante dans de nombreux domaines de l'apprentissage automatique et de la *computer vision*. Le système le plus important issu de cette approche est sans doute les *ConvNet* qui, comme nous l'avons vu, reposent sur un ensemble de « couches convulationnelles ». La concaténation de ces couches permet d'appliquer des filtres sur une image en fonction de plusieurs millions de paramètres, à condition que ces réseaux soient entraînés au préalable. Des cartes de caractéristiques (*features maps*) sont successivement calculées par rapport à une image qui correspond à un « input ».

Comme nous l'avons remarqué, un *ConvNet* traite et analyse des représentations de caractéristiques (*features represantation*) des images selon plusieurs niveaux intermédiaires jusqu'à un résultat final (cf. Seguin, 2018 : 13)[492]. Si ces techniques informatiques sont déjà utilisées depuis la fin des années 1980, leurs applications dans le traitement automatique d'images sont plus récentes. Nous avons vu que plusieurs travaux ont étudié la puissance des caractéristiques de ces réseaux en les appliquant à de larges ensembles d'images numériques[493] . Selon l'auteur, on peut remarquer trois raisons principales à cet engouement : (1) l'augmentation de la taille des bases de données d'images, (2) l'évolution de la puissance des ordinateurs, (3) l'amélioration des algorithmes pour programmer les opérations.

Plus particulièrement, l'auteur s'intéresse à l'application des méthodes de traitement aux images relatives à l'histoire de l'art. Mais, selon lui, les applications de la *computer vision* aux images d'art concernent souvent « des tentatives de traduction directe de tâches typiques de computer vision dans le domaine de l'art » (Seguin, 2018 : 14). D'aucuns ont donc tenté d'appliquer à des collections d'images artistiques des systèmes automatiques de classification ou de détection d'objets, en les reliant

[492] Voir aussi la discussion sur le *deep learning* dans le **chap. 4**.

[493] L'auteur renvoie, entre autres, aux travaux suivants : Deng *et al.* (2009) ; Simonyan & Zisserman (2014) ; He *et al.* (2016) ; Oquab *et al.* (2014).

notamment à des bases de données concernant l'histoire de l'art. Dans cette perspective, les *ConvNet* sont utilisés pour calculer un descripteur global pour chaque image d'une collection numérisée, et donc, pour une image donnée, son descripteur peut être comparé aux descripteurs des autres images de la collection, et ainsi de suite.

9.1.3. Deuxième étape : de l'archive physique à l'archive numérique

Dans cette partie (Seguin, 2018 : 19-56), l'auteur considère le problème de la numérisation efficace de larges collections de photographies. Avant le développement du réseau Internet, ce type d'archives d'images rassemblait les principales reproductions d'œuvres d'art, ce qui évite de nombreux déplacements à ceux qui voulaient ces ressources. Ces collections photographiques comprennent une grande variété de formats et de structures (dont des images, des textes manuscrits et imprimés) qui, selon l'auteur, dépendent généralement des méthodes et des instruments d'organisation adoptés par les institutions qui les gèrent (Seguin, 2018 : 19). À titre d'exemple, mais aussi de cas d'étude, il propose donc de considérer la collection photographique archivée à la Fondation Cini, et qui est le résultat d'une « agrégation » de différents fonds d'archives d'images.

D'emblée, on peut mettre en évidence trois obstacles à la numérisation de cette collection d'images photographiques : (1) l'hétérogénéité au sein de la collection, notamment en ce qui concerne les images, et l'absence de standardisation quant à la taille ou à la forme du support des images ; (2) la nature « double-face » de ces ressources, notamment en tant qu'outils de travail pour des historiens de l'art, qui les annotaient et inscrivaient des informations complémentaires sur ou en marge de ces images ; (3) le débit (ou la puissance) informatique de traitement pour réaliser une telle entreprise (numériser plusieurs milliers d'images doit, en quelque sorte, rester une expérience raisonnable dans la durée)[494].

[494] Ainsi, pour réduire l'entreprise, ce projet de numérisation se concentre sur une partie définie de la collection : les *schedoni*, des cartons standardisés sur lesquels sont collées les photographies et sur lesquels sont inscrites les métadonnées (auteur, lieux, institution, date, description, etc.).

En effet, pour réaliser ce projet de numérisation, il a été nécessaire de concevoir un instrument unique[495]. Il s'agit d'un dispositif, une table circulaire et rotative, comprenant quatre plaques pour déposer les images à numériser. La rotation est contrôlée par un moteur à vitesse variable, et les plaques sont conçues pour accueillir divers formats (tailles) d'images. Une équipe de deux personnes participent à la numérisation automatique des images (l'une pose les images, l'autre les enlève) ; en outre, un système de capteurs détecte la position du document sur la plaque. Cet instrument de numérisation permet aussi de numériser le recto et le verso des documents, car des dispositifs de numérisation sont placés au-dessus et en dessous de la table[496]. Concrètement, ce processus de numérisation a débuté en juillet 2016 et s'est terminé en août 2017 (13 mois), processus durant lequel 337 000 documents sont numérisés, pour aboutir sur à peu près 50 TB de données brutes.

D'un autre côté, ce projet a mis en place un *pipeline* pour extraire automatiquement les informations contenues dans ces reproductions numériques. Étant donné la quantité de données, un tel système automatisé est nécessaire ; ce processus est d'ailleurs avantagé par le fait que la plupart des informations sont déjà structurées dans les documents originaux[497]. La première étape concerne une tâche de conversion du « format brut » (*raw files*) au format JPEG, pour réduire la taille de ces données. Ensuite, il s'agit d'extraire automatiquement les différentes parties de l'image : pour ce faire, l'auteur utilise un système de segmentation automatique basée sur une architecture computationnelle de type *ConvNet*. À une image numérique donnée, le système est capable de prédire, pour chaque pixel, si celui-ci appartient à l'arrière-plan, au papier cartonné ou à l'image. Ce système se base sur (plus précisément, rassemble des parties de) deux différentes architectures (U-Net et Resnet)[498] qui sont souvent utilisées en segmentation d'images et qui compilent plusieurs paramètres du système de codage pré-entraîné sur d'autres collections d'images. Ces architectures computationnelles

[495] Cet appareil a été commandé à l'entreprise *Factum Arte*, installée à Madrid, et qui a pour but de concevoir des instruments techniques pour améliorer la conservation du patrimoine culturel ; cf. http://www.factum-arte.com/.

[496] Les appareils produisent des reproductions de haute qualité (5424 x 3616 pixels) ; par conséquent, les données produites sont importantes (120 MB par document numérisé).

[497] Pour l'auteur, il y a trois parties structurées qui sont pertinentes dans ces documents : (1) le papier carton, (2) l'image photographique, (3) les métadonnées textuelles.

[498] Cf. He *et al.* (2015) et Ronneberger, Fischer & Brox (2015).

constituent un réseau, auquel il s'agit de soumettre les images numérisées de la collection, pour l'entraîner à prédire la probabilité d'appartenance de chaque pixel à l'une des trois parties (arrière-plan, carton, image). Cependant, pour transformer ces prédictions en « régions visuelles », une opération « manuelle » est nécessaire, afin d'éviter de confondre « l'arrière-plan dans une photographie » de « l'arrière-plan dans la numérisation ».

Après avoir séparé les éléments visuels dans ces documents, l'étape suivante est d'extraire et de lire les métadonnées textuelles, qui sont systématiquement situées dans un cadre tabulaire, où chaque cellule contient quelques mots à propos de la photographie. Si ces informations sont disposées de façon systématique, elles varient toutefois d'un document à l'autre, comme il s'agit parfois d'écriture manuscrite ou imprimée ; c'est pourquoi ce projet se centre surtout sur les inscriptions imprimées, quoique l'auteur discute la possibilité d'améliorer et de généraliser cette approche pour l'appliquer aux inscriptions manuscrites.

Un logiciel de reconnaissance optique de caractères (*Google Vision OCR*) est utilisé et a produit des listes de mots à partir d'éléments textuels extraits dans la numérisation : pour obtenir une information structurée, chaque élément textuel extrait doit être affecté à une « étiquette de métadonnée » (en pratique, la position de ces métadonnées étant stable, l'extraction peut facilement faire correspondre les données extraites à des catégories de métadonnées déjà présentes sur le carton original ; cf. fig. 102 ; cf. Seguin, 2018 : 28)[499]. La performance de ce logiciel a été très bonne en ce qui concerne cette tâche : le taux d'erreur étant de l'ordre de 3% ; selon l'auteur, si cette qualité du système d'extraction automatique des métadonnées est acceptable, l'on peut néanmoins s'interroger sur ce « taux d'erreur », qui n'est pas forcément acceptable, par exemple, pour historien de l'art.

[499] Les métadonnées sont catégorisées comme ceci : *Location, Author, Fondo' Stamp, Institution, Description, Photographer, Identifiers, Date.*

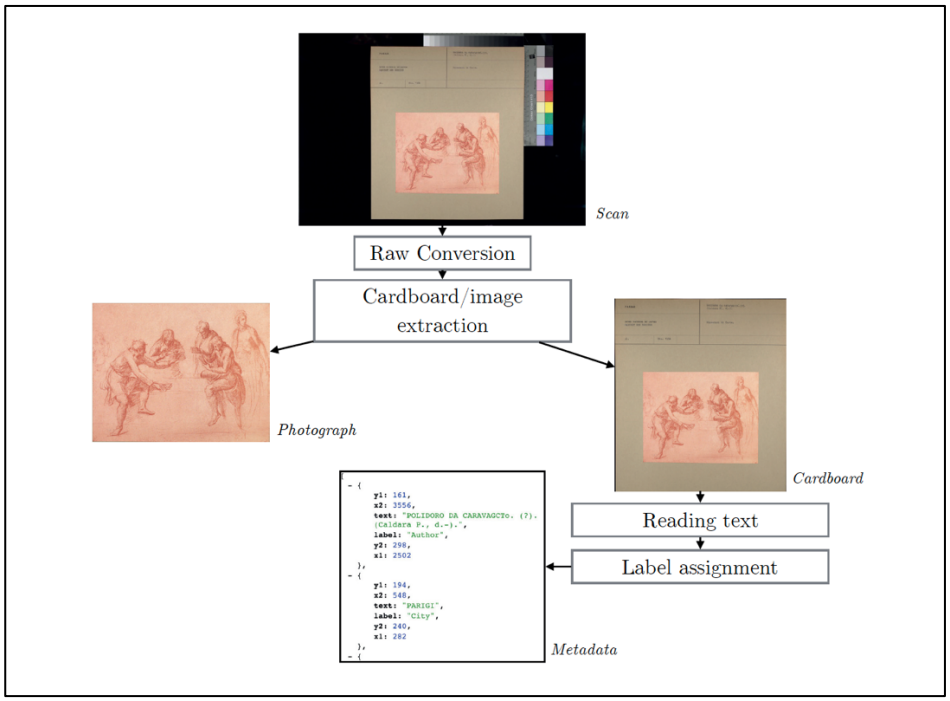


Figure 102 : Schéma de Seguin représentant les différentes opérations du *pipeline* permettant d'extraire les éléments visuels et textuels des documents numérisés (cf. Seguin, 2018 : 28).

Ce processus (extraction des éléments visuels et textuels) a par la suite été appliqué, après l'entraînement des réseaux sur un échantillon d'images de la collection, aux 330 000 reproductions numériques provenant de la Fondation Cini. Le traitement informatique a duré une semaine et a nécessité un ordinateur puissant, du fait de la haute résolution des images. À la fin du traitement, chaque image numérique a été convertie en trois éléments : (1) une reproduction complète du document original, (2) une image recadrée de la photographie et (3) un dictionnaire de mots-clés représentant les métadonnées des documents.

Ensuite, la dernière étape de ce *pipeline* consiste à relier l'information extraite avec une base de données. Pour ce faire, l'auteur a utilisé des bases de données disponibles en ligne, et qui correspondent à la thématique de la collection (peinture italienne) : notamment *Union List of Artist Names* et *Wikidata*[500]. Cependant, la variation des noms d'auteurs au sein de ces ressources documentaires et les connaissances implicites et contextuelles qui permettent de différencier ces noms rendent difficile l'homogénéisation des métadonnées. Afin de pallier ce problème, l'auteur propose un processus complémentaire dont le but est d'utiliser un dictionnaire (un thésaurus) pour

[500] *Union List of Artist Names* : https://www.getty.edu/research/tools/vocabularies/ulan/ ; *Wikidata* : https://www.wikidata.org/wiki/Wikidata:Main_Page.

relier les différents noms à leurs entités correspondantes, mais aussi pour mettre en place, à partir de ces correspondances, un second dictionnaire. Les noms sans correspondances à la suite de la première phase sont confrontés à ce second dictionnaire durant une deuxième phase de traitement.

Ce processus complémentaire fonctionne relativement bien selon l'auteur (c'est-à-dire, pour à peu près les trois quarts des données). S'il souligne qu'une « analyse correcte de la collection de photographies n'entre pas dans le cadre de ce projet », Seguin remarque cependant qu'« avec 74% des affectations d'images accomplies, nous sommes dans une position unique pour examiner pour la première fois de manière quantitative le contenu global de la collection » (cf. Seguin, 2018 : 45). D'une manière générale, cette étape a donc permis de représenter quantitativement des « tendances » à travers l'espace et le temps, ou selon les artistes, les contextes de production ou de traitement de larges ensembles d'images. Puisque de larges ensembles d'images font l'objet d'une numérisation intensive, on remarque donc de nouveaux usages possibles. Cependant, la plupart des instruments d'indexation ou de recherche dans les institutions se basent sur des métadonnées textuelles, des « étiquettes » (*tags*), et dans ce chapitre, l'auteur propose au contraire d'adopter une approche qui s'intéresse à l'information visuelle dans ces documents. Grâce aux méthodes du *deep learning* et de la *computer vision*, il envisage la possibilité de mettre en place un processus de classification et d'organisation d'images à partir de leurs contenus visuels.

9.1.4. Troisième étape : organiser l'information visuelle

Cette partie du travail de Seguin (2018 : 55-102) vise à pallier deux problèmes spécifiques :

- La recherche d'images qui se base sur des similarités visuelles peut dépendre d'une ressemblance des couleurs et des formes des éléments représentés, ainsi que du style de l'œuvre d'art, voire de son contenu d'un point de vue sémantique.
- Rassembler plusieurs archives d'images et bases de données implique que différentes images correspondent au même objet, comme cela implique que des images reproduisent seulement des parties d'un même objet, et donc que ces reproductions seront considérées comme des images différentes.

Dans un premier temps, l'auteur propose une formalisation théorique qui, pour une collection d'images donnée, permet de représenter la complexité des informations sous la forme d'un graphique qui représente des rapports, les « connexions visuelles », entre les images (Seguin, 2018 : 62-68). La détection de la similarité entre deux images est pour l'auteur un concept généralisable : l'idée est de mettre en évidence automatiquement la propagation de motifs visuels, de *patterns* à travers de larges collections d'images. Cependant, il faut remarquer une certaine diversité dans la manière dont deux images peuvent apparaître « connectées ».

Pour l'auteur, deux raisons principales expliquent cette diversité dans les connexions entre images. D'une part, des images peuvent réutiliser une même composition globale, quoique cette composition est différente en ce qui concerne des éléments locaux. D'autre part, des images réutilisent parfois un élément similaire, tandis que le reste de la composition est modifiée. Seguin rappelle néanmoins que, dans les pratiques, « les choses ne sont pas aussi tranchées, car la relation visuelle entre deux œuvres d'art est souvent un mélange d'éléments globaux et locaux » (Seguin, 2018 : 62). Pour remédier à ce problème, l'auteur propose donc une formulation mathématique de la similarité visuelle dans un ensemble d'images, reposant sur plusieurs définitions qui mettent en évidence trois propriétés principales ; nous illustrons les formules telles quelles, et dans la langue originale.

Le but de ces définitions est de mettre en évidence que pour certaines paires de connexions visuelles dans une collection d'images, il est possible de déterminer si l'une ou l'autre de ces paires est plus « forte » qu'une autre. D'un point de vue mathématique, cela se traduit par un « ordre partiel strict » (*strict partial order*) pour un ensemble d'images donné, et selon l'auteur, on peut rapprocher cet ordre au jugement subjectif d'une personne qui devrait effectuer cette tâche (comparer la similarité visuelle au sein d'un ensemble d'images). En outre, l'auteur propose une formalisation mathématique pour standardiser les graphiques qui représentent les relations visuelles, les connexions entre les images : en ce sens qu'un graphique stable est un graphique dont tous les éléments sont stables, et ce que l'auteur appelle les *monographs* sont les classes de graphiques possibles pour une collection d'images.

Selon l'auteur, en effet, l'utilisation de graphiques pour représenter les relations de similarité visuelle présente au moins trois avantages : (1) il s'agit d'une structure de représentation de données compréhensible à la fois par un être humain et par un

ordinateur ; (2) ces structures sont facilement modifiables, annotables et manipulables par des utilisateurs ; (3) les graphiques représentent visuellement les informations afin de les comparer et de les diffuser au plus grand nombre par le moyen de sites Web. D'un autre côté, il souligne que la mise en place d'un système de recherche et d'exploration qui se base sur des graphiques représentant des collections d'images est un processus complexe, que son approche tente de résoudre, au moins partiellement ; effectivement, la suite du chapitre explique comment coupler l'utilisation de graphiques pour organiser la collection d'images avec un puissant système de recherche et de visualisation qui se base sur les méthodes du *deep learning* et de la *computer vision*.

Le cadre de référence du processus proposé comprend trois composantes : (1) le graphique des connexions et annotations réalisées par les utilisateurs (*morphograph*) ; (2) des techniques de vision par ordinateurs, qui apprennent à organiser les images sur la base des graphiques ; (3) une interface de recherche qui permet aux utilisateurs de naviguer à travers la collection et d'annoter les graphiques. De sorte que ce processus crée une « boucle rétroactive » : plus il y a de connexions dans le graphique, plus le modèle visuel de représentation sera amélioré grâce à l'apprentissage automatique, et plus la performance du moteur de recherche sera performante pour mettre en évidence de nouvelles connexions, et améliorer davantage le modèle visuel, etc.

Lorsqu'il s'agit de détecter des photographies d'une même œuvre d'art, l'auteur remarque que c'est la « cohérence spatiale » qui est le meilleur indicateur pour prédire la ressemblance (*sameness*) entre deux images. En entraînant un programme de classification reposant sur cette cohérence spatiale, le processus a permis de détecter plus de 100 000 paires de photographies correspondantes dans la collection de la Fondation Cini. Davantage, il est possible d'obtenir d'importantes améliorations des performances des *ConvNet* en leur appliquant les contraintes de similarité encodées par les connexions et les annotations des usagers.

Par conséquent, numériser et traiter automatiquement des photographies d'objets artistiques, agréger de multiples collections qui se chevauchent et proviennent de différentes institutions, met en évidence des ressemblances physiques pour détecter la similarité visuelle apprise par des machines afin d'améliorer l'expérience de navigation et de recherche au sein des grandes collections d'images. Dans cette partie, l'auteur montre que cela peut être fait sans forcément considérer les métadonnées textuelles des documents. En outre, ce processus permet de détecter les images redondantes dans ces

grands corpus. Mais pour l'auteur, la fonction de similarité visuelle est l'élément le plus important de ce système de recherche, qui est en mesure d'améliorer continuellement ses performances au fur et à mesure de l'expansion des connexions visuelles.

9.1.5. Quatrième étape : le moteur de recherche *Replica*

Si l'objectif initial du projet est d'utiliser les approches propres à la *computer vision*, l'auteur affirme que l'utilisation d'informations textuelles (notamment l'extraction de métadonnées à partir des cartons originaux) est particulièrement bénéfique pour améliorer les performances de recherche du processus informatique. Pour ce faire, l'auteur présente le moteur de recherche qui a été conçu dans le cadre du projet Replica : une interface pour naviguer, explorer les images de la collection, mais aussi pour les annoter (cf. Seguin, 2018 : 103-125). En effet, il est essentiel de mettre en place un outil permettant aux usagers d'interagir avec ces larges ensembles d'images. Ce système de recherche doit fonctionner comme un outil d'indexation d'archives d'images qui fonctionne comme un moteur de recherche dans un ensemble d'images et de métadonnées relatives, mais aussi comme un dispositif de visualisation des résultats de la recherche.

Dans un premier temps, l'auteur traite le problème de l'indexation, c'est-à-dire celui de l'organisation des données, mais en évitant de collecter et de diffuser « manuellement » les données. Pour ce faire, il utilise le cadre normatif mis en place par l'IIIF[501] ; rappelons-le, ce cadre de référence propose d'utiliser un « langage commun » pour organiser des ressources visuelles à l'aide de l'informatique, ce qui permet également de standardiser ces ressources et favoriser l'homogénéisation des métadonnées. Nous avons vu que l'apport principal de l'IIIF est de standardiser les outils de transmission d'une image à une application, en passant par une base de données. Chaque ressource d'image est représentée par un lien URL qui propose un accès standardisé à différentes qualités de cette image (*Image API*) ; d'un autre côté, chaque objet est représenté à travers un « manifeste » (*manifest*) comprenant une description de l'objet et des consignes sur la manière dont le contenu doit être présenté. L'auteur se focalise surtout sur les deux APIs proposés par l'IIIF : *Image API* et *Presentation API*, car selon lui elles offrent des solutions et une interopérabilité durables. En effet, l'auteur

[501] **Chap. 6, sect. 6.2**.

précise que le cadre normatif de l'IIIF est adéquat pour organiser les données collectées par son *pipeline* :

> L'API de présentation permet une publication simple des ressources visuelles de manière unifiée entre les institutions, tout en conservant une référence à leur propre site web avec des informations supplémentaires. De même, l'API Image leur permet d'être le véritable fournisseur des images de manière directe, même si leur contenu est intégré dans une autre application, ce qui permet de conserver une citation correcte de la ressource (Seguin, 2018 : 106 ; nous traduisons de l'anglais).

L'auteur a donc converti ces données en vue de les rendre compatibles avec les normes IIIF, afin d'optimiser l'organisation de la collection d'images, ainsi que de divers ensembles de métadonnées récoltées sur Internet ou par le moyen du processus d'extraction automatique des métadonnées sur les documents originaux. Ceci permet de formaliser cet ensemble de données et de métadonnées en des répertoires standardisés, et de distinguer, d'une part, tous les numérisations et traitements appliqués à la collection, et d'autre part, l'indexation des ressources extraites pendant ce traitement. En fin de compte, l'auteur estime que la distribution de ressources visuelles à travers les outils de l'IIIF fonctionne bien, même s'il remarque des limitations liées au manque d'harmonisation des métadonnées d'une institution à une autre, malgré un effort constant de standardisation des métadonnées.

Dans un second temps, l'auteur présente les capacités de recherche de l'interface Replica, dont la visée est de permettre une flexibilité maximale aux utilisateurs pour explorer la collection. Dans cette perspective, l'interface comprend plusieurs « modes de recherche » :

- La recherche par métadonnées textuelles (le résultat de la recherche est décidé à partir des métadonnées extraites du manifeste IIIF, auxquelles sont rapportées les images correspondantes).
- La recherche à partir d'une image (à l'aide de la procédure détaillée dans la deuxième étape ; cf. **supra**) ou de plusieurs images (en utilisant une procédure de classification supplémentaire, en l'occurrence : *Support Vector Machine classifier*, un classificateur binaire).

L'auteur propose également une expérience dont l'objectif est de tester la performance de l'interface Replica afin de montrer que même un utilisateur qui n'est pas formé à

l'histoire de l'art peut rapidement apprendre à utiliser cette interface et à identifier de nombreuses connexions visuelles. Selon l'auteur, les résultats de cette expérience mettent en évidence « le pouvoir de généralisation de l'algorithme d'apprentissage et son utilité pour aider à la navigation dans les données », et surtout, elle « indique qu'un système plus précis apprend à partir des annotations précédemment saisies et aide les utilisateurs à trouver davantage de connexions visuelles » (cf. Seguin, 2018 : 122). Dans la conclusion de son travail, l'auteur souligne toutefois l'importance de reconnaître la « double nature de la photographie », en tant que représentation et en tant qu'objet ; c'est notamment cette distinction qui a permis de formaliser mathématiquement la similarité visuelle pour qu'elle soit compréhensible par l'ordinateur.

Somme toute, l'important travail de Seguin se répartit en quatre parties ou étapes qui visent à expliquer comment mettre en place un processus « multicouche » (*complex layered process*) permettant de rendre « explorable » et « navigable » une large collection d'images d'archives photographiques (fig. 103).

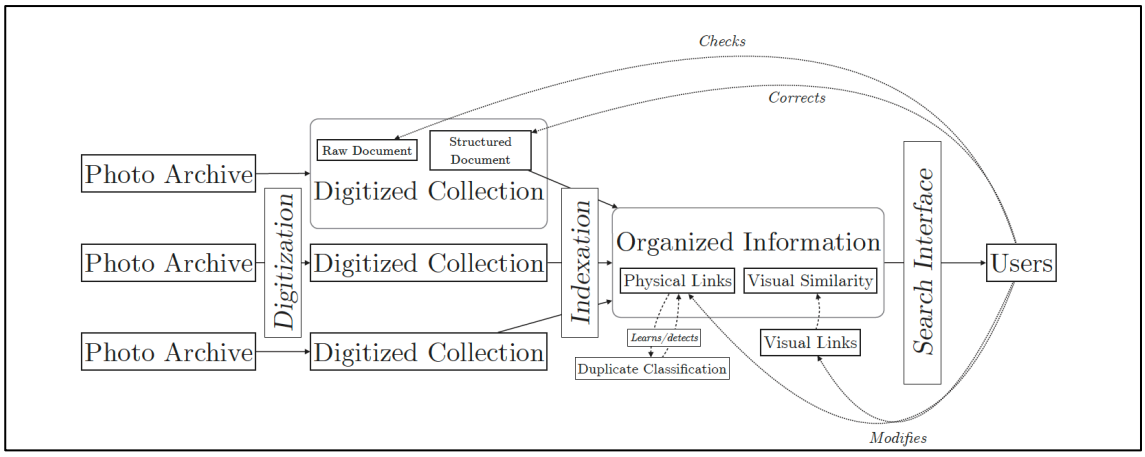


Figure 103 : Présentation schématique détaillée du projet Replica (cf. Seguin, 2018 : 6-8) ; en plus de convertir des archives photographiques en ressources numériques et d'indexer ces ressources, nous pouvons observer que le but est de proposer une interface de recherche, et surtout d'organiser ces ressources en fonction de caractéristiques visuelles qui sont à la fois mises en évidence par la machine (détection automatique) et par les utilisateurs (corrections et modifications).

La première étape consiste à proposer un état des lieux de la littérature à propos du traitement automatique de larges ensembles d'images. Plus précisément, il s'agit de mettre en évidence les principales caractéristiques et fonctionnalités du traitement informatique d'images (notamment la recherche d'image, mais aussi en ce qui concerne

« l'apprentissage de métrique » ou metric *learning*)[502], ainsi que l'évolution et les applications récentes de la « vision par ordinateur »[503] (*Computer vision*) dans le domaine de l'histoire de l'art.

La deuxième étape propose la solution qui a été utilisée pour convertir efficacement une archive photographique en données numériques. En plus de transformer plus de 300 000 photographies en images numériques, l'auteur propose un instrument pour segmenter des éléments visuels et textuels dans ces ressources physiques ; cet instrument se base sur des techniques avancées de traitement computationnel (*deep learning*). Ainsi, cette étape met en évidence la continuité entre processus de numérisation et extraction (semi)automatique de ressources documentaires de cette institution.

La troisième étape présente une approche mathématique de formalisation de la notion d'espace de l'image par rapport à la notion de similarité formelle. Cette formalisation du principe de ressemblance visuelle s'appuie sur des « connexions », c'est-à-dire des liens au sein d'un graphique (*morphograph*) où sont réparties spatialement différentes images. Cette partie du travail de Seguin est importante, car elle présente la méthodologie et le processus technique (*pipeline*) qui permet de généraliser l'information obtenue, soit sur base d'une relation de ressemblance (p. ex. : deux photographies représentent le même objet physique ; l'auteur parle dans ce cas de *physical connections*), soit d'une relation de similarité (visuelle) relative (*visual connections*). Pour ce faire, l'auteur utilise encore une fois des techniques avancées de traitement automatique d'images : le but étant de montrer comment une machine informatique peut apprendre une fonction de détection de similarité visuelle. Ainsi, cette troisième étape se concentre sur l'organisation et l'indexation de larges collections d'images, mais en introduisant un modèle formel pour représenter la complexité des relations (des « connexions ») entre ces images.

La quatrième étape détaille l'approche qui a permis de mettre en place une interface Web permettant de rechercher et de naviguer à travers une large collection d'images numériques. L'auteur explique que les progrès récents dans les technologies du

[502] Dans le domaine de l'apprentissage informatique en vue d'applications de classification, la « métrique », également appelée « distance » ou « similarité », est un principe qui permet de mesurer le degré de parenté de deux éléments d'un même ensemble.

[503] Le but de cette branche de l'intelligence artificielle est de permettre à un ordinateur d'analyser, de traiter et de comprendre une ou plusieurs images.

Web (notamment grâce à l'IIIF) fournissent des standards performants en vue de l'indexation de différents types de collections d'images. Ainsi, cette quatrième étape détaille la conception et les fonctionnalités de cette interface Web avec laquelle l'usager peut interagir et annoter les ressources documentaires.

En ce qui concerne la présentation des résultats de recherche, le but de l'interface est d'aider l'utilisateur à trouver des connexions visuelles entre des images et de les annoter dans un graphique, afin d'améliorer la « métrique de similarité visuelle » que la machine doit apprendre. Le moteur de recherche propose principalement deux solutions pour visualiser plusieurs images aux descripteurs visuels plus ou moins similaires :

- La sélection d'images est présentée par le moyen d'une « grille d'images » (*grid* ; cf. fig. 105).
- La sélection est représentée dans l'espace d'un système d'abscisses et d'ordonnées (*map* ; cf. fig. 106). Dans ce cas, les utilisateurs peuvent librement explorer et changer d'échelle dans la « carte de visualisations » (*map visualization interface*) afin d'explorer l'espace de similarité visuelle correspondant à la recherche.

D'une manière générale, lorsqu'une image est sélectionnée, celle-ci est présentée par une page Web qui regroupe les principales métadonnées sur le producteur, la date de production, la technique utilisée, le lieu de conservation, le type d'image, etc., ainsi que l'origine de ces métadonnées descriptives (fig. 104).

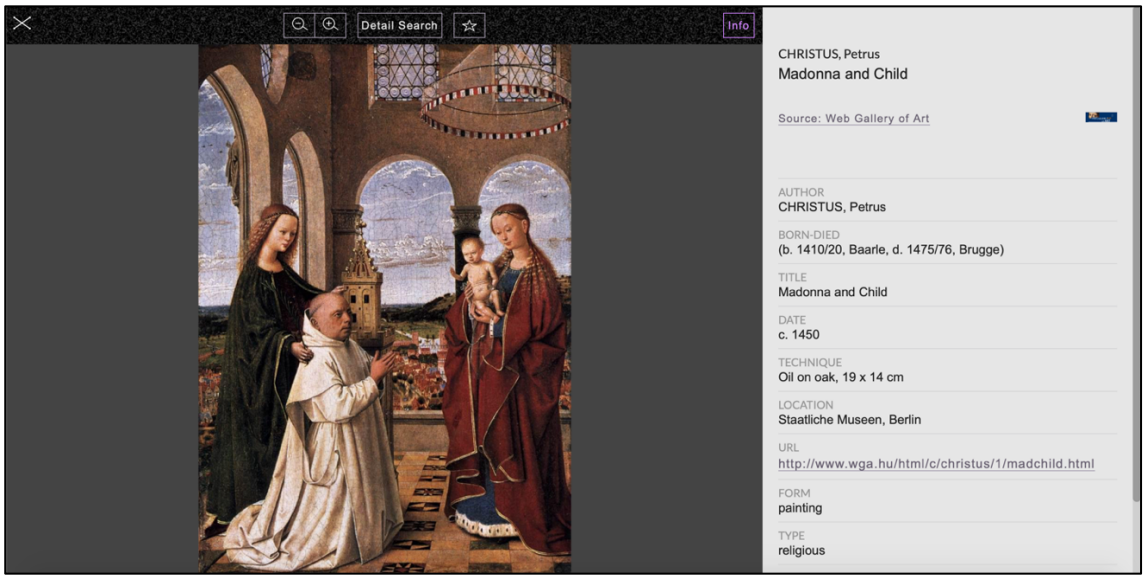


Figure 104 : Exemple de présentation d'une ressource (Replica ; à gauche, une reproduction numérique de l'image, à droite, les principales métadonnées descriptives).

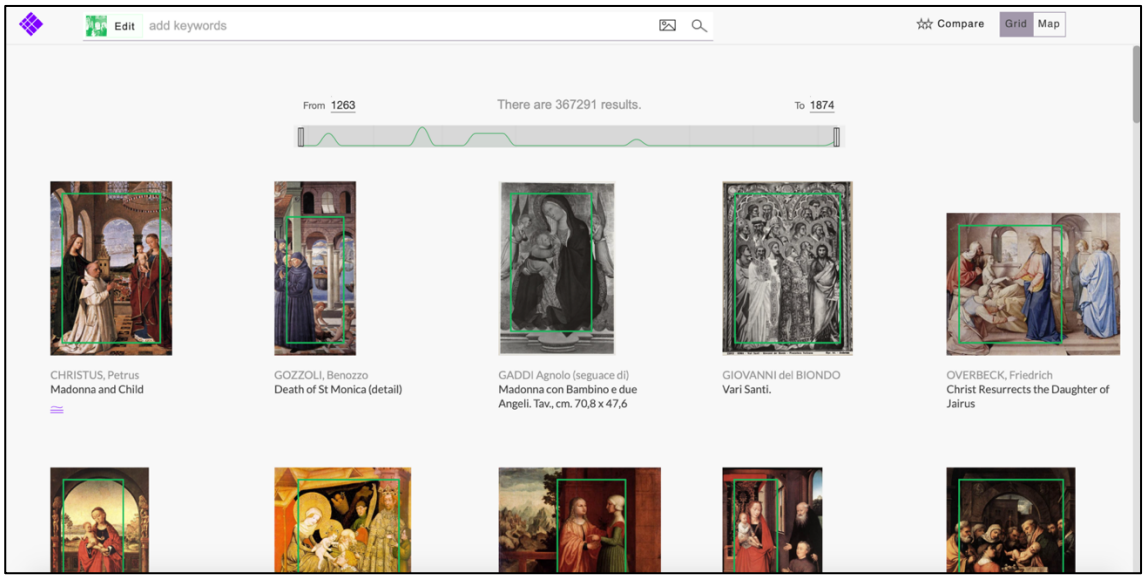


Figure 105 : Exemple d'une recherche à partir d'une image et de la présentation des résultats par le moyen d'une grille d'images (*grid*) (Replica) ; le système informatique sélectionne les images qui correspondent à la requête de l'utilisateur, qui peut éventuellement filtrer les résultats en fonction d'une période temporelle.

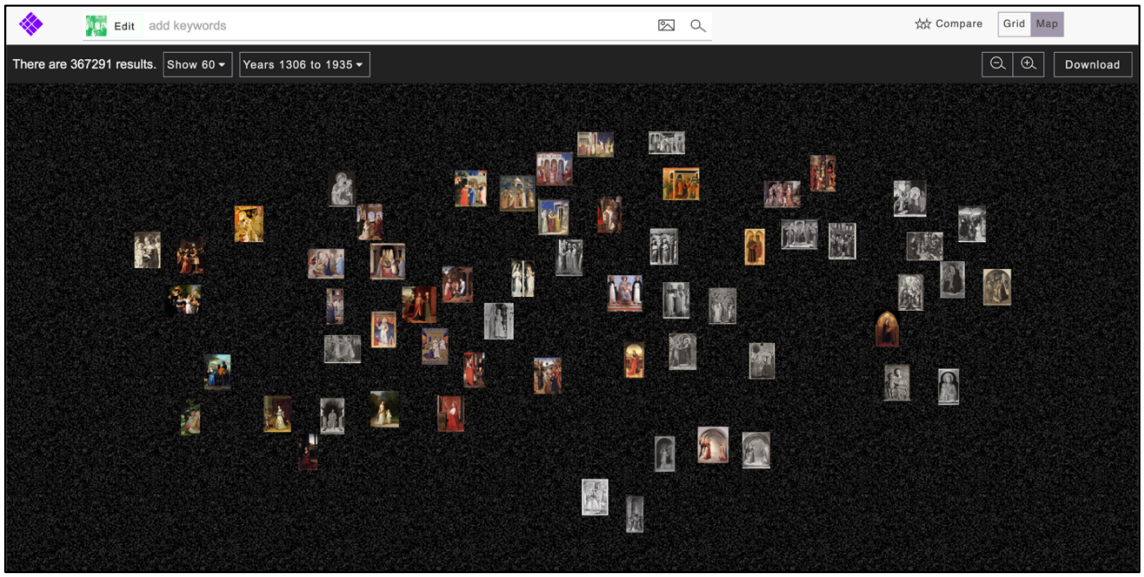


Figure 106 : Sélection de 60 résultats par le moyen d'une visualisation (*map visualization interface* ; les images sont situées dans l'espace de visualisation en fonction de la similarité visuelle établie par le traitement informatique des images).

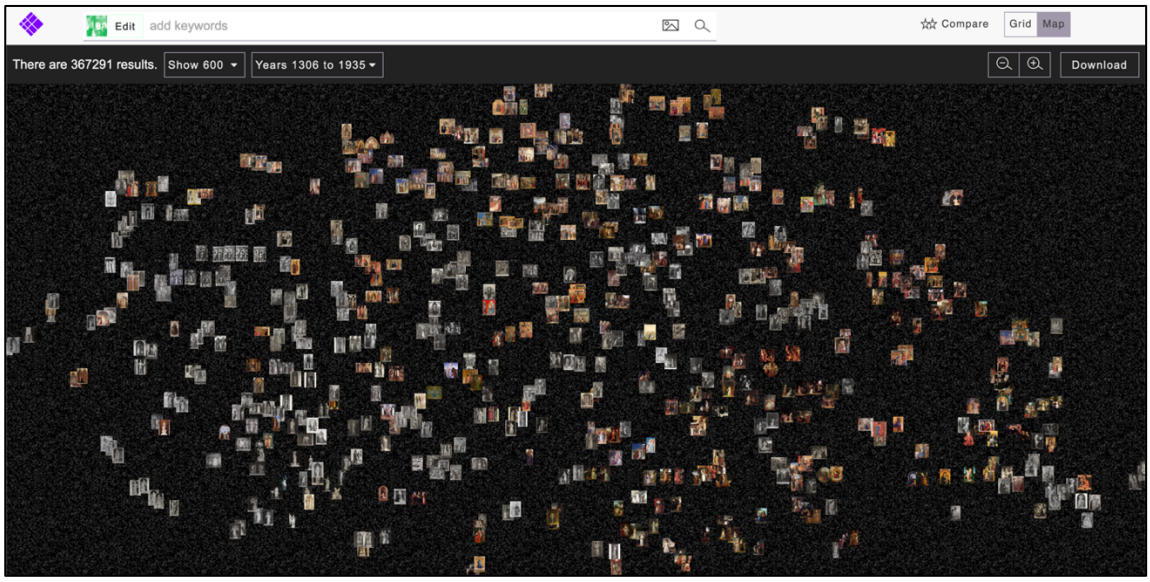


Figure 107 : Sélection de 600 résultats (*map visualization interface*).

Le moteur de recherche Replica s'inscrit résolument dans la même économie visuelle (le même *style de visibilité*) que certains dispositifs de visualisation que nous avons présentés dans le **chapitre précédent**. Ce dispositif de communication en ligne permet, d'une part, de sélectionner des images qui sont susceptibles de partager des caractéristiques visuelles (soit à partir d'une recherche par mots-clés soit en proposant directement une image soit les deux), et d'autre part, de visualiser le résultat de cette sélection. Dans cette perspective, le moteur de recherche Replica est un instrument puissant qui offre à l'utilisateur un dispositif de communication en ligne pour découvrir des connexions visuelles entre des images[504], mais aussi pour sélectionner des éléments pertinents, les comparer et les annoter pour créer plus de nouvelles connexions, de telle sorte que la participation des utilisateurs permet à terme d'améliorer la performance du système informatique.

Au terme de notre enquête, il est indéniable que la conversion numérique et la communication en ligne des archives et des collections d'images offrent des perspectives inédites en ce qui concerne l'étude contrastée des formes visuelles dans le but de *cartographier* les thèmes, les motifs et les formes de l'art à travers différentes époques. C'est du moins ce que nous suggérions en lisant les remarques d'André Chastel : nous

[504] À ce propos, Seguin présente quelques découvertes intéressantes qui ont été permises grâce à l'application de ce processus sur la collection (cf. *ibid.* : 133 *sq.*).

sommes désormais dans l'*ère numérique* de l'analyse des images[505]. Cette technologie est susceptible de renouveler profondément les pratiques d'analyse de larges ensembles d'images, puisque nous constatons qu'elle permet de comparer et de confronter des images dans un espace de visualisation. Ce point est important lorsqu'il est mis en parallèle de l'évolution des recherches en sciences humaines sur l'étude du visuel et des images. Nous l'avons vu, ces recherches affirment qu'il faut s'éloigner du « paradigme langagier » afin de favoriser un « paradigme de la présence des images ». Certes, dans l'analyse informatique d'images, cette présence des images n'est pas une présence matérielle, c'est-à-dire que des caractéristiques spécifiques au support matériel ne sont pas prises en compte dans un environnement informatique. Pourtant, l'avantage de ces techniques est sans doute de produire automatiquement des contrastes entre les images en fonction de paramètres plastiques ou formels, ce qui permet notamment de renouer avec les travaux récents en iconologie et en sémiotique.

Pour terminer cette section, nous pouvons rapidement mettre en évidence les possibilités de comparaison qu'offre ce moteur de recherche à travers deux exemples.

Premièrement, nous avons sélectionné une image disponible dans les archives photographiques de la Fondation Cini (fig. 108). Il s'agit d'une peinture du Maître de la Madeleine qui représente la vierge Marie et qui date du XIV^e^ ou du XV^e^ siècles. En utilisant cette image à la place d'une description verbale dans le champ de recherche, le moteur de recherche a proposé en retour plus de 360 000 images similaires qui ont été produites entre 1216 et 1745 (cf. fig. 109). Lorsque les résultats sont filtrés en fonction d'une période plus courte (1250 à 1327), le nombre de résultats est réduit à plus ou moins 14 000 images similaires (cf. fig. 110). Si la plupart des résultats concernent des images similaires, il est possible que le moteur de recherche propose des images qui représentent un autre personnage que la célèbre *Madonna*. Pour contrer ce problème, l'interface permet justement d'exclure un certain type d'images durant la navigation ; en plus du « filtre temporel », ce « filtre d'exclusion » peut être utilisé à la fois dans la présentation des résultats sous la forme d'une grille d'image (cf. fig. 111) que dans la visualisation des images (cf. fig. 112). L'instrument est efficace dans le sens où il permet de constituer rapidement des corpus d'images qui partagent des caractéristiques visuelles. Mieux, nous

[505] **Chap. 3, sect. 3.4.**

observons que le corpus peut indiquer des récurrences, mais aussi des différences minimes dans la composition formelle des images (par exemple : positions et postures des personnages)[506]. Comme dans l'*Atlas Mnémosyne*, le dispositif de visualisation d'images peut être considéré comme un « terrain d'expérimentation et de confrontation visuelle des images », sauf que dans ce cas-ci, c'est le système informatique qui crée des rapprochements entre les images.

Deuxièmement, nous avons choisi deux peintures sur le même sujet (cf. fig. 113 & 115), mais deux images qui ont été réalisées à deux époques différentes : *La Madone à la prairie* de Raphael, qui date du début du XVI^e^ siècle, et *Madonna* d'Edvard Munch, datant de la fin du XIX^e^ siècle. L'idée est d'éprouver l'instrument afin de voir comment celui-ci va rapprocher ces deux images d'autres images de la collection. Encore une fois, nous pouvons observer que le moteur de recherche propose des résultats qui rassemblent des images qui partagent une même composition formelle. Cette similarité ne concerne pas seulement d'autres images de *Madonna*, mais encore des images qui partagent la même dynamique dans la composition des formes (fig. 114 & 116). Notons en effet que la *Madonna* de Munch est considérée comme similaire à différentes représentations de Jésus-Christ, voire à une sculpture.

[506] Nous rejoignons ici les discussions sur l'expérience de Impett & Moretti (2017) qui cherche à utiliser le traitement informatique d'images pour étudier la notion de *Pathosformel* ; cf. **chap. 3, sect. 3.1.4.**

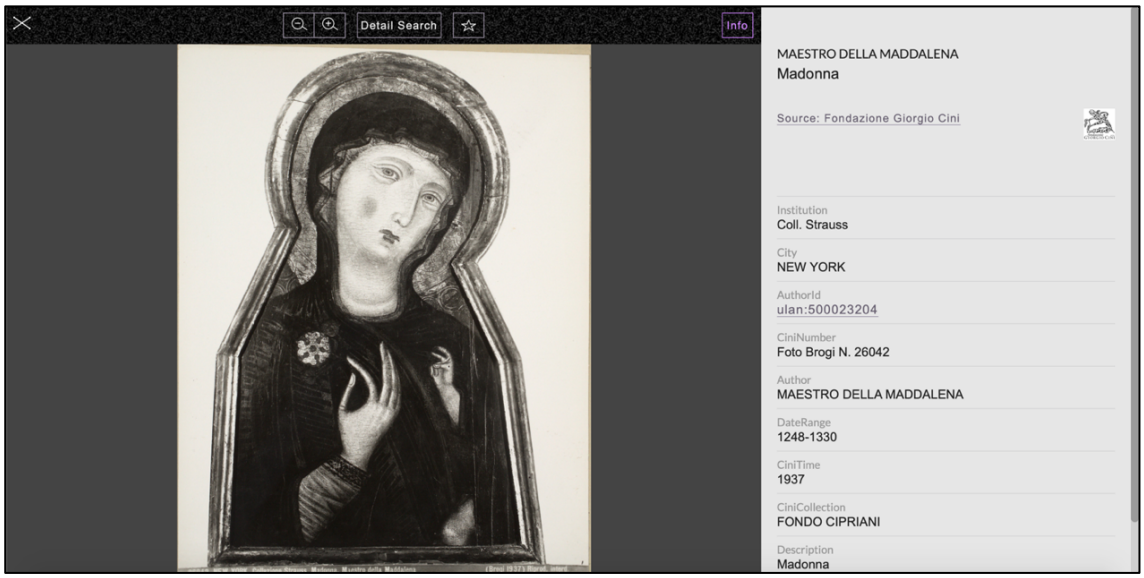


Figure 108 : Présentation d'une ressource de la *Fondazione Giorgi Cini* ; à droite, quelques métadonnées descriptives.

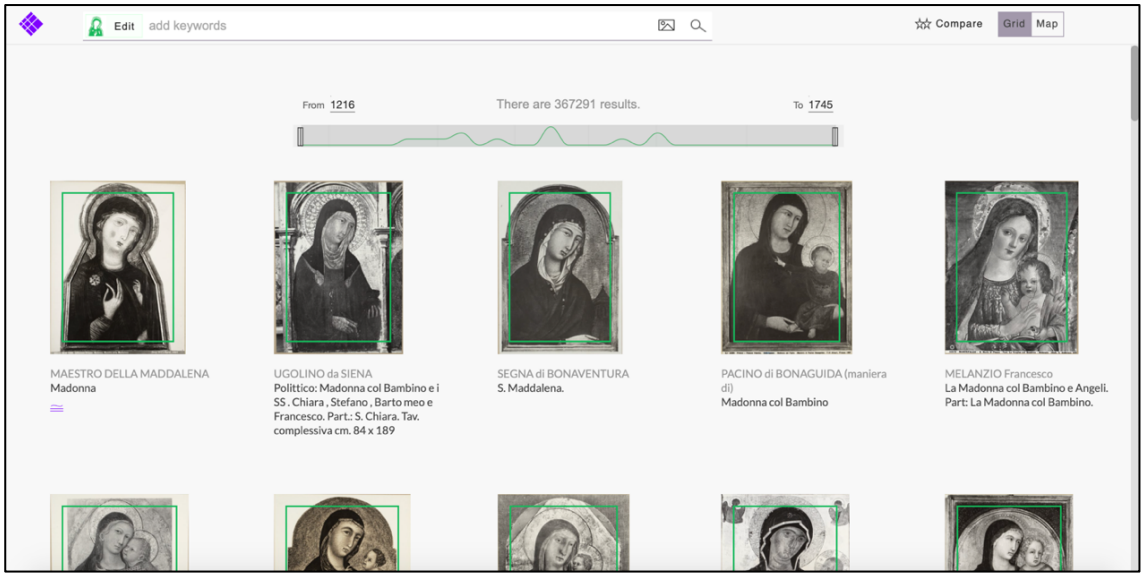


Figure 109 : Résultat d'une recherche par correspondance visuelle en fonction d'une période donnée, dans ce cas à peu près cinq siècles (plus de 360 000 résultats ; présentation *grid*).

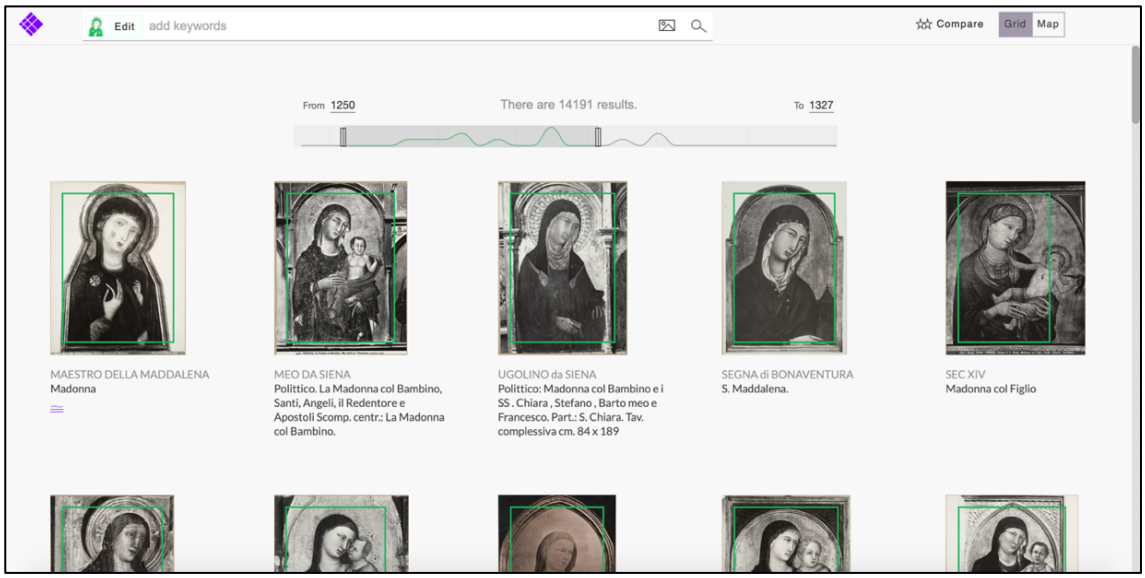


Figure 110 : Résultat d'une recherche par correspondance visuelle en fonction d'une période donnée, dans ce cas à peu près un siècle (plus de 14 000 résultats ; présentation *grid*).

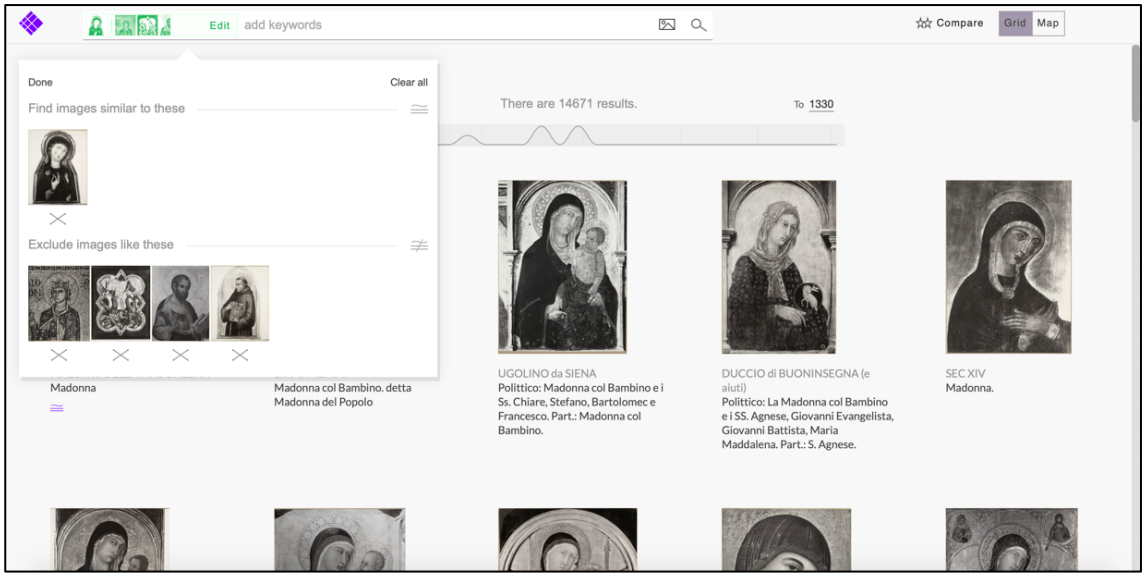


Figure 111 : Le moteur de recherche Replica permet aussi de filtrer les résultats en excluant certaines images (cf. à gauche).

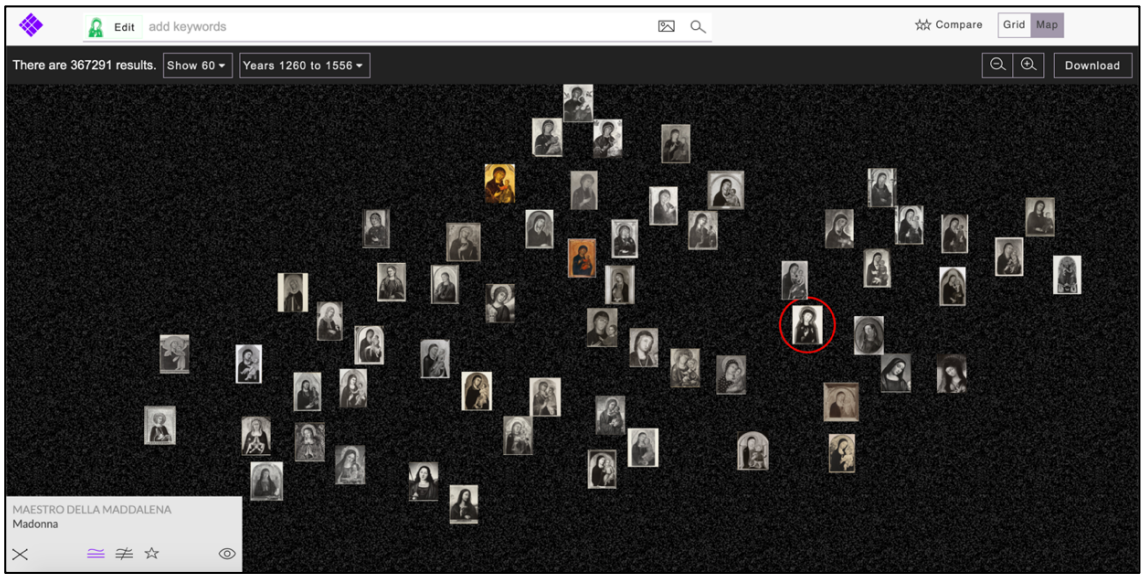


Figure 112 : Sélection de 60 images qui résulte d'une recherche par correspondance visuelle en fonction d'une période donnée, dans ce cas à peu près trois siècles (plus de 360 000 résultats ; présentation *map visualization interface*).

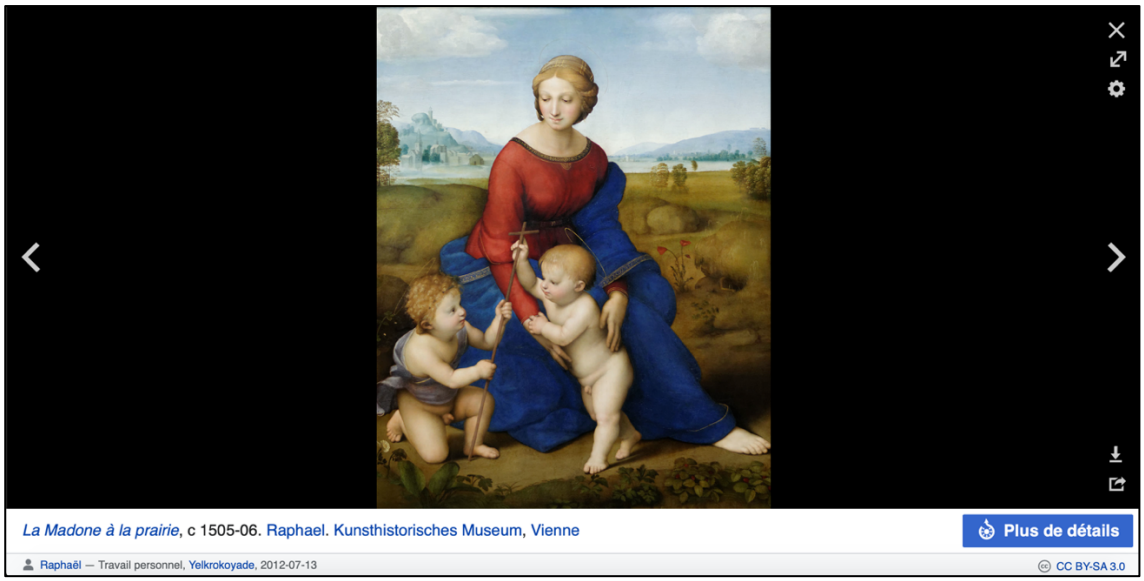


Figure 113 : *La Madone à la prairie* de Raphael (disponible sur *WikimediaCommons*).

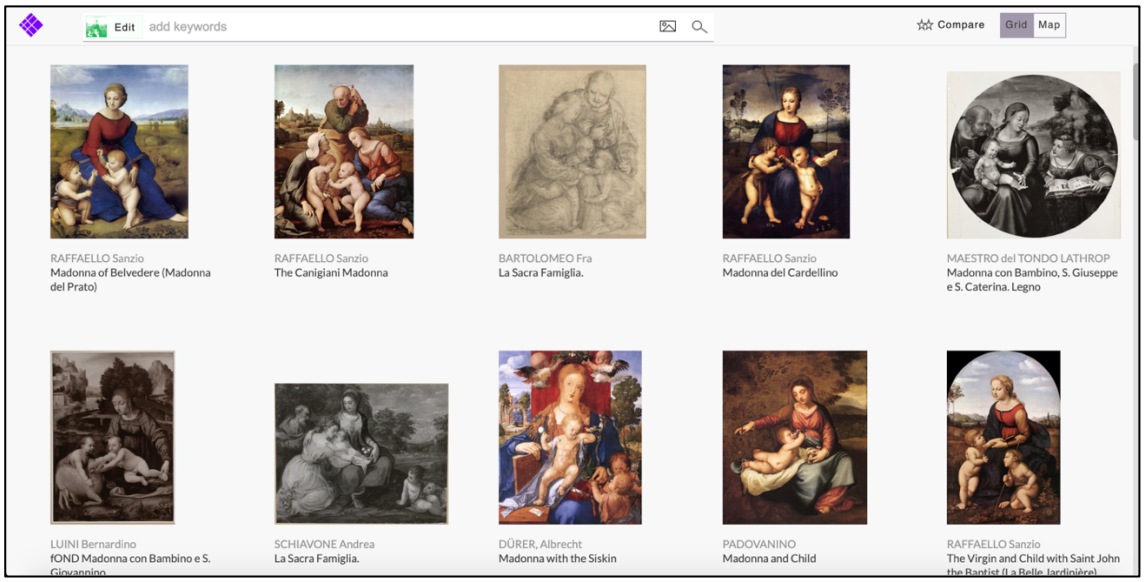


Figure 114 : Résultat d'une recherche sur Replica à partir d'une « image extérieure » (*WikimediaCommons*).

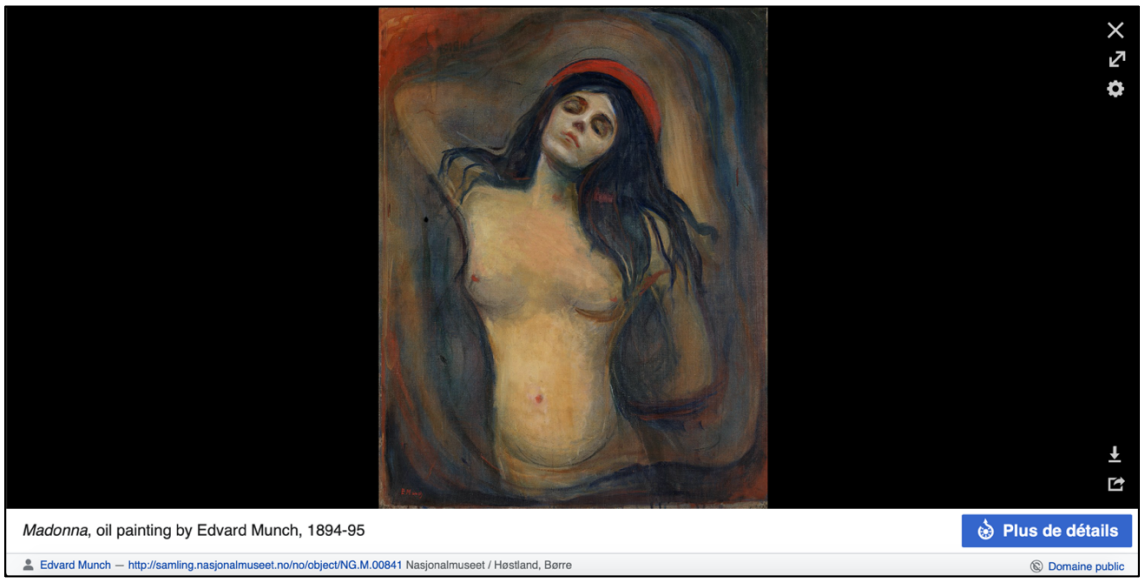


Figure 115 : Madonna d'Edvard Munch (disponible sur *WikimediaCommons*).

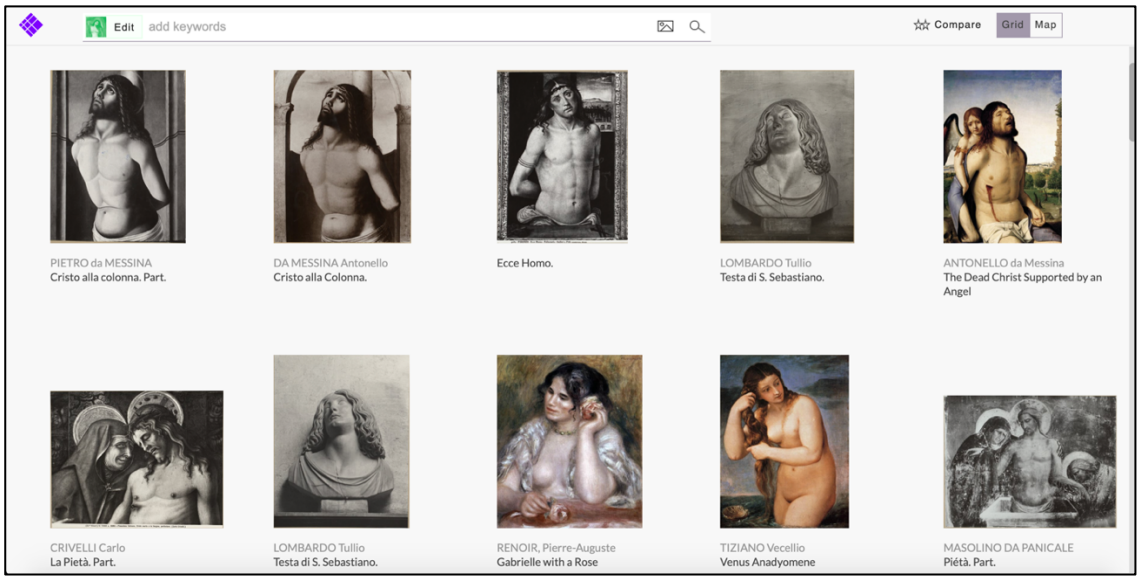


Figure 116 : Résultat d'une recherche sur Replica à partir d'une « image extérieure » (*WikimediaCommons*).

En résumé, le projet Replica a cherché à répondre à deux questions : « comment rendre les collections photographiques accessibles ? » et « comment aider les historiens de l'art pour explorer de ces collections ? ». Les sections précédentes font ressortir que ce projet a répondu à ces questions en proposant quatre étapes principales : (1) utiliser un instrument (scanner) spécifique permet la numérisation efficace des ressources documentaires ; (2) traiter automatiquement les numérisations et les convertir en données structurées ; (3) développer une formalisation du concept de similarité visuelle ; (4) concevoir une interface Web pour rendre accessible ce large ensemble d'images et de données numériques.

Dans cette perspective, il semble évident que ce type de moteur de recherche constitue un instrument qui peut aider des historiens de l'art à *explorer* et à *observer* un large ensemble d'images. Cette dimension instrumentale doit justement être comprise comme un support pour la recherche et non pas comme une technique qui produit automatiquement des analyses de formes visuelles. En effet, même si un système informatique met en évidence des relations entre des images en fonction de leurs similarités visuelles, l'*utilisateur-expérimentateur* doit nécessairement corriger, aménager ou réduire les résultats qui sont souvent beaucoup trop nombreux. Cependant,

ces résultats « bruts » sont un point de départ possible pour étudier l'évolution de certaines compositions formelles par rapport à des images qui partagent ou non ces caractéristiques formelles, mais aussi en fonction d'une période plus ou moins longue.

Dans ce **dernier chapitre**, nous avons surtout voulu mettre en évidence que les possibilités du traitement informatique sont convaincantes, même si ces techniques et instruments ne sont pas parfaits (d'ailleurs, nous avons largement souligné que la recherche d'une méthode parfaite est une utopie). Pour les améliorer, il s'agit sans aucun doute de favoriser davantage la collaboration entre les chercheurs en sciences humaines et les informaticiens, car c'est seulement en coopérant que ces nouvelles possibilités pourront aider les chercheurs à mieux comprendre les images et le visuel. D'un côté, les sciences humaines doivent indiquer aux informaticiens « ce qu'ils attendent » du traitement automatique d'images. Néanmoins, les sciences humaines doivent prendre en compte la dimension instrumentale (et non miraculeuse) du numérique en se formant à ces nouvelles technologies. De l'autre, les informaticiens gagneraient à prendre en compte certaines recherches récentes en sciences humaines et, surtout, à ne pas se limiter à développer uniquement leurs propres théories et méthodes. En outre, ce rapprochement permettrait sans doute de mettre plus clairement en évidence les avantages et les limites du traitement informatique d'images, surtout lorsqu'il s'agit d'expliquer ces avantages et ces limites à des spécialistes du visuel qui n'ont pas encore de connaissances en informatique.

9.2. Autres projets qui se basent sur la *media visualization* et sur les algorithmes

Nous pouvons en outre observer que les résultats de ce projet ont encouragé d'autres initiatives similaires. À titre d'exemple, le *Digital Humanities Lab* (DHLab) est un service mis en place par l'Université de Yale pour aider les chercheurs à utiliser des outils et des méthodes informatiques dans des recherches en sciences humaines. Dans cette perspective, plusieurs projets de cette université font appel à des technologies en lien avec le *deep learning* pour répondre à des questions sur l'histoire culturelle. Parmi ces projets, *PixPlot*[507] est une interface qui a pour but de faciliter l'exploration dynamique de

[507] Voir : https://dhlab.yale.edu/projects/pixplot/ et http://www.catherinederose.com/pixplot.

plusieurs milliers d'images numériques en s'inspirant notamment des travaux de Benoît Seguin.

L'idée principale est de visualiser une collection d'images sur le patrimoine culturel (*Meserve-Kunhardt Collection*) en utilisant des techniques propres à la *computer vision* et à la *data visualization* pour constituer des *clusters* (des « regroupements ») d'images similaires. *PixPlot* est conçu comme un logiciel libre dont le but est de visualiser des collections d'images à grande échelle, mais par le moyen d'une mise en scène interactive. Chaque image a été traitée à l'aide de *ConvNet* entraînés à l'aide d'*ImageNet*[508]. Les visualisations agglomèrent les images visuellement similaires dans une même zone de cette mise en scène.

Nous l'avons vu, le traitement informatique consiste à utiliser des algorithmes pour réaliser certaines tâches sur des collections d'images (visualiser des images similaires dans un grand corpus d'images ; identifier des objets dans des images ; trouver des images similaires dans un corpus). Ainsi, les utilisateurs de *PixPlot* peuvent naviguer librement dans cet espace de visualisation qui met en évidence des groupes d'images (des *clusters*). En outre, il peut aussi sélectionner des *hot spots* (points d'intérêts) qui présentent une image représentative de chaque groupe identifié par l'ordinateur (cf. fig. 117, 118 & 119).

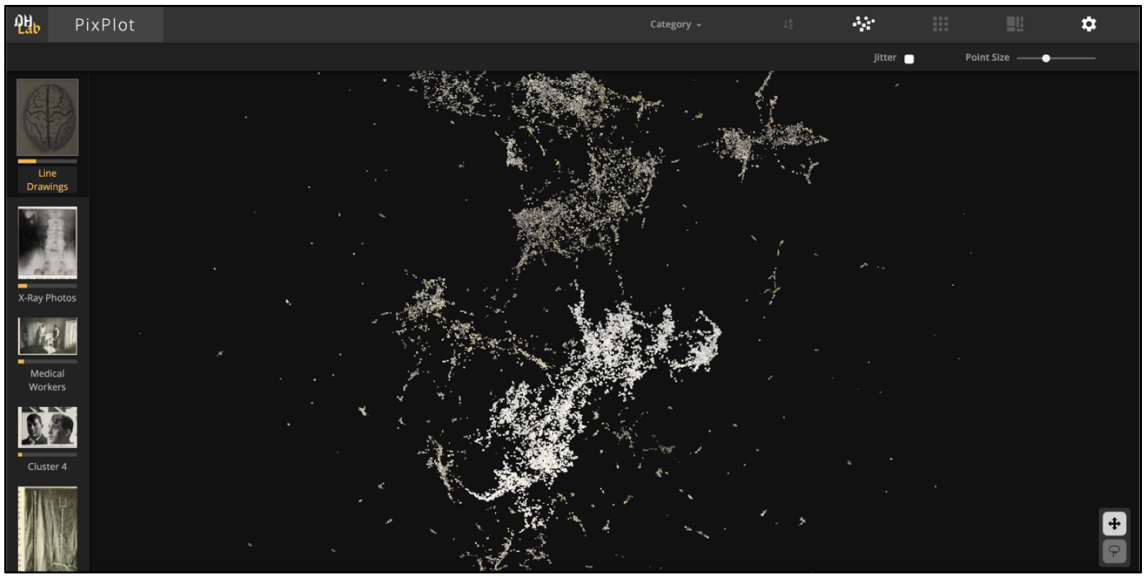


Figure 117 : Visualisation en réseaux qui regroupe les images en fonction de leurs similarités visuelles (*PixPlot*).

[508] **Chap. 4, sect. 4.2.3.**

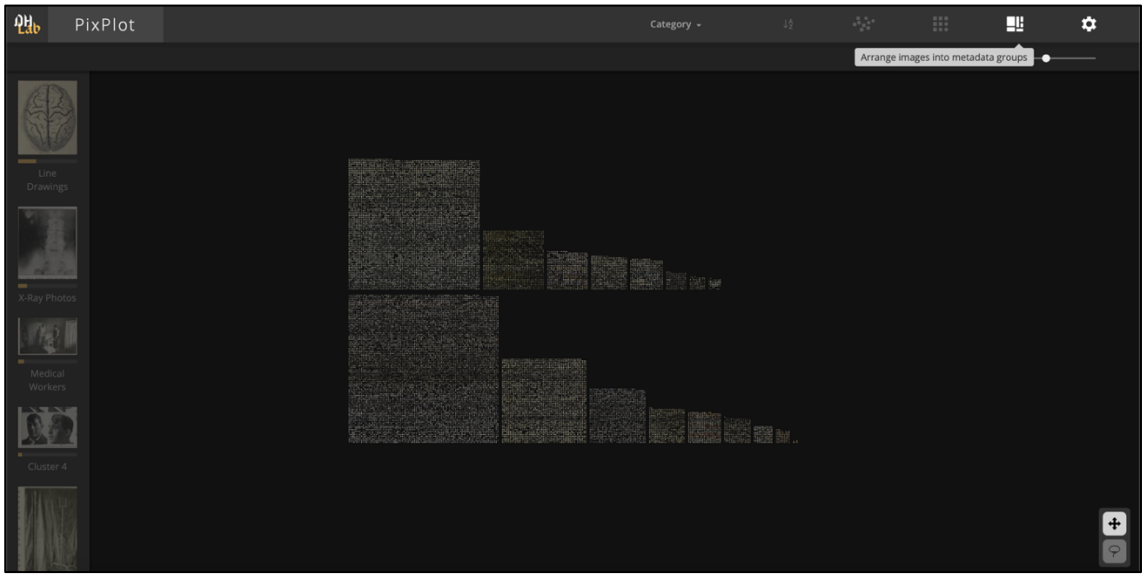


Figure 118 : Visualisation qui se fonde sur le regroupement des images en fonction de leurs métadonnées (*PixPlot*).



Figure 119 : Visualisation qui classe les images en fonction du nom du fichier, c'est-à-dire par ordre alphabétique (*PixPlot*).

À titre d'exemple, nous pouvons remarquer deux applications de *Pixplot* qui illustrent bien le *style de visibilité* que nous avons observé dans ce **chapitre**.

En se basant sur un algorithme t-SNE pour classer automatiquement des images, ce projet a mis en place un dispositif pour créer des visualisations bidimensionnelles ou tridimensionnelles (cf. fig. 120-123), dispositif de visualisation qui permet d'observer plus ou moins 24 000 photographies de presse et d'actualité quotidiennes qui ont été produites vers 1910 par le *Bain News Service*[509]. Dans la même perspective, un dispositif de visualisation permet d'explorer une sélection d'à peu près 13 000 images (sculptures, peintures, dessins) provenant de différents musées et galeries de la *Smithsonian Institution*, encore une fois, à l'aide d'un traitement informatique qui se fonde sur des algorithmes. Comme dans l'exemple précédent, le dispositif permet de choisir entre différents modèles de visualisation des images en fonction du type d'algorithme utilisé pour la classification et la visualisation des images (cf. fig. 124 & 125).

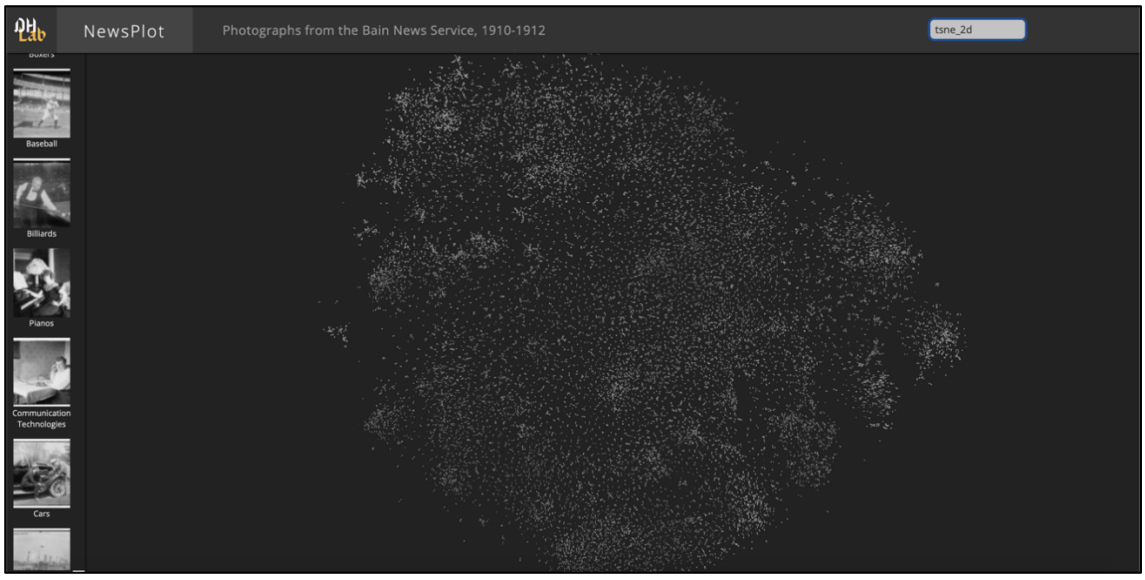


Figure 120 : Autre exemple de visualisation (*NewsPlot*).

[509] Il s'agit d'un ensemble d'images qui proviennent de la *George Grantham Bain Collection* qui est conservée à la *Library of Congres*.



Figure 121 : Même exemple de visualisation (*NewsPlot* ; vue rapprochée)

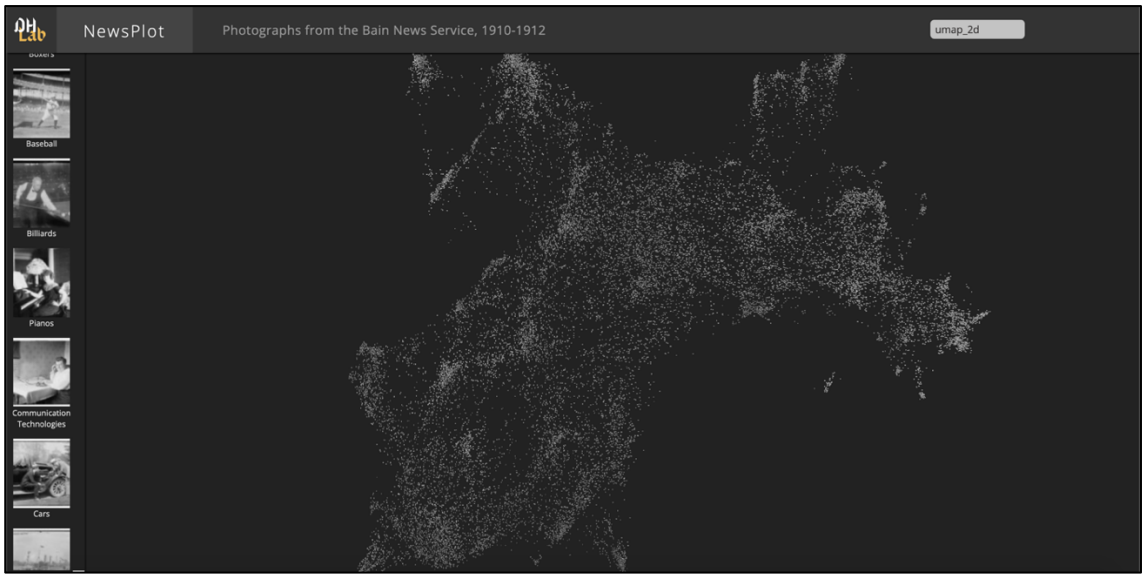


Figure 122 : Autre exemple de visualisation (*NewsPlot*).



Figure 123 : Autre exemple de visualisation (*NewsPlot*)

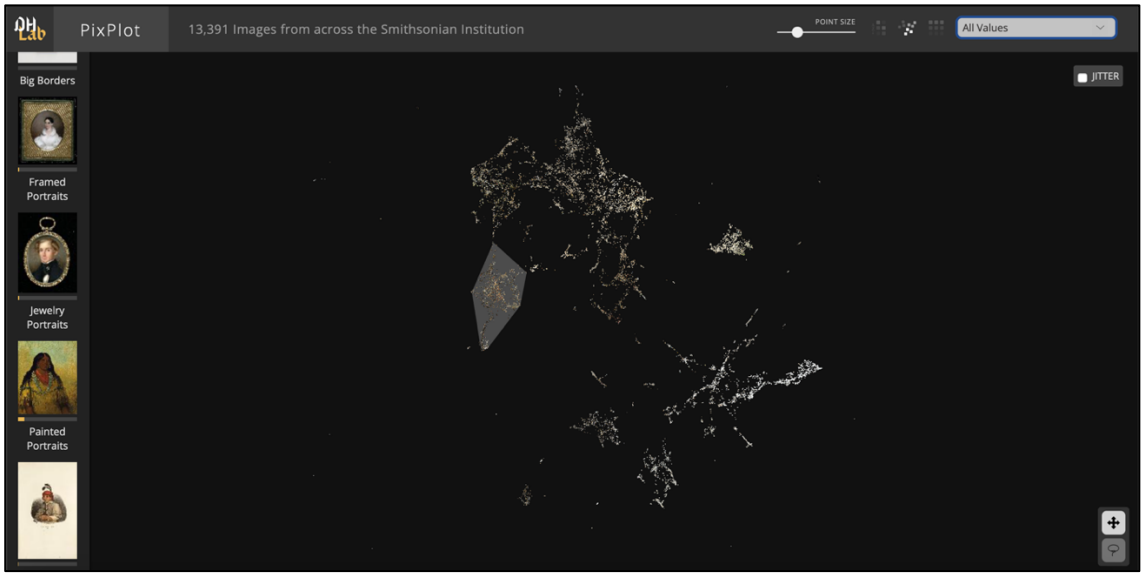


Figure 124 : Autre exemple de visualisation (*PixPlot*)

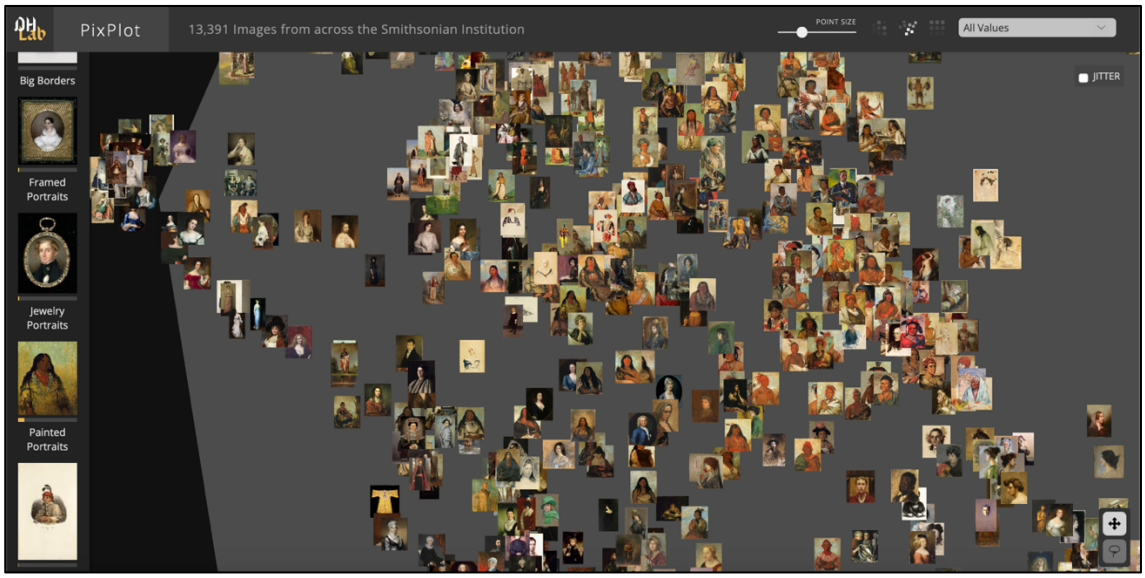


Figure 125 : Autre exemple de visualisation (*PixPlot*)

Nous avons surtout présenté des projets qui concernent la classification et la visualisation d'images fixes, quoique d'autres initiatives proposent d'utiliser le traitement informatique d'images sur d'autres types de ressources visuelles.

Par exemple, le but du *Sensory Moving Image Archive Project* (ou *SEMIA Project*) est d'étudier comment explorer, naviguer et réutiliser divers « sous-ensembles spécifiques » d'objets du patrimoine, notamment par le moyen des images en mouvement, mais en utilisant des outils computationnels d'analyse et de visualisation des caractéristiques et des relations visuelles dans les images. Le projet *SEMIA* a été mis en place par Bert Spaan et Nanne van Noord en collaboration avec différentes institutions (University of Amsterdam ; Amsterdam University of Applied Sciences ; Netherlands Institute for Sound and Vision ; Studio Louter ; EYE Filmmuseum). Le but est d'utiliser la *computer vision* et le *deep learning* pour analyser des images en fonction de quatre paramètres : couleur la plus proéminente dans le plan (*color*), les formes détectées dans l'image (*shape*), la direction du mouvement principal dans l'image (*movement*), la complexité visuelle du plan (*clutter*). Pour ce faire, ce projet utilise deux collections patrimoniales : la collection *Desmet de EYE* (films muets) et les émissions d'actualités de *NTS/NOS de Sound and Vision*. Ce projet propose un site Web[510] qui permet de visualiser la similitude au sein d'un corpus de 100 000 images extraites de 6 969 vidéos disponibles

[510] Cf. https://bertspaan.nl/semia/#/.

dans *Open Images*, un ensemble de ressources audiovisuelles sous licence *Creative Commons* (le même algorithme t-SNE a été utilisé pour créer des visualisations bidimensionnelles en fonction de la similarité entre tous les plans ; cf. fig. 126).

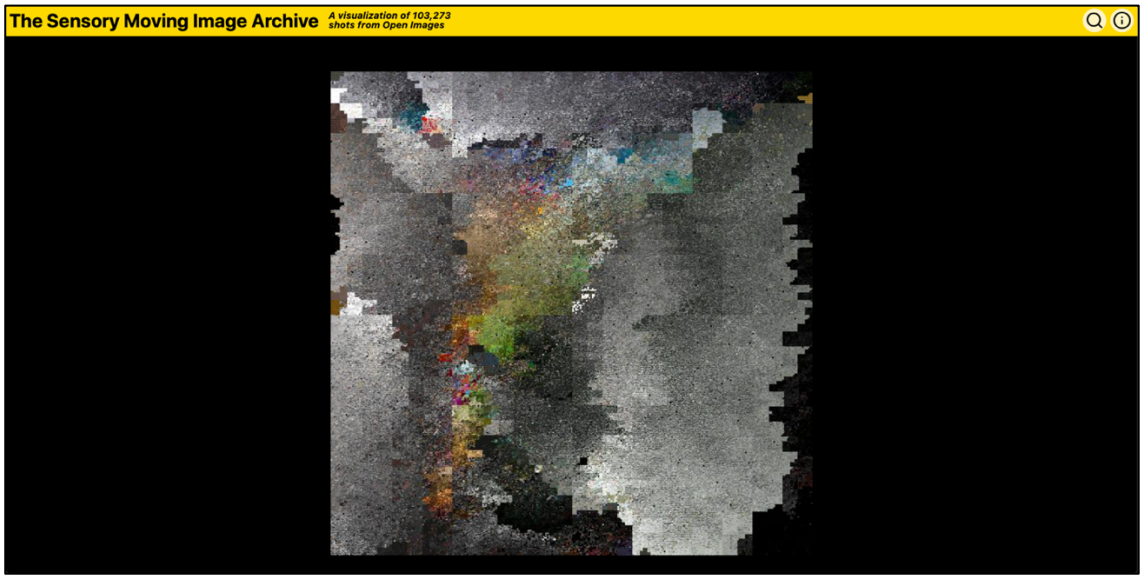


Figure 126 : Exemple de visualisation en fonction de la couleur du plan (*The Sensory Moving Image Archive*).

Dans une autre perspective, *Recognition* est un projet culturel qui a employé le traitement informatique pour comparer et trouver des similitudes entre des œuvres d'art conservées par le *Tate* et des photographies contemporaines d'actualités de l'agence de presse *Reuters*. Ce programme a été actif du 2 septembre au 27 novembre 2016 sous forme de site Web et d'une installation (*Tate Britain*). Cette installation interactive a été mise en place pour que les visiteurs du musée puissent également faire correspondre une œuvre d'art avec une image d'actualité. S'il s'agit de réaliser la même tâche que le système informatique de *Recognition*, l'ensemble de données proposé aux utilisateurs est plus restreint. L'objectif de l'installation était de comparer les décisions prises par *Recognition* à celles prises par des êtres humains. Grâce à des algorithmes qui simulent la manière dont les humains voient et interprètent des images, ce programme recherche des similitudes visuelles et thématiques dans la base d'images du *Tate*. Développés par les *computer scientists* de *JoliBrain* (*DeepDetect*), ces algorithmes se basent sur l'application d'un ConvNet qui, d'une part, détecte des objets dans des images, et d'autre part, associe des descriptions simples à ces images. Nous l'avons vu, ce type de traitement automatique d'images permet d'extraire des caractéristiques visuelles saillantes dans un grand nombre d'images. Dans le même sens, des algorithmes de reconnaissance faciale

(*Microsoft Cognitive Services*) ou de traitement de textes (pour trouver des correspondances sémantiques dans les légendes ou descriptions des images) ont été utilisés pour traiter les images. Ce processus de reconnaissance d'objets est couplé à un moteur de recherche qui permet d'établir des correspondances par similarité entre des images traitées par la machine. Durant les trois mois de fonctionnement du processus informatique, Recognition a produit plusieurs milliers de correspondances entre les œuvres d'art de la collection du *Tate* et les images d'actualité (7271 correspondances).

Nous avons rapidement présenté quelques projets qui se basent sur la *computer vision*, sur la *media visualization* et le *deep learning*. Nous en omettons sans aucun doute beaucoup d'autres. Nous avons seulement voulu mettre en évidence la multiplication de ce type de projet.

Somme toute, le présent travail a surtout considéré l'impact du numérique sur la communication en ligne et l'analyse d'images. Cependant, les progrès dans le traitement informatique d'images aboutissent également sur des performances remarquables en ce qui concerne la génération (ou la création) automatique d'images. Même si la question de la production d'images à l'aide de l'informatique ne fait pas partie des objectifs que nous nous sommes fixés, nous présenterons brièvement deux initiatives incontournables dans ce domaine[511].

DALL-E est un modèle informatique développé par l'entreprise *OpenAI* qui est spécialisée dans l'utilisation du *deep learning* pour générer des images numériques à partir de descriptions verbales (celles-ci sont nommées « *prompts* »). La première version de DALL-E (janvier 2021) utilise une version de GPT-3[512] qui a été modifiée pour générer des images. En 2022, DALL-E 2 permet de générer des images plus réalistes (en fonction de concepts, d'attributs et de « styles ») et en plus haute résolution. En 2023, *OpenAI* a annoncé que la troisième version (DALL-E 3) sera capable de comprendre « significativement plus de nuances et de détails » que dans les versions précédentes[513].

[511] Notons que ces services de génération automatique d'images sont payants.

[512] GPT (actuellement, la quatrième version GPT-4) est un « modèle multimodal à grande échelle » (*large-scale multimodal model*) capable de traiter des images et du texte en *input* et de produire du texte en *output*. Cf. le rapport technique d'*OpenAI* : https://arxiv.org/pdf/2303.08774.pdf.

[513] Cf. https://openai.com/dall-e-2/ & https://openai.com/dall-e-3/.

À ce jour, DALL-E est une implémentation de GPT-3 avec 12 milliards de paramètres qui « échange du texte contre des pixels » ; le modèle est entraîné sur des couples « texte-image » provenant d'Internet[514]. DALL-E peut générer différents types d'images, notamment des « photographies » (cf. Saharia, *et al.* 2022) et des « peintures », ou encore des variations d'une image originale. La compréhension des prompts par DALL-E 2 est limitée (il est possible que le modèle confonde des concepts ou ne les représente pas de façon cohérente dans l'image générée ; une requête comprenant trop d'objets peut entraîner des erreurs)[515].

Dans la même perspective, Midjourney[516] est un service d'intelligence artificielle générative créé et hébergé par le laboratoire de recherche éponyme (l'équipe est dirigée par David Holz). Comme DALL-E, Midjourney génère des images à partir de descriptions verbales. Les utilisateurs créent des images avec Midjourney à l'aide d'un *bot* Discord[517]. Pour générer des images, les utilisateurs utilisent une commande (*imagine*) et encodent les descriptions à l'aide d'un *prompt* ; ensuite, le « bot » renvoie quatre images à l'utilisateur qui peut alors choisir les images qu'il souhaite transformer dans une résolution plus grande. En outre, Midjourney propose d'autres commandes à envoyer au bot (la commande *blend* qui permet à l'utilisateur de mélanger deux images ; la commande *shorten* qui permet d'obtenir des suggestions sur la façon de raccourcir une description trop longue).

Comme dans l'analyse et la classification d'images, les usages du *deep learning* offrent certainement de nouvelles possibilités pour générer des images. Toutefois, ces usages suscitent aussi des problématiques et des questionnements que les sciences humaines et les sciences informatiques doivent dès lors prendre en compte. Quels sont

---

[514] DALL-E a été développé conjointement avec CLIP (*Contrastive Language-Image Pre-training*), un autre modèle basé sur le *deep learning* et qui a été entraîné sur 400 millions de couple texte-image. Le rôle de ce second modèle est de comprendre et de classer les résultats de DALL-E en prédisant quelle légende (dans une liste de plusieurs dizaines de milliers de légendes sélectionnées au hasard dans l'ensemble de données) est la plus appropriée pour une image. CLIP est utilisé pour filtrer la liste d'images générées par DALL-E dans le but de sélectionner les *outputs* les plus appropriés à la requête de l'utilisateur. Pour plus d'informations sur la dernière version de DALL-E, voir notamment Belter *et al.*, « Improving Image Generation with Better Captions » (cf. https://cdn.openai.com/papers/dall-e-3.pdf).

[515] Sur les limitations de DALL-E, cf. Strickland (2022).

[516] https://www.midjourney.com/.

[517] *Discord* est un logiciel gratuit de messagerie instantanée (https://discord.com). En anglais, un *bot* est un programme informatique dont le but est d'automatiser des tâches répétitives en utilisant des algorithmes qui simulent le comportement d'un humain ; cf. **chap. 4, sect. 4.2.2.**

les avantages et les limites de ces techniques ? Quels secteurs d'activités sont susceptibles d'utiliser l'intelligence artificielle pour esquisser rapidement des images ? Les images créées automatiquement dans un environnement virtuel sont-elles comparables à celles qui sont produites par les gestes du corps ? Est-ce que ces techniques « dévalorisent » le travail créatif original (du dessinateur, du peintre, du photographe, etc.) et par conséquent, les contextes pratiques et les acteurs qui produisent habituellement des images ? Ou alors, au contraire, est-ce que ces techniques « augmentent » à tel point les possibilités de création d'images qu'elles nous engagent vers une révolution dans la culture visuelle ?

Toutes ces questions (et encore beaucoup d'autres) méritent d'être examinées. Nous ne doutons pas que des recherches en cours s'attellent à trouver des réponses sur ce sujet, et que d'autres recherches à venir parviendront sans doute à expliquer l'impact du numérique sur la culture visuelle, qu'il s'agisse de génération ou d'analyse d'images.

9.3. Conclusions

Dans ce **dernier chapitre**, nous avons constaté que la troisième tendance stratégique met en évidence des solutions innovantes qui proposent des instruments de navigation et d'exploration des ressources numériques. Ces instruments s'appuient sur des opérations de calculs et des algorithmes dont le but est souvent de mettre en relation un large ensemble d'images afin de proposer un point de vue sur toutes les images à l'aide d'un modèle de visualisation. Nous avons montré différents projets qui utilisent ce type d'analyse computationnelle dans le but de classer des images en fonction de la similarité visuelle. Nous pouvons remarquer que ces techniques offrent un moyen pour décrire et classer automatiquement des images par rapport à des variables numériques. Si la réalisation de ce processus peut nécessiter une durée plus ou moins longue, le processus est évidemment plus rapide qu'un ou plusieurs humains qui devraient réaliser la même tâche.

Rappelons que le traitement des images concerne uniquement des matrices de pixels et non pas les images en tant que telles, matrices sur lesquelles il s'agit d'appliquer un même filtre convolutif, c'est-à-dire une transformation des images qui consiste à transformer une image en une représentation des caractéristiques des images (*feature*

map)[518]. Un filtre de convolution est défini dans le but d'atteindre un objectif précis ; dans la plupart des cas, il s'agit de fonctions qui sont habituellement utilisées en *image processing* (segmentation, détection de contours, opérations morphologiques, etc.). Pour répondre à un problème que le traitement informatique essaye de résoudre, plusieurs opérations de filtrage et de réduction sont enchaînées, apprises puis appliquées sur toutes les matrices de pixels. Ce sont donc les grandeurs descriptives qui résultent de ce traitement qui rendent compte des principales caractéristiques des images (p. ex. : la prépondérance d'une couleur ou la proportion de traits horizontaux ou verticaux). Autrement dit, ces techniques remplacent l'analyse d'images par une analyse des informations quantitatives sur les images. Il faut cependant garder à l'esprit que ces techniques nécessitent un apprentissage, c'est-à-dire qu'il est nécessaire d'entraîner au préalable le réseau de neurones convolutifs (*ConvNet*) pour que celui-ci puisse traiter correctement les images en fonction du but à atteindre.

Dans les deux premières tendances stratégiques, l'organisation et les dispositifs de recherche dans les collections et archives d'images patrimoniales se basent surtout sur les métadonnées descriptives. Par conséquent, nous avons vu que ces dispositifs limitent les possibilités d'exploration par le biais de ce paradigme langagier, parce qu'ils ignorent souvent des caractéristiques matérielles, formelles, plastiques, qui sont pourtant indispensables pour bien comprendre des images. Dans cet **ultime chapitre**, nous avons montré que des projets plus récents envisagent la conception d'interfaces Web dans une toute autre perspective. D'une part, nous avons vu que des techniques de visualisation sont utilisées pour mettre en évidence des relations entre des ressources numériques. D'autre part, d'autres projets ont mis en place des techniques de visualisation qui permettent à des utilisateurs d'explorer des collections et des archives d'images numérisées en fonction de caractéristiques visuelles telles que la lumière, la couleur, les formes des images.

Nous pouvons finalement remarquer que l'évolution de ces trois tendances stratégiques converge sur une économie des larges ensembles d'images numériques qui est une synthèse provisoire de ces trois tendances. En effet, l'objectif principal de la première tendance est de proposer un dispositif de recherche et nous avons vu que cette fonctionnalité est maintenue dans la plupart des projets, jusqu'aux projets les plus récents.

[518] **Chap. 4, sect. 4.2.2.**

Nous pouvons toutefois faire ressortir deux problématiques qu'il s'agira probablement de prendre en compte pour améliorer ces instruments. D'une part, la pertinence de ces visualisations se fonde sur l'objectivité qui est attribuée à la technologie numérique et à l'automatisme de la raison computationnelle, même si nous avons vu que les spécialistes remarquent les limites du traitement et de la classification automatique d'images. D'autre part, si ces techniques sont actuellement mises en place pour montrer les performances de l'informatique dans ce domaine, elles ne permettent pas nécessairement de répondre à des questionnements précis sur la culture visuelle, mais plutôt de faire ressortir des relations de similarité visuelle qu'il faut encore interpréter et mettre en rapport avec les connaissances en vigueur sur les images et le visuel.

CONCLUSION

Somme toute, le présent travail a montré que l'impact de l'informatique sur l'étude des images et de la culture visuelle est indéniable pour au moins deux raisons. D'une part, la conversion numérique permet d'accéder à de gigantesques ressources visuelles et audiovisuelles via Internet. D'autre part, le traitement informatique est utilisé pour mettre en place des instruments d'analyse d'images.

Pourtant, à ce stade, il est difficile de proposer une conclusion, même provisoire, sur les conséquences de ce bouleversement. Dans les domaines de recherche évoqués, tous s'accordent sur le fait qu'il reste encore beaucoup à faire.

En guise de conclusion, nous proposons d'abord de revenir dans les grandes lignes sur notre démarche afin de souligner ce que l'état de l'art du présent travail a mis en évidence. Ensuite, nous réexposons les principales caractéristiques des trois tendances stratégiques que l'étude de notre corpus fait ressortir. Enfin, nous présentons quelques considérations sur les recherches futures concernant l'utilisation de la communication en ligne et de l'analyse informatique d'images.

Dans un premier temps, l'état de l'art nous a permis de regrouper quelques travaux sur l'accroissement des images (Debray ; Gervereau ; Moles), notamment à la suite de l'utilisation intensive de la reproduction photographique du patrimoine culturel (Babelon & Chastel ; Malraux ; Stiegler). Nous avons montré que les pratiques d'archivage

d'images ont d'abord pour but de rassembler en grande quantité divers supports d'images, parce que ces images sont considérées comme des preuves d'un passé révolu (Moles). Dans ce sens, une archive est généralement comprise comme un ensemble de documents qui sont conservés et organisés par une institution patrimoniale.

Dans la foulée, nous avons vu qu'un travail sur la gestion des archives d'images datant des années 2000 (Kattnig) indique que le traitement documentaire d'images est lié à des disciplines comme l'histoire de l'art ou les sciences du langage, surtout la phase du traitement intellectuel, c'est-à-dire les opérations d'analyse et de description des images. Nous avons vu qu'au départ l'analyse d'images est souvent rapprochée des travaux en iconologie (Panofsky ; Gombrich), une approche scientifique qui a pour but d'interpréter des œuvres d'art visuel en s'attachant surtout à la description des images et à l'étude du contexte de production des œuvres, notamment en étudiant des sources littéraires. En outre, ce travail mentionne également des approches en relation avec une théorie du langage : la sémiologie de l'image (Barthes) et la sémiotique (Peirce). Toutefois, nous avons souligné dès le départ que différents régimes médiatiques se sont succédé dans l'histoire du visuel et des images, de telle sorte que les archives d'images rassemblent aujourd'hui plusieurs types de supports issus de différentes époques, dont certains sont difficilement abordables d'après les méthodes de l'histoire de l'art et des sciences du langage.

D'un autre côté, nous avons vu que les pratiques d'archivage peuvent être considérées comme une pratique énonciative d'où surgit une multiplicité d'énoncés ou de discours (Foucault). Ces énoncés sont manifestés par le moyen de supports matériels (les documents, les images ou tout autre type de ressources documentaires) qui sont conservés dans les archives. Néanmoins, l'intelligibilité de ces supports n'est jamais définitivement acquise. C'est pourquoi les pratiques d'archivage ont pour objectif de conserver les supports des archives tout en maintenant les conditions d'intelligibilité nécessaires pour les interpréter. Pour que les archives restent compréhensibles, elles sont donc nécessairement en relation avec des communautés discursives qui visent à maintenir les conditions d'interprétation des énoncés de l'archive (Maingueneau ; Giannachi ; Bachimont).

Étant donné que notre propre travail s'appuie sur la sémiotique, il était important de présenter quelques considérations générales à propos de cette approche scientifique dès le départ de notre enquête. Nous avons remarqué que les discours et l'énonciation

sont progressivement devenus des concepts centraux dans la science du langage initiée par Ferdinand de Saussure et poursuivie par Louis Hjemslev. Du reste, nous nous sommes limités à remarquer que la sémiotique d'Algirdas J. Greimas a développé une théorie du langage qui a progressivement accordé une place centrale à l'énonciation (Benveniste) lorsqu'il s'agit d'analyser des discours. Les discours sont le résultat d'un acte de langage qui est manifesté par le moyen de textes qui sont, selon François Rastier, l'unité fondamentale de l'analyse sémiotique. Ainsi, un texte est analysé par rapport à un contexte de production, ou encore par rapport à d'autres textes.

Afin de définir le cadre méthodologique de notre enquête, nous avons pourtant choisi l'approche sémiotique de Jacques Fontanille. En mettant l'accent sur l'énonciation en acte, l'intérêt de cette approche est de proposer une alternative au point de vue qui considère que les textes sont les seules unités fondamentales de l'analyse sémiotique des discours. En effet, même si la cohérence des discours est comprise au niveau du texte (le texte-énoncé est alors considéré comme le résultat d'un acte d'énonciation), l'énonciation en cours se réalise avant tout au sein d'une scène pratique, où l'interprète se confronte à plusieurs formes immanentes à travers des objets-supports qui sont eux aussi influencés par des normes et des usages pratiques. En suivant cette perspective, nous avons vu que le support des textes informatiques se rapporte à deux objets-supports différents : d'une part, les composants électroniques et le codage informatique, et d'autre part, l'écran et la page-écran. Le but de notre travail est de montrer que la communication en ligne repose sur la mise en scène de cette page-écran pour construire une relation d'énonciation entre un utilisateur et une interface Web (Pignier).

Ensuite, nous avons présenté l'approche théorique du support proposée par Bruno Bachimont, approche qui s'appuie à la fois sur la philosophie et la sémiotique. D'une manière générale, cette approche théorique du support met en évidence que c'est grâce à la matérialité du support que le sens est manifesté et que les contenus sont physiquement manipulables. Néanmoins, ce support n'est jamais constitué de façon aléatoire : le support matériel doit être adéquat pour inscrire les signes constituant les textes ; de plus, il doit être intégré dans le cadre d'une utilisation pratique qui correspond à certaines normes d'usage et d'interprétation. Cet auteur affirme qu'un contenu est lié à un support d'inscription, de telle sorte que les inscriptions n'ont pas de sens en dehors d'un processus d'interprétation qui est toujours en relation avec une rationalité particulière. Chaque type de supports entraîne une rationalité qui dépend des caractéristiques matérielles du

support, mais aussi d'habitudes sociales et culturelles. Autrement dit, la production et l'interprétation du sens se fondent sur des supports matériels qui manifestent des énoncés au sein de scènes pratiques, quoique ces supports sont souvent structurés de manière formelle en vue d'usages pratiques. C'est également ce qu'indiquent les travaux récents en sémiotique de l'énonciation à travers le couple support matériel/support formel (Fontanille ; Dondero & Reyes).

La plupart des travaux récents sur l'archivage et le numérique insistent sur l'importance des changements médiatiques qui ont lieu dans de nombreux domaines de la vie, de la culture et de la science. Pourtant, la conversion numérique n'évacue pas la question de la manifestation des inscriptions sémiotiques ni celle de l'interprétation des formes et des processus sémiotiques. Comme d'autres supports, le support numérique impose par conséquent un type de rationalité. En effet, certains travaux envisagent que l'informatique comprend un potentiel réflexif, mais une réflexivité qu'il ne faut pas forcément confondre avec d'autres types de rationalité. En théorie, la principale différence entre, d'une part, la rationalité graphique et la rationalité audiovisuelle, et d'autre part, la rationalité numérique, concerne les possibilités de manipulation automatiquement des contenus, parce que ceux-ci sont convertis en données numériques (Bachimont). Ce sont uniquement ces données qui sont traitées par l'ordinateur, afin d'augmenter considérablement les possibilités de manipulations et d'analyse des contenus qui ont été numérisés. À ce propos, trois travaux nous ont permis de considérer le numérique selon une approche sémiotique et surtout de situer la réflexivité du numérique par rapport à celle du langage (Kumiko Tanaka-Ishii ; Valle & Mazzei). Dans le même sens, la réflexivité du visuel et des images repose sur différents styles de visibilité des images en fonction des époques et des contextes, c'est-à-dire qu'il s'agit de prendre en compte l'économie visuelle qui permet d'accorder du sens à ces images (Bordron).

Si nous avons d'abord rapproché l'archive de l'énonciation et du numérique, c'est pour souligner qu'il s'agit de prendre en compte la dimension sémio-pragmatique (Klinkenberg ; Odin) et l'épaisseur rhétorique (Saemmer) de la communication en ligne des archives et des collections d'images. Nous avons montré que la sémiotique de la communication en ligne doit prendre en compte la relation énonciative (Pignier) et la dimension stratégique (Fontanille) de cette communication, c'est-à-dire l'aménagement d'un objet-support dans le but de construire le point de vue de l'utilisateur d'un service

Web. Nous avons proposé d'étudier cette communication à l'aide d'une approche sémio-rhétorique qui prend en compte une réflexion critique sur les stratégies de communication en ligne des archives et des collections d'images.

Dans le chapitre suivant, nous avons remarqué que plusieurs recherches remarquent qu'il est nécessaire de prendre en compte et d'étudier ces stratégies. Dans un premier temps, nous avons présenté les principaux objectifs des humanités numériques. Dans la foulée, nous avons remarqué que plusieurs travaux en sciences humaines ont discuté l'un des avantages évidents de l'informatique : l'accessibilité à distance à des ressources documentaires (Deloche ; Howells ; Belting). La conversion des archives en ressources numériques favorise en effet l'idée d'une dématérialisation des médias (Debray). Il s'agit donc d'adapter la médiation patrimoniale à de nouvelles pratiques médiatiques et non pas de comprendre le numérique comme un dispositif transparent qui donne directement accès à des archives et des collections d'images. C'est dans cette perspective que les institutions patrimoniales sont bouleversées par les possibilités qu'apporte le numérique et l'environnement informatique (Lavédrine), notamment parce que cet environnement engage l'utilisateur des archives dans de nouvelles opérations de lecture (Brunet).

La question du patrimoine numérique est à la fois un problème d'intelligibilité et d'accessibilité. Le numérique est à la fois un dispositif technique et un principe de conversion des contenus (Treleani). Par conséquent, il est attendu que la communication en ligne comprend plusieurs types d'inscriptions et de formes sémiotiques. Nous avons donc considéré la communication en ligne comme une communication multimédia (Stockinger ; Romescu). Comme d'autres dispositifs, il s'agit chaque fois de mobiliser un objet-support dans le cadre de pratiques qui font sens pour des humains et non uniquement pour des informaticiens et des ordinateurs. Dans tous les cas, la communication en ligne se fonde sur une programmation complexe à laquelle s'ajoute la conception d'une interface visant à la fois à scénariser l'information (Cambone) et à instaurer une relation énonciative entre l'utilisateur et l'interface (Drucker).

À ce propos, nous avons remarqué l'importance de certaines normes qui standardisent la programmation de l'environnement informatique. La communication informatique se base sur un processus complexe qu'on peut répartir en trois étapes principales : le service à réaliser à l'aide du réseau informatique, le protocole, c'est-à-dire la conception et la programmation informatique pour réaliser le service et l'interface à

travers lesquels les utilisateurs manipulent les ressources numériques. Nous avons également vu que l'utilisation de cette communication informatique dépend d'une série de processus intermédiaires qui s'appuient sur des processus informatiques qui conditionnent la lisibilité d'un contenu numérique : d'une part, le codage binaire du contenu et le codage des métadonnées et des formats fichiers, d'autre part, les applications et les plateformes informatiques pour manipuler des ressources numériques (Bachimont).

C'est pourquoi l'expérience interprétative de ressources numériques repose toujours sur un dispositif de lecture qui est la seule face visible du processus d'interaction entre l'humain et la machine. Nous avons souligné que ce type de dispositifs gagnerait à être compris comme une médiation ouverte (Eco) qui aboutit sur un nouveau type d'expérience des archives et collections d'images. Cette expérience est aménagée au sein d'une mise en scène qui se fonde en partie sur le calcul et l'automatisme et en partie sur la participation active de l'utilisateur.

C'est dans ce sens que nous avons postulé que l'influence de l'informatique dans les pratiques de communication ou d'analyse des archives d'images est observable à travers l'évolution des choix stratégiques des institutions patrimoniales. Pour mettre en évidence quelques conséquences de ces choix, nous nous sommes concentrés sur le cas particulier du traitement, de la classification et de la visualisation des larges ensembles d'images issues d'institutions patrimoniales, ce qui réduit évidemment la portée et le type d'images qui ont été considérées dans le présent travail[519].

Nous avons constaté que les possibilités offertes par la communication en ligne évoluent assez vite. Qu'il s'agisse de bases d'images numériques des institutions patrimoniales qui proposent d'explorer un inventaire des images conservées sous la forme de ressources numériques, d'expositions virtuelles proposant des mises en scène de collections d'images, ou qu'il s'agisse de dispositifs de visualisation qui présentent les résultats d'une analyse computationnelle sur des milliers d'images numériques, il est toutefois attendu qu'un dispositif de communication en ligne soit proposé à l'utilisateur. C'est toujours à travers ce dispositif que l'utilisateur construit son propre parcours de navigation, même dans un environnement complètement programmé à l'avance.

[519] Sans aucun doute, il serait nécessaire de considérer ce que l'informatique d'images apporte dans d'autres domaines scientifiques ou secteurs culturels qui sont confrontés à la classification et à l'analyse d'un nombre important de ressources visuelles.

Ainsi, pour comprendre l'impact du numérique sur l'étude et la compréhension des archives et collections d'images, nous avons proposé de considérer quelques dispositifs qui sont mis en place pour créer la communication en ligne d'un patrimoine numérique constitué de larges ensembles d'images. Cependant, avant de faire ressortir trois tendances générales qui sont, selon nous, caractéristiques de l'évolution des stratégies de conversion et de diffusion de ces larges ensembles d'images numériques, il était nécessaire de considérer l'impact de l'environnement numérique sur l'évolution des méthodes d'analyse d'images. En effet, le second objectif de cette thèse a été de montrer qu'il faut prendre en compte à la fois les ruptures et les continuités entre le traitement informatique d'images et les méthodes utilisées pour analyser des images avant le numérique.

Dans un deuxième temps, nous avons présenté deux types d'approches qui ont surement été décisives dans les méthodes d'analyse d'images : d'un côté, l'iconologie, de l'autre, la sémiologie et la sémiotique.

D'une part, nous avons vu que l'iconologie propose en principe d'interpréter les images artistiques à partir de l'observation, l'étude des sources littéraires et du contexte de l'œuvre (Panofsky). Progressivement, la plupart des approches ont favorisé un changement d'orientation dans les recherches et dans la méthodologie de cette discipline (Didi-Huberman). Ce changement favorise un dialogue entre plusieurs disciplines et différents points de vue théoriques concernant l'interprétation les images et le visuel (Howells ; Gervereau ; Havelange). Cette ouverture au débat interdisciplinaire a fait ressortir l'importance accordée au langage dans les premières approches. C'est notamment ce qu'ont mis en évidence les travaux que regroupe l'expression *iconic turn* (Alloa). Sur ce sujet, nous avons retenu trois points importants. Premièrement, le visuel possède une réflexivité qui lui est propre. Cette réflexivité ne se fonde pas forcément sur langage verbal, mais plutôt sur les mécanismes et la logique propres au visuel (Hagelstein). Deuxièmement, une image est avant tout un dispositif matériel qui s'intègre toujours dans une relation avec le regard du spectateur et dans des usages pratiques (Belting). Troisièmement, l'étude de la culture visuelle doit aussi prendre en compte l'évolution des moyens de production des images et les changements des conditions de visibilité de ces images (Arasse).

D'autre part, nous avons remarqué que des travaux en sémiologie cherchent à faire ressortir le message transmis par l'image. Dans cette perspective, l'image est définie comme un message analogique (Barthes) dont il faut décrire le contenu au moyen du langage verbal (Mounin). Cependant, si la représentation est l'un des principaux pouvoirs de l'image, celle-ci est toujours manifestée par une surface-support proposant un parcours du regard à un spectateur (Marin). En suivant une autre perspective (celle de la sémiotique greimassienne), la sémiotique visuelle a proposé de distinguer deux types de lecture : la lecture figurative, qui se base sur la reconnaissance d'objets qui sont représentés dans une image ; la lecture plastique, qui se concentre sur les caractéristiques plastiques (positions, formes, couleurs) de la surface de l'image (Greimas). Nous retrouvons cet intérêt pour la dimension plastique des images dans des travaux ultérieurs en sémiotique qui, lorsqu'ils se sont intéressés à la question de la rhétorique des signes visuels, ont mis en évidence trois éléments principaux de ce type de signes : la forme, la couleur et la texture (Groupe μ). Enfin, nous avons vu que l'analyse énonciative des images met en évidence deux points : d'une part, le point de vue d'où l'on produit un énoncé visuel est un lieu d'où des choses sont vues (lieu de savoir) et à partir duquel le spectateur construit sa propre vision ; et d'autre part, les images disposent d'une réflexivité autonome du langage et de la description verbale (Dondero).

Ensuite, nous avons remarqué que le problème de la traduction des images en mots est également mis en évidence par des travaux sur l'indexation d'archives d'images (Bachimont ; Gros). Cette traduction des archives d'images en termes linguistiques a permis de développer des outils d'indexation qui classent des images en fonction de métadonnées descriptives relatives à des images. Pourtant, ce type d'indexation contraint à ne pas prendre en compte les mécanismes et la réflexivité propre au visuel ; nous l'avons vu, cette réflexivité est pourtant placée au centre des recherches plus récentes sur le visuel et les images. Pour ces travaux, l'étude contrastée de corpus d'images peut faire ressortir les mécanismes inhérents au visuel à travers différentes productions visuelles.

En effet, nous avons remarqué que, depuis les premiers travaux scientifiques en histoire de l'art (Focillon ; Warburg ; Wölfflin ; Panofsky ; Schapiro) jusqu'aux travaux récents en sciences humaines (Francastel ; Arnheim ; Chastel), plusieurs auteurs insistent sur le fait que, pour comprendre le visuel et les images, il faut prendre en compte les continuités, mais aussi les contrastes entre les formes visuelles mobilisées dans les images au fur et à mesure des époques. C'est là un point essentiel, car nous avons suggéré que

les approches récentes envisagent d'utiliser d'autres méthodes pour analyser des corpus d'images. L'ouverture des sciences humaines aux progrès dans des domaines spécifiques de l'informatique permet de développer des protocoles informatiques capables d'analyser rapidement une image, mais surtout de reproduire le processus sur une grande quantité d'images. L'un des avantages est que l'ordinateur permet aux spécialistes du visuel d'étayer leurs argumentations à l'aide de résultats statistiques issus du traitement informatique de larges corpus d'images.

Nous avons montré que les progrès en informatiques sont mis en relation avec des spécialistes du visuel et des images, ce qui aboutit à des expériences et des projets dont le but est d'utiliser l'informatique pour comprendre la culture visuelle. C'est notamment ce que remarquent certains théoriciens des médias. Si, dans un premier temps, les travaux d'Edmond Couchot ont proposé des remarques intéressantes sur la relation entre image et numérique, nous avons également mis en évidence les travaux plus récents de Lev Manovich sur les nouveaux médias et les *software studies*.

Pour Manovich, les domaines qui étudient le visuel et les images doivent prendre acte des progrès récents en informatique, notamment lorsqu'il s'agit d'analyser une grande quantité de données. Néanmoins, il faut mettre en place une réflexion critique sur l'usage de ces nouveaux médias. En s'inspirant des travaux en *computer vision*, des domaines comme la *digital art history* et la *computationnal semiotics* montrent déjà que l'informatique est actuellement utilisée dans une perspective analytique. En effet, le traitement informatique d'images offre des dispositifs pour analyser et classer des images en fonction des métadonnées descriptives ou en fonction de caractéristiques formelles et plastiques, mais à condition que ces images soient d'abord converties en données numériques. À partir d'une quantification complète de la surface des images, il s'agit uniquement d'analyser des valeurs numériques qui sont des représentations numériques des caractéristiques plastiques de cette surface. Par conséquent, le passage aux méthodes d'analyse quantitative laisse de côté la matérialité des images. De plus, le point central du traitement informatique est que celui-ci s'appuie sur l'automatisme. C'est grâce à la programmation informatique que cet automatisme est produit, afin de coordonner des opérations de calcul portant sur une très grande quantité de données numériques. Pourtant, si ces opérations ne sont pas directement accessibles pour l'utilisateur, nous avons mis en évidence que les usages du support numérique reposent toujours sur des pratiques comprenant une part d'interprétation humaine.

Rappelons que le traitement des images concerne uniquement des matrices de pixels et non pas les images en tant que telles, matrices sur lesquelles il s'agit d'appliquer un même filtre convolutif, à savoir une transformation des images qui consiste à transformer une image en une représentation numérique de ses caractéristiques (*feature map*). Un filtre de convolution est défini dans le but d'atteindre un objectif précis ; dans la plupart des cas, il s'agit de fonctions qui sont habituellement utilisées en *image processing* (segmentation, détection de contours, opérations morphologiques, etc.). Plusieurs opérations de filtrage et de réduction sont enchaînées, apprises puis appliquées sur toutes les matrices de pixels. Ce sont donc les grandeurs descriptives qui résultent de ce traitement qui rendent compte des principales caractéristiques des images (p. ex. : la prépondérance d'une couleur ou la proportion de traits horizontaux ou verticaux). Autrement dit, ces techniques remplacent l'analyse d'images par une analyse des informations quantitatives sur les images. Il faut cependant garder à l'esprit que ces techniques nécessitent un apprentissage, c'est-à-dire qu'il est nécessaire d'entraîner au préalable le réseau de neurones convolutifs (*ConvNet*) pour que celui-ci puisse traiter correctement les images en fonction du but à atteindre. Si ces techniques informatiques d'analyse d'images influencent les pratiques de manipulation et d'interprétation des images à travers des opérations computationnelles sur des données qui ne sont pas réalisables par un humain, il faut garder à l'esprit que la visée du traitement informatique d'images est d'être utilisé par un humain.

Ainsi, nous avons constaté que la principale différence entre les approches en sciences humaines et en informatique concerne la manière de caractériser l'image et le type d'opérations analytiques appliquées à l'image. Néanmoins, le traitement informatique d'images et ces approches se rejoignent sans aucun doute sur un intérêt commun : étudier des images à travers les formes visuelles qui composent des images en comparant de nombreuses images. Par contre, étant donné que ces systèmes fonctionnent à l'aide de langages de programmation et de logiciels, les spécialistes doivent faire face à des considérations techniques s'ils veulent comprendre les opérations d'analyse réalisées par le moyen d'un traitement informatique. En outre, la plupart des dispositifs se limitent au mieux à classer automatiquement des images en fonction d'un principe de similitude visuelle (similitude rudimentaire et qui est certainement critiquable) et à présenter cette classification par le moyen d'un dispositif de visualisation. Malgré les

limites du traitement informatique d'images, nous avons vu qu'il s'agit toutefois d'un début prometteur.

Somme toute, dans cet état de l'art, nous avons pris le parti de défendre une approche sémio-rhétorique lorsqu'il s'agit de considérer les modalités d'interaction entre des utilisateurs et des archives et des collections d'images numériques. Comme nous l'avons suggéré, la conversion des larges ensembles d'images est combinée à des stratégies et à des dispositifs de communication qui créent et aménagent une relation d'énonciation entre l'internaute et une interface Web. Dans tous les cas, il s'agit d'une énonciation multimédia, mais une énonciation qui permet de construire différents types d'utilisateurs en fonction de la stratégie qui est déployée.

En considérant plusieurs exemples d'interfaces Web mises en place par des institutions patrimoniales, nous avons d'abord analysé une première tendance stratégique. Quatre exemples de bases d'images numériques indiquent clairement que cette stratégie dépend d'une indexation préalable des images qui repose sur un métalangage de description des archives d'images (*POP* ; *Memobase* ; *Getty Search Gateway* ; *BALaT*). Ce modèle d'indexation contraint les dispositifs de lecture qui sont proposés par les institutions patrimoniales. En effet, le fonctionnement des bases d'images numériques est limité parce qu'il est seulement fondé sur une description verbale du contenu des images grâce à des ontologies qui rassemblent l'ensemble des notions et concepts utilisés pour indexer les documents. Dans une certaine mesure, ce choix découle des pratiques d'archivage et des outils de gestion des archives d'images qui précèdent l'utilisation de l'informatique. Avant la conversion numérique des archives, les inventaires étaient déjà utilisés dans les pratiques d'archivage afin d'établir la liste des documents possédés par l'institution et de les retrouver plus facilement à travers un instrument de recherche décrivant les ressources dans un langage contrôle. Les ressources sur les images sont encodées, structurées et sélectionnées en fonction d'informations verbales.

Le but de la première tendance stratégique est surtout de *stocker* de larges ensembles d'images numériques. Nous avons vu que cette tendance se base sur un dispositif de communication en ligne qui instrumentalise l'interface Web pour mettre en place des inventaires de leurs ressources ou des vitrines des activités de l'institution. Les

bases d'images numériques des institutions patrimoniales comprennent deux fonctionnalités principales : l'*exploration* et la *scénarisation de l'information*.

Ensuite, nous avons observé une évolution en ce qui concerne des services Web donnant accès à de larges ensembles d'images. Nous avons présenté trois projets qui s'appuient sur une volonté de renouvellement des services Web en lien avec des archives d'images (*WikimediaCommons* ; *WikiArt* ; *Europeana*). Les projets que nous avons rattachés à ce renouvellement mettent en évidence des limitations similaires à celles des premiers choix stratégiques des institutions patrimoniales. Par rapport aux stratégies liées à des institutions patrimoniales, nous avons cependant constaté quelques différences, notamment en ce qui concerne le type d'images qui constitue les ressources ou la communauté qui encadre les images et les ressources sur ces images.

Dans un cas, la stratégie se fonde sur une *coopération entre les utilisateurs* et le principe du *libre partage* des données, notamment en favorisant le modèle de l'encyclopédie sur Internet. Nous avons vu que ce sont les utilisateurs qui fournissent les ressources qui constituent la base d'images numériques, qui est toutefois organisée en fonction de catégories et de thématiques générales et qui est surveillée par une communauté d'administrateur. Ce projet renouvelle donc le type d'images qui sont accessibles à l'aide d'un service Web en fournissant une abondante médiathèque de ressources visuelles. En outre, nous avons remarqué qu'un projet parallèle propose une encyclopédie en ligne des arts visuels.

Dans l'autre cas, il s'agit de défendre l'idée d'une *coopération entre les institutions*, c'est-à-dire que la conversion numérique des ressources patrimoniales est l'occasion d'un effort de coopération entre institutions pour agglomérer et standardiser ces ressources, mais surtout pour défendre un patrimoine culturel à travers une identité européenne. Ce changement renouvelle la portée des ressources numériques concernées. Dans ce second cas, nous avons également remarqué la mise en place de *discours de vulgarisation* concernant des images. Par exemple, en intégrant des parcours de lecture qui se basent sur des textes écrits qui concernent des ressources agglomérées par une ou plusieurs institutions. D'une manière ou d'une autre, ces textes sont produits par divers auteurs, mais ils s'intègrent plus généralement dans le discours des institutions qui participent au projet. Plus généralement, ces discours sont évidemment destinés aux utilisateurs, afin de les aider à comprendre les ressources numériques. Mais, encore une fois, cette stratégie indique que la compréhension des ressources visuelles se fonde sur

une compréhension linguistique de ces ressources. Même s'ils s'inscrivent dans un renouvellement des bases d'images numériques, nous avons rattaché ces deux projets à la première tendance stratégique, sauf que nous remarquons une composante discursive supplémentaire dans la scénarisation (discours de vulgarisation). Le but est encore une fois l'agglomération et le stockage d'un ou de plusieurs larges ensembles d'images numériques.

Tendance stratégique (*But*)	*Fonctionnalités de l'interface*	*Types d'utilisateur*

Base d'images numériques (« Stocker »)	- *Exploration*
	  - Structuration des ressources
	  - Sélection/présentation des ressources
		Explorateur
		↓

	- *Scénarisation*
	  - Présentation de l'institution (missions et objectifs) et du service Web
	  - Contextualisation des ressources (métadonnées descriptives)
	  - Discours de vulgarisation
		Lecteur

Pourtant, dans un cas comme dans l'autre, l'interface Web est complètement encadrée par un même *style de visibilité* des images qui incite à les comprendre par le moyen d'un *espace qu'il faut lire*. L'environnement informatique est uniquement employé pour scénariser toute l'information véhiculée à travers des formes sémiotiques principalement héritées des pratiques d'écriture et de lecture. Même si cette deuxième tendance renouvelle parfois les fonctionnalités proposées dans les interfaces (notamment les filtres, les modalités de présentation des ressources ou les applications qui sont intégrés dans l'interface), cette stratégie favorise également une économie qui s'appuie entièrement sur un principe de *traduction en mots du visuel*. Certes, les parcours de l'utilisateur sont en relation avec une rationalité computationnelle, mais il ne s'agit pas d'envisager la compréhension des images en dehors de la rationalité graphique et d'une logique documentaire.

Il s'agit toutefois d'une étape importante dans l'évolution des stratégies de communication en ligne des larges ensembles d'images sur Internet. Ce renouvellement des bases d'images numériques a notamment permis de regrouper différentes ressources visuelles dans le but de standardiser les archives d'images et les métadonnées de plusieurs

institutions, même si ces ressources ont été rassemblées pour deux raisons différentes (d'un côté, proposer une médiathèque où des ressources en libre partage sont fournies par les internautes, de l'autre, défendre un sentiment d'identité européenne à travers l'idée d'un patrimoine culturel en commun).

Si ce renouvellement a d'abord favorisé la conception d'espaces qui s'apparentent à des médiathèques en ligne ou des blogs qui peuvent contenir des discours de vulgarisation, d'autres projets ont progressivement pris acte des caractéristiques et des possibilités spécifiques de l'environnement informatique. En effet, nous avons suggéré que les humanités numériques favorisent un contexte d'échange entre les sciences humaines et les techniques informatiques. C'est pourquoi nous avons considéré davantage cette impulsion de renouvellement stratégique en remarquant qu'il s'agit de plus en plus de concevoir des interfaces Web qui mettent l'accent sur la dimension interactive de la navigation Web (*Rijksmuseum* ; *12 Sunset* ; *Ikonik Analysis Toolkit* ; *Rekall*). L'évolution de la deuxième tendance stratégique nous a permis de remarquer certains changements importants dans l'utilisation du support numérique. Selon nous, ces changements sont à l'origine du déplacement vers la troisième tendance, celle qui concerne les dispositifs de visualisation. Ce développement de la dimension interactive est un point important en ce qui concerne la troisième tendance stratégique, parce que celle-ci a principalement pour objectif d'améliorer les possibilités d'interaction avec un large ensemble d'images numériques.

Dans les expositions virtuelles, une fonctionnalité s'ajoute aux fonctionnalités des bases d'images : l'*interaction simulée*. Parmi les services Web présentés, nous avons vu que la mise en scène du site Web est conçue comme un espace où le point de vue de l'utilisateur n'est pas seulement construit comme un espace lisible, mais aussi comme un espace à observer, à scruter, autrement dit : un espace qui donne *à voir* plutôt qu'*à lire*. L'idée est notamment d'utiliser des applications Web dans le but de *montrer autrement* les archives et les collections d'images numériques. Néanmoins, comme nous l'avons remarqué, ce changement n'évacue pas complètement le problème de la classification des images qui, nous l'avons vu, se base encore sur un système informatique qui classe uniquement des images selon leurs métadonnées descriptives. Qu'il s'agisse de fournir des informations descriptives sur des images ou de mettre en scène des expositions ou des ballades virtuelles, les séquences d'interactions sont toujours programmées pour

légitimer le discours sur les images de l'institution ou d'autres acteurs participant au projet, au service. Même si les interfaces Web changent d'aspect, le service vise toujours à proposer en premier lieu un instrument d'exploration de ressources numériques qui est fondé sur une compréhension linguistique du visuel.

Tendances stratégiques (*But*)	*Fonctionnalités de l'interface*	*Types d'utilisateur*

Base d'images numériques (« Stocker »)	- *Exploration*
	  - Structuration des ressources
	  - Sélection/présentation des ressources
		Explorateur
		↓

	- *Scénarisation*
	  - Présentation de l'institution (missions et objectifs) et du service Web
	  - Contextualisation des ressources (métadonnées descriptives)
	  - Discours de vulgarisation
		Lecteur
		↓

Expositions virtuelles (« Montrer »)	- *Exploration*
	  - Structuration des ressources
	  - Sélection/présentation des ressources
	- *Scénarisation*
	  - Présentation de l'institution (missions et objectifs) et du service Web
	  - Contextualisation des ressources (métadonnées descriptives)
	  - Discours de vulgarisation
	- *Interaction simulée*
	  - Applications Web
		Observateur

Grâce à la présentation de ces deux premières tendances (d'une part, les bases d'images numériques et médiathèques en ligne, de l'autre, les musées et les expositions virtuelles), nous avons cependant souligné une volonté de renouveler les interfaces à travers l'intégration d'applications interactives dans les services et les interfaces Web. Dans le même sens, la troisième tendance vise toujours à concevoir des interfaces pour explorer de façon interactive de larges ensembles d'images numériques, mais en transformant complètement le *style de visibilité* de ces images. Dans les deux premières tendances stratégiques, l'organisation et les dispositifs de recherche dans les collections et archives d'images patrimoniales se basent surtout sur les métadonnées descriptives. Par conséquent, nous avons vu que ces dispositifs limitent les possibilités d'exploration

parce qu'ils ignorent souvent des caractéristiques matérielles, formelles, plastiques, qui sont pourtant indispensables pour bien comprendre des images. Nous avons montré que la troisième tendance essaye progressivement de s'éloigner de cette classification d'images en fonction de métadonnées descriptives, au profit d'une classification qui se base sur les caractéristiques visuelles des images.

La troisième tendance s'inscrit dans la continuité des *cultural analytics* et de la *media visualization*, c'est-à-dire dans la continuité d'un domaine de recherche qui a pour but d'utiliser le traitement informatique pour aider à comprendre la culture visuelle à travers l'étude des médias. Encore une fois, il s'agit d'utiliser l'informatique pour explorer de larges ensembles d'images numériques, mais dans ce cas-ci nous avons remarqué un déplacement important en ce qui concerne la manière de classer et de présenter des images.

Dans la continuité de la deuxième tendance stratégique, nous avons d'abord vu que l'idée est de proposer des services Web qui favorisent la participation active de l'utilisateur, mais des services qui sont encore largement influencés par une exploration en fonction de la description verbale des images (*E.A.T. Datascape* ; *Object : Photo*). L'un des avantages de ces progrès en informatique est qu'ils offrent des instruments pour traiter des images en grand nombre, à partir de description verbale, mais aussi en fonction des éléments plastiques (formes, couleurs) des images. Cependant, nous avons montré que des projets plus récents envisagent la conception d'interfaces Web dans une tout autre perspective : il s'agit de proposer des *dispositifs de visualisation d'images*.

Nous avons constaté surtout un intérêt envers des opérations analytiques qui visent à classer automatiquement des images et à les présenter à l'aide de montages et de diagrammes d'images qui peuvent faire ressortir des relations sémantiques, mais aussi des relations de similarités formelles et plastiques entre ces images (*Phototrails* ; *Vy*). Ces opérations se basent sur un langage de programmation, tandis que les résultats du traitement sont accessibles à travers des écrans d'ordinateur et des interfaces. Dans ces processus d'analyse computationnelle, les images sont définies sous la forme de nombres, parce que les ordinateurs ne traitent pas les images telles que l'humain les perçoit.

Nous avons souligné que les diagrammes et les réseaux occupent une place centrale dans cette utilisation du numérique. La notion de diagramme peut aider à

comprendre comment l'image et d'autres représentations visuelles sont susceptibles d'être un support de travail dans le but de faire émerger de nouvelles formes pour aider le chercheur. Dans le cas de la visualisation d'un large ensemble d'images numériques, un diagramme articule les opérations computationnelles qui sont appliquées sur chaque image avec une représentation globale des relations de similarité entre toutes les images. Ces visualisations proposent un point de vue à travers un processus exploratoire schématisé, où le *raisonnement diagrammatique* peut être compris comme l'observation et la transformation des relations entre les parties et le tout d'un ensemble d'images.

Nous avons vu que ces techniques de visualisation sont utilisées pour mettre en évidence des relations entre des ressources numériques. Plusieurs projets ont ainsi l'ambition de mettre en place des services qui permettent à des utilisateurs d'explorer des collections et des archives d'images numérisées en fonction de la similitude visuelle que partagent certaines images (*Replica* ; *Pixplot*). Mais l'application du traitement informatique pour analyser des formes visuelles est cependant limitée par deux obstacles : d'une part, la reconnaissance des formes ne produit pas une interprétation, elle met seulement en évidence un système de similitudes ; d'autre part, ces instruments prennent seulement en compte des moyennes statistiques de caractéristiques visuelles, parce qu'ils sont justement programmés pour faire abstraction de la configuration réelle des éléments qui composent les images. En outre, nous avons vu que les techniques d'apprentissage mobilisées dans ce type de projet reposent encore en grande partie sur la qualité des métadonnées descriptives des images, voire sur un principe d'étiquetage (tags) qui se base sur le langage verbal.

Au départ, l'enjeu des premières stratégies institutionnelles était surtout d'adapter les anciennes pratiques d'archivage au support numérique, mais en utilisant seulement une partie des possibilités qu'offre ce support. Dans cette perspective restreinte, le dispositif de lecture est conçu comme un accès direct et objectif aux ressources, mais un dispositif où la distribution des rôles est complètement déterminée par l'institution patrimoniale, sa mission et ses objectifs. Les services Web des institutions patrimoniales visent avant tout à fournir un dispositif de recherche qui se limite à traiter des descriptions simplifiées des images. L'interaction avec le dispositif se réduit à une manipulation rudimentaire de l'interface Web : la machine sélectionne des images et les présente dans l'interface sous la forme de liste de résultats. Pourtant, à travers une volonté évidente de renouvellement des fonctionnalités des interfaces Web, nous remarquons que la

communication en ligne n'est plus seulement considérée comme la diffusion d'un stock d'images. Nous avons montré que les services Web ont de plus en plus favorisé la coopération entre les institutions, et surtout la participation active des utilisateurs par le moyen du réseau Internet. Davantage, nous avons vu que ce renouvellement stratégique a plus récemment pour but de proposer une multiplicité de points de vue sur des ressources numériques, puisque l'idée défendue est que ces dispositifs révèlent des relations précédemment inaperçues entre de grandes quantités d'images.

Depuis les années 2000-2010, nous avons montré que plusieurs projets utilisent le traitement informatique d'images dans le but de classer automatiquement des images. Même si ce traitement peut mettre en relation des images en fonction de caractéristiques descriptives et visuelles, il va sans dire que, jusqu'à présent, les caractéristiques descriptives ont été privilégiées par rapport aux caractéristiques visuelles. Néanmoins, nous avons constaté que la troisième tendance stratégique met en évidence des solutions innovantes. En effet, le but de la troisième tendance est l'*analyse*. Pour ce faire, l'interaction simulée s'appuie de plus en plus sur des algorithmes et des dispositifs de visualisation dont le but est de contraster un large ensemble d'images numériques dans un espace dynamique de représentation et d'expérimentation. Ces instruments s'appuient sur des opérations de calculs pour mettre en évidence des relations au sein d'un large ensemble d'images, tout en proposant à la fois un point de vue global et un point de vue local sur toutes les images à l'aide d'un modèle de visualisation. Ces techniques sont évidemment plus rapides qu'un ou plusieurs humains qui devraient réaliser la même tâche, mais tout notre travail a été de montrer que ces techniques n'ont pas sens en dehors d'un usage pratique. Autrement dit, les méthodes d'analyse computationnelle d'images sont susceptibles d'aboutir sur des interprétations, c'est-à-dire des connaissances sur les culturelles visuelles, si et seulement si ces méthodes sont mobilisées dans le cadre de pratiques qui permettent à l'humain de manipuler et d'utiliser les résultats du traitement informatique.

Nous pouvons finalement remarquer que l'évolution de ces trois tendances stratégiques converge sur une économie des larges ensembles d'images numériques qui est, en quelque sorte, une *synthèse provisoire* de ces trois tendances. En effet, nous y retrouvons les fonctionnalités des interfaces précédentes, ainsi qu'une fonctionnalité supplémentaire : l'*expérimentation*.

Tendances stratégiques (*But*)	*Fonctionnalités de l'interface*	*Types d'utilisateur*

Base d'images numériques (« Stocker »)	- *Exploration*
	  - Structuration des ressources
	  - Sélection/présentation des ressources
		Explorateur
		↓

	- *Scénarisation*
	  - Présentation de l'institution (missions et objectifs) et du service Web
	  - Contextualisation des ressources (métadonnées descriptives)
	  - Discours de vulgarisation
		Lecteur
		↓

Expositions virtuelles (« Montrer »)	- *Exploration*
	  - Structuration des ressources
	  - Sélection/présentation des ressources
	- *Scénarisation*
	  - Présentation de l'institution (missions et objectifs) et du service Web
	  - Contextualisation des ressources (métadonnées descriptives)
	  - Discours de vulgarisation
	- *Interaction simulée*
	  - Applications Web
		Observateur
		↓

Big visual data (« Analyser »)	- *Exploration*
	  - Structuration des ressources
	  - Sélection/présentation des ressources
	- *Scénarisation*
	  - Présentation de l'institution (missions et objectifs) et du service Web
	  - Contextualisation des ressources (métadonnées descriptives)
	  - Discours de vulgarisation
	- *Interaction simulée*
	  - Applications Web
	  - Visualisations de données et d'images
	- *Expérimentation*
	  - Analyse/comparaison d'un corpus d'images (raisonnement diagrammatique)
		Expérimentateur

Concernant les recherches futures, nous pouvons faire ressortir plusieurs problématiques qu'il faudra prendre en compte pour améliorer les dispositifs de communication en ligne et l'analyse informatique d'images.

(1) L'image est avant toute chose un support matériel. Or la matérialité des images est évidemment perdue lors de la conversion numérique. Comment s'agit-il d'envisager cette perte ? Pourrait-on même la compenser d'une quelconque manière, puisque l'image numérique se fonde sur une transformation de l'image en valeurs numériques ? Jusqu'à présent, nous pouvons remarquer que la solution est d'employer un métalangage de description qui consigne les informations sur le support original de la ressource numérique. En effet, nous avons constaté que la description des images par le moyen du langage verbal guide encore largement l'organisation et la classification des larges ensembles d'images numériques, même dans certains projets récents. Une fois de plus, ce rôle central du langage indique que la conversion numérique des archives d'images n'élimine pas le fait que celles-ci sont organisées dans le cadre de communautés et de pratiques discursives, cadre qui permet de maintenir les conditions d'intelligibilité des ressources documentaires.

Le numérique n'est pas, par conséquent, une solution pour sortir définitivement le visuel et les images de l'emprise du langage. Au contraire, plusieurs services Web se basent surtout sur des processus de classification des images en fonction de mots-clés, quoique la dernière tendance stratégique met en avant des processus de classification en fonction des caractéristiques plastiques. Dans tous les cas, étant donné que la pertinence de ces dispositifs se fonde sur l'automatisme (caractéristique principale de la raison computationnelle), nous devons adopter une position critique vis-à-vis de cette dimension automatique du traitement informatique d'images. Ce que nous voulons dire, c'est que l'automatisme peut être très utile pour classer ou analyser une grande quantité d'images, mais la médiation numérique n'est pas une traduction neutre ni une méthode infaillible pour comprendre les images et le visuel. Par conséquent, il faut bien comprendre que le traitement informatique est un outil parmi tant d'autres, mais il s'agit sans aucun doute de l'outil le plus récent, le plus actuel.

(2) Nous l'avons remarqué, le numérique est une innovation médiatique, parce qu'il offre un nouveau principe de transformation des médias en ressources numériques. De plus, ce support ouvre sur une innovation concernant les méthodes pour interpréter des ressources,

parce qu'il offre un instrument complexe d'analyse et de visualisation de ces ressources. Comme nous l'avons remarqué, ces innovations sont nécessairement en lien avec une stratégie de communication en ligne. Toutefois, la rencontre des approches en sciences humaines avec le traitement informatique est récente. Les informaticiens ont développé leurs instruments dans un certain cadre théorique et méthodologique, qui n'est pas le même que celui des approches qualitatives. Il en résulte que, si des dispositifs sont actuellement mis en place pour montrer les performances de l'informatique dans ce domaine, ces performances ne répondent pas nécessairement à des questionnements précis sur la culture visuelle. Il s'agit au mieux de faire ressortir des relations de similarité visuelle qu'il faut encore interpréter et mettre en rapport avec les connaissances en vigueur sur les images et le visuel. Néanmoins, nous avons vu que de plus en plus de projets visent à rapprocher les progrès récents en informatique de thématiques et problématiques considérées par les historiens de l'art et autres spécialistes des images.

(3) Dans la mesure du possible, il semble évident que les informaticiens cherchent à optimiser les services qu'ils proposent, même s'il est également remarquable que ces mêmes informaticiens ne sont pas davantage influencés par les recherches en sciences humaines. Les spécialistes en *computer vision* et les spécialistes du visuel devraient continuer de favoriser des approches pluridisciplinaires dont l'objectif est de comparer des corpus d'images issues d'époques différentes. C'est dans cette perspective que l'apport de l'iconologie et de la sémiotique visuelle sera déterminant, car ce sont justement ces deux approches qui peuvent fournir un point de départ solide, non pas pour développer de meilleurs instruments d'analyse, mais pour donner une orientation théorique et méthodologique à ces instruments, notamment en renouant avec les travaux sur l'analyse des formes visuelles qui prennent acte du fait que les images peuvent être analysées en fonction de la dimension plastique des images.

Nous avons présenté plusieurs stratégies dont le but est de proposer un dispositif qui sert de support pour une analyse, à condition de garder à l'esprit que le traitement informatique n'est pas une analyse complète, mais seulement une étape vers une interprétation. Pour ce faire, les stratégies les plus récentes cherchent à accorder une place plus importante au visuel dans la communication en ligne des images, notamment en proposant des expositions virtuelles qui intègrent des applications interactives, ou encore des dispositifs de visualisation qui servent de lieu d'expérimentation pour contraster un

corpus d'images. Dans le second cas, même s'il s'agit d'utiliser l'analyse informatique dans une perspective réflexive, il faut nécessairement qu'un utilisateur apporte en retour ses actions et sa propre expérience interprétative pour donner du sens aux résultats de cette analyse. En effet, il ne faut pas comprendre que l'humain, à travers sa rationalité et ses cultures, s'oppose radicalement à la technique : dans la réalité de la pratique, celle-ci n'est jamais très éloignée de ses motivations et de ses envies, c'est-à-dire : *de ce que l'humain veut réaliser par le moyen d'une technique*[520].

(4) À ce propos, il faudrait que les chercheurs en sciences humaines et les informaticiens travaillent sur le principe de similarité visuelle que plusieurs projets récents mettent en évidence à travers des dispositifs de visualisation. La nouveauté de ce principe nous oblige à accepter ses limites. D'autant plus que la mise en place de techniques de segmentation et de classification des images complètement automatisée n'est pas encore réalisable.

En conclusion de son ouvrage *La pensée visuelle*, Arnheim (1976 [1969] : 328) précise que cette pensée doit « voir dans les formes qui s'offrent au regard les images des modèles de forces qui sous-tendent notre existence – le fonctionnement des esprits, des corps ou des machines, la structure des sociétés ou des idées ». Ainsi, il serait avantageux de développer des protocoles informatiques capables d'intégrer le fait que les images comprennent une dimension autoréflexive, qui ne se fonde pas nécessairement sur la réflexivité du langage verbal, mais qui se fonde sur les modèles de forces et les formes propres au visuel. Il faudrait toutefois que les chercheurs s'accordent sur ce qu'on entend exactement par la reconnaissance des formes visuelles (*pattern recognition*), parce que ces formes sont très différentes si elles sont considérées en tant que relations de structures

[520] Tous les dispositifs techniques sont évidemment le fruit de la rationalité humaine. Et c'est uniquement par rapport à cette rationalité qu'on peut envisager les possibilités réflexives qu'offre le numérique. Notre rapport à la technique n'est donc pas statique ni sans relation avec la réflexivité humaine. Au contraire, comme le souligne Gilbert Simondon, on pourrait dire que « culture et technicité sont deux modes d'analyse, et que l'homme doit apprendre à traiter les problèmes selon ces deux processus, modes extrêmes qui permettent de saisir les limites des domaines complexes de la réalité » (Simondon, 2014 (1965) : 329). Ou, comme le remarque Jean-Pierre Séris, « la technique, dans son interaction de plus en plus étroite et constante avec les sciences sous tous leurs aspects, est une activité intellectuelle d'aussi haut niveau que la science » (Séris, 2017 (1994) : 280). Ces auteurs se réfèrent sans doute à d'autres techniques que le numérique ; mais, nous l'avons vu, le numérique peut être considéré comme un dispositif technique. Entre sens et matérialité, la technique concerne les outils, les instruments, mais aussi les contenus qui nous permettent autant d'agir sur le « monde des choses » que sur le « monde des représentations ».

ou si elles sont comprises comme appartenant à des mouvements dynamiques, c'est-à-dire aux forces qui sont actives dans l'image. Quelques traits suffisent parfois pour manifester ces structures et ces forces, qui peuvent tout autant se dissimuler sous une importante profusion de détails. La question est par conséquent la suivante : sur la base du seul plan d'expression de ces images, quelle connaissance peut-on produire à l'aide de cette reconnaissance automatique[521] des formes visuelles mobilisées dans des images ?

Le présent travail montre que pour *déchiffrer* les larges ensembles d'images numériques qui s'accumulent ces dernières décennies, il faut les intégrer dans une large réflexion critique qui chercherait à comprendre *comment* et *dans quel but rhétorique* nous les *chiffrons*. Pourtant, que s'agit-il de demander à ces procédures de segmentation et de visualisation ? Que s'agit-il de leur faire dire ? C'est décidément à ces deux questions que sont confrontées les recherches actuelles et futures. Apprendre à un ordinateur un protocole d'analyse informatique d'un large ensemble d'images n'est qu'une partie du problème, car il faut en retour tirer au clair les motivations humaines qui sont à l'origine de ces instruments d'analyse. En effet, les applications de l'automatisme et du calcul dans l'analyse des images et l'étude de la culture visuelle n'ont pas de sens en dehors d'un usage pratique, sans lequel la conversion numérique serait vouée à n'être qu'une série de nombres obscurs enfermés dans les différentes composantes électroniques et dans le codage informatique qui constituent les ordinateurs et les réseaux informatiques.

[521] L'automatisme ne doit pas nous leurrer à ce sujet : la quantification complète de la surface d'une image réduit les caractéristiques visuelles de cette image à des séries de nombres qui sont, par conséquent, une représentation restreinte de ces caractéristiques. Le numérique *augmente* les possibilités de manipulation et de traitement des images, mais en même temps il se fonde sur la *diminution* des formes visuelles et des caractéristiques plastiques qui suscitent en grande partie notre intérêt pour ces images. Il s'agit, par exemple, de segmenter l'image en régions ou zones d'intérêt. Il s'agit aussi de réduire la dimension propre des images au profit de la visualisation de corpus d'images dans une interface. Par conséquent, *observer de loin* ou *à distance* un large ensemble d'images exclut de *voir de près* chaque image, et inversement.

RÉFÉRENCES BIBLIOGRAPHIQUES ET WEBOGRAPHIE

Sources primaires : projets et sites Web

Corpus principal

12 Sunset : Exploring Ed Ruscha's Archive (https://12sunsets.getty.edu/map/...)

Belgian Art Links and Tools (http://balat.kikirpa.be/...)

Europeana project (https://www.europeana.eu/fr/collections/)

Experiment in Art and Technology Datascape (http://jiminy.medialab.sciences-po.fr/eat_datascape)

Getty Search Gateway (https://search.getty.edu/gateway/landing)

Memobase (https://memobase.ch)

Object : Photo (https://www.moma.org/interactives/objectphoto/#home)

Phototrails (http://phototrails.info/visualizations/)

Plateforme ouverte du patrimoine (https://mediathequepatrimoine.culture.gouv.fr/...)

Replica project (https://diamond.timemachine.eu/replica/...)

Rijksstudio et les applications Web du *Rijksmuseum* (https://www.rijksmuseum.nl/en/rijksstudio)

Vy (http://vy.nasjonalmuseet.no)

WikimediaCommons (https://commons.wikimedia.org/wiki/Accueil)

Corpus étendu

Banque d'images de la *Bibliothèque nationale de France* (cf. https://images.bnf.fr/#/home/)

Bibliothèque numérique de l'*Institut National de l'Histoire de l'Art* (https://bibliotheque-numerique.inha.fr/collection-toutes)

Collections en ligne de *Paris Musées* (https://www.parismuseescollections.paris.fr/fr)

Google Images (https://images.google.com)

Ikonikat

PixPlot (https://dhlab.yale.edu/projects/pixplot/)

Recognition

Rekall

TinyEye (https://tineye.com)

WikiArt (https://www.wikiart.org)

Autres projets et sites Web mentionnés

Cultaral Analytics Lab (http://lab.culturalanalytics.info)

DALL-E (https://openai.com/dall-e-3/)

Digital Art History Directory (https://dahd.hcommons.org/digital-art-history-directory/)

Google Arts & Culture (cf. https://artsandculture.google.com/explore)

Guggenheim Museum (https://www.guggenheim.org/collection-online)

ImageNet

Midjourney (https://www.midjourney.com/)

Musée national de Norvège (https://www.nasjonalmuseet.no/en/collection/)

Museum of Modern Art (https://www.moma.org/collection/)

Museum of New Zealand (https://collections.tepapa.govt.nz)

Tate Museum (https://www.tate.org.uk/art)

The British Museum (https://www.britishmuseum.org/terms-use/copyright-and-permissions/images-and-photography)

The Metropolitan Museum of Art Collection (cf. https://www.metmuseum.org/art/collection)

Universalis (https://www.universalis.fr)

Web Design Museum (https://www.webdesignmuseum.org/all-websites)

Wikipédia (https://fr.wikipedia.org/wiki/Wikipédia:Accueil_principal)

Sources secondaires : articles et monographies

AA. VV., *Photographie et patrimoine*, dans *Les Nouvelles du Patrimoine*, Bruxelles, Association des Amis de l'UNESCO en Belgique, n°117, 2007.

AA. VV., *Meta/morphoses. Les archives, bouillons de culture numérique*, dans *La Gazette des archives*, n°245, 2017.

AA. VV., *Dossier : « L'œuvre d'art entre structure et histoire » et « Greimas et la sémiotique de l'image »*, dans *La part de l'Œil*, n°32, 2019.

AAGE BRANDT, Per, « Qu'est-ce que la sémiotique ? Introduction à l'usage des non-initiés courageux », dans *Actes sémiotiques* [En ligne], n°121, 2018. [URL : https://www.unilim.fr/actes-semiotiques/5961]

ABLALI, Driss, « Hjemslev et Greimas : deux sémiotiques universelles différentes », dans *Linx*, n°44, 2001, pp. 39-53.

—, « Sémiotique et Sic : je t'aime, moi non plus », dans *Semen* [En ligne], n°23, 2007, mis en ligne le 18 janvier 2012. [URL : http://journals.openedition.org/semen/4911]

ABLALI, Driss & BERTIN, Erik (dirs), *Sociabilités numériques*, Louvain-la-Neuve, Academia, coll. « Extensions sémiotiques », 2020.

ALBERTI, Leon Battista, *De Pictura*, tr. fr. de Jean-Louis SCHEFER, Paris, Éditions Macula, 2014 [1435].

ALLOA, Emmanuel, « Changer de sens. De quelques effets du "tournant iconique" », dans *Critique*, n°759-760, 2010, pp. 647-658.

ALLOA, Emmanuel (éd.), *Penser l'image*, Dijon, Les Presses du réel, coll. « Perceptions », 2010.

—, *Penser l'image II. Anthropologies du visuel*, Dijon, Les Presses du réel, coll. « Perceptions », 2015.

—, *Penser l'image III. Comment lire les images ?*, Dijon, Les Presses du réel, coll. « Perceptions », 2017.

ANIS, Jacques & ZARA, Alessandro, « L'ordinateur support textuel, le texte informatique comme processus », dans ARABYAN & KLOCK-FONTANILLE (dirs), *L'écriture entre support et surface*, Paris, L'Harmattan, 2005, pp. 71-85.

ARASSE, Daniel, *Le détail. Pour une histoire rapprochée de la peinture*, Paris, Flammarion, coll. « Champs arts », 1996.

—, « Interpréter l'art : entre voir et savoirs », dans *Les renouvellements de l'observation dans les sciences contemporaines* [Vidéo], *Canal-U*, 2001. [URL : https://www.canal-u.tv/35165]

—, *L'homme en jeu. Les génies de la renaissance*, Paris, Hazan, coll. « Bibliothèque Hazan », 2009.

—, *L'homme en perspective. Les primitifs d'Italie*, Paris, Hazan, coll. « Les incontournables Hazan », 2019 [2008].

ARISTOTE, *Catégories. De l'interprétation. (Organon I et II)*, tr. fr. et présentation par Jules TRICOT, Paris, Vrin, coll. « Bibliothèque des textes philosophiques », 2008.

ARNHEIM, Rudolph, *La pensée visuelle*, tr. fr. de Claude NOËL & Marc LE CANNU, Paris, Flammarion, coll. « Champs », 1976.

AUMONT, Jacques, *L'image*, Paris, Armand Colin, coll. « Cinéma/arts visuels », 3e édition, 2011.

BABELON, Jean-Pierre & CHASTEL, André, *La notion de patrimoine*, Paris, Liana Levi, 1994.

BACHIMONT, Bruno, « Image et audiovisuel : la documentation entre technique et interprétation. Critiques et perspectives », dans *Documentaliste-Science de l'Information*, vol. 42, n°6, 2005, pp. 348-353.

—, *Ingénierie des connaissances et des contenus : le numérique entre ontologies et documents*, Paris, Hermès-Lavoisier, coll. « Science informatique et SHS », 2007a.

—, « Nouvelles tendances applicatives : de l'indexation à l'éditorialisation », dans GROS (dir.), *L'indexation multimédia : description et recherche automatiques*, Paris, Hermès Sciences, coll. « Traité IC2, série Signal et image », 2007b.

—, *Le sens de la technique : le numérique et le calcul*, Paris, Les Belles Lettres, coll. « Encre marine. À présent », 2010.

—, « L'archive et la massification des données : une nouvelle raison numérique », dans *La Gazette des archive*, n°245, 2017a, pp. 27-43.

—, *Patrimoine et numérique : technique et politique de la mémoire*, Bry-sur-Marne, INA, 2017b.

BADIR, Sémir, *Épistémologie sémiotique. La théorie du langage du Louis Hjemslev*, Paris, Honoré Champion, 2014.

—, « La typologie sémiotique des modalités. Une mise au point », dans *Semiotica*, n°234, 2020, pp. 79-101.

BADIR, Sémir & DONDERO, Maria Giulia (dirs), *Le Groupe μ entre rhétorique et sémiotique*, dans *Protée*, vol. 38, n°1, 2010. [URL : https://constellation.uqac.ca/id/eprint/2401/2/Vol_38_no_1.pdf]

BADIR, Sémir & COUÉGNAS, Nicolas, « Qu'est-ce que Wikipédia ? L'approche sémiotique d'un "objet scientifique non identifié" », dans ABLALI & BERTIN (dirs), *Sociabilités numériques*, Louvain-la-Neuve, Academia, coll. « Extensions sémiotiques », 2020, pp. 243-259.

BADULESCU, Cristina & DE LA VILLE, Valérie-Inés, « La médiation muséale au prisme du numérique », dans *Revue française des sciences de l'information et de la communication* [En ligne], n°16, 2019. [URL : http://journals.openedition.org/rfsic/5581]

BARDIN, Laurence, *L'analyse du contenu*, Paris, Presses universitaires de France, coll. « Quadrige Manuels », 2013 [1977].

BARDIOT, Clarisse, « Noting/Annotating. Documenting the technology-based performing arts », dans *Art press*, n°2, Paris, Art press, Les enjeux de la conservation des arts technologiques, 2009, pp. 30-36.

—, « Rekall: An environment for notation / annotation / denotation », dans *Performance Research*, n°20/6, 2015a, pp. 82-86.

—, « Video recording and documentation of the performing arts: from the annotation to the visualization of metadata, the example of the Rekall software », dans *Acoustic Space Journal*, n°14 , 2015b, pp. 159-168.

—, « Arts de la scène et Big Data. Retracer et analyser le processus de création d'un spectacle grâce à la visualisation de données », dans *H2PTM*, 2015c, pp. 231-242.

BARDIOT, Clarisse, CODUYS, Thierry, JACQUEMIN, Guillaume & MARAIS, Guillaume, « Rekall : un environnement open-source pour documenter, analyser les processus de création et simplifier la reprise des œuvres scéniques », dans *Journées d'Informatique Musicale*, AFIM, 2014, Bourges, France, 2014, pp.119-129.

BARRAL I ALTET, Xavier, *L'histoire de l'art*, Paris, Presses universitaires de France, coll. « Que sais-je ? », 2020 [1989].

BARTHES, Roland, « Le message photographique », dans *Communications*, n°1, 1961, pp. 127-138.

—, « Rhétorique de l'image », dans *Communications*, n°4, 1964, pp. 40-51
—, *La chambre claire : note sur la photographie*, Paris, Gallimard, 1980.

BASSO FOSSALI, Pierluigi & DONDERO, Maria Giulia, *Sémiotique de la photographie*, Limoges, Presses universitaires de Limoges, 2011.

BASSO FOSSALI, Pierluigi, BORDRON, Jean-François, DONDERO, Maria Giulia, KLINKENBERG, Jean-Marie, PROVENZANO, François & TORE, Gian Maria (dirs), *SIGNATA*, n°1, *Cartographie de la sémiotique actuelle*, Liège, Presses universitaires de Liège, 2010. [URL : https://journals.openedition.org/signata/272]

BASSO FOSSALI, Pierluigi, BORDRON, Jean-François, DONDERO, Maria Giulia, KLINKENBERG, Jean-Marie, PROVENZANO, François & TORE, Gian Maria (dirs), *SIGNATA*, n°2, *La sémiotique, entre autres*, Liège, Presses universitaires de Liège, 2011. [URL : https://journals.openedition.org/signata/516]

BASSO FOSSALI, Pierluigi, BORDRON, Jean-François, DONDERO, Maria Giulia, KLINKENBERG, Jean-Marie, PROVENZANO, François & TORE, Gian Maria (dirs), *SIGNATA*, n°4, *Que peut le métalangage ?*, Liège, Presses universitaires de Liège, 2013. [URL : https://journals.openedition.org/signata/539]

BATCHEN, Gregory, *Each Wild Idea*, Cambridge, MIT Press, 2001.

BAXANDALL, Michael, *Formes de l'intention. Sur l'explication historique des tableaux*, tr. fr. de Catherine FRAIXE, Nîmes, Jacqueline Chambon, 1991.

BECKER, Howard S., *Les mondes de l'art*, tr. fr. de Jeanne BOUNIORT, Paris, Champs-Flammarion, 2010 [1982].

BELTING, Hans, *Pour une anthropologie des images*, tr. fr. de Jean TORRENT, Paris, Gallimard, coll. « Le Temps des images », 2004.

BENJAMIN, Walter, *Sur la photographie*, tr. fr. de Jörn CAMBRELENG, Arles, Éditions Photosynthèses, coll. « Argentique », 2012.

BENVENISTE, Émile, *Problèmes de linguistique générale 1*, Paris, Gallimard, coll. « Tel », 1966.
—, *Problèmes de linguistique générale 2*, Paris, Gallimard, coll. « Tel », 1974.

BERRY, David M., *Understanding Digital Humanities*, Basingstoke, Palgrave Macmillan, 2012.

BERTIN, Jacques, *Sémiologie graphique : les diagrammes – les réseaux – les cartes*, Paris-La Haye, Mouton, 1967.

BERTRAND, Denis, *Précis de sémiotique littéraire*, Paris, Nathan, coll. « Fac linguistique », 2000.

BEYER, Andreas, COHN, Danièle & VLADOVA, Tania (dirs), *Esthétique et sciences de l'art*, dans *Trivium* [En ligne], n°6, 2010. [URL : https://journals.openedition.org/trivium/3599]

BIDEAULT, Maryse (dir.), *L'iconographie du Caire dans les collections patrimoniales françaises*, Paris, Publications de l'Institut national d'histoire de l'art, 2010.

BIRKIN, Jane, « Institutional Metadata and the Problem of Context », dans *Digital Culture and Society*, vol. 6, n°2, 2021.

BODE, Katherine & ARTHUR, Paul Longley, *Advancing Digital Humanities: Research, Methods, Theories*. Houndmills, Basingstoke, Palgrave Macmillan, 2014.

BOEHM, Gottfried, « Par-delà le langage ? Remarques sur la logique des images », dans *Trivium* [En ligne], n°1, 2008. [URL : http://journals.openedition.org/trivium/252]

BOEHM, Gottfried, *Wie Bilder Sinn erzugen. Die Macht des Zeigens*, Berlin, Berlin University Press, 2007.

BOEHM, Gottfried, BRANDSTETTER, Gabriele & MÜLLER, Achatz von (éds), *Figur und Figuration. Studien zu Wahrnehmung und Wissen*, Munich, Wilhelm Fink, 2007.

BORDRON, Jean-François, *L'iconicité et ses images*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2011.

—, *Image et vérité : essais sur les dimensions iconiques de la connaissance*, Liège, Presses universitaires de Liège, coll. « Sigilla », 2013.

BOULLIER, Dominique, GHITALLA, Frank, GKOUSKOU, Pergia, LE DOUARIN, Laurence & NEAU, Aurélie, *L'outre-lecture : Manipuler, (s')approprier, interpréter le Web*, Bibliothèque publique d'information-Centre Pompidou, 2003.

BOULOGNE, Arlette, JADOUL, Marie-Berthe, MASSE, Claudine, MÉNILLET, Dominique & PAPPANO, Hortensia, « Journée d'étude ADBS audiovisuel : Indexation des images et des sons: le numérique a-t-il changé les pratiques ? », dans *Documentaliste-Sciences de l'Information*, vol. 43, n°2, 2006, p. 136-142.

BOURDIEU, Pierre & DARBEL, Alain, *L'amour de l'art. Les musées européens et leurs publics*, Paris, Éditions de Minuit, coll. « Le sens commun », 1972 [1966].

BOUTAUD, Jean-Jacques (dir.), *Les applications de la sémiotique à la communication des organisations*, dans *Communication et organisation*, n°39, 2011.

BOUTAUD, Jean-Jacques & BERTHELOT-GUIET, Karine (dirs), *La vie des signes au sein de la communication : vers une sémiotique communicationnelle*, dans *Revue française des sciences de l'information et de la communication*, n°3, 2013.

BRIGGS, Asa & BURKE, Peter, *A social history of the media : from Gutenberg to the Internet*, Cambridge, Malen, Polity, 2009.

BROSSAUD, Claire & REBER, Bernard, *Humanités numériques : Volume 1, Nouvelles technologies cognitives et épistémologie*, Paris, Hermes Science Publications, 2007a.

—, *Humanités numériques : Volume 2, Socio-informatique et démocratie cognitive*, Paris, Hermes Science Publications, 2007b.

BRUCKMANN, Denis, « Conclusion. Sources matérielles et ressources numériques : inventer, réinventer », dans ROUSTAN (dir.), *La recherche dans les institutions patrimoniales : sources matérielles et ressources numériques*, Villeurbanne, Presses de l'ENSSIB, coll. « Papiers, État de l'art », 2016.

BRUNET, François, *La photographie. Histoire et contre-histoire*, Paris, Presses universitaires de France, 2017.

BURDICK, Anne, DRUCKER, Johanna, LUNENFELD, Peter, PRESNER, Todd & SCHNAPP, Jeffrey, *Digital_Humanities*, Cambridge, MIT Press, 2012. [URL : https://mitpress.mit.edu/books/digitalhumanities]

CAMBONE, Marie, « La médiation patrimoniale à l'épreuve du "numérique" : médiation patrimoniale, médiation documentaire et médiation expérientielle », dans *Revue française des sciences de l'information et de la communication*, n°16, 2019. [URL : http://journals.openedition.org/rfsic/5689]

CAPURRO, Carlotta & PLETS, Gertjan « Europeana, EDM, and the Europeanisation of Cultural Heritage Institutions », dans *Digital Culture and Society*, vol. 6, n°2, 2021.

CARBOU, Guillaume, « Pour une fondation anthropologique du rapport entre SIC et sémiotique », dans *Revue française des sciences de l'information et de la communication* [En ligne], n°3, 2013. [URL : http://journals.openedition.org/rfsic/446]

CARPO, Mario, « Big Data and the End of History », dans *DAHJ*, n°3, 2018.

CARTIER-BRESSON, Anne, *L'objet photographie. Une invention permanente*, Arles, Actes Sud, coll. « Photo Poche », 2012.

CASEMAJOR LOUSTAU, Nathalie & GELLEREAU, Michèle, « Les visites en ligne du patrimoine photographique. Quelle évolution des dispositifs de médiation ? », dans *Documentation et Bibliothèques*, vol. 55, n°4, 2009, pp. 189-199.

CASEMAJOR LOUSTAU, Nathalie, « Diversifier les figures du public : l'appropriation du patrimoine culturel sur le Web », dans *Communications* [En ligne], n°29/2, 2012. [URL : http://communication.revues.org/2709].

—, « Valorisation du patrimoine photographique : entre régime documentaire et régime artistique », dans *Culture & Musée*, n°21, 2013, pp. 45-65.

CAUNE, Jean, *Esthétique de la communication*, Paris, Presses universitaires de France, coll. « Que sais-je », 1997.

CHABANNE, Jean-Charles, « Approches interdisciplinaires des enseignements culturels et artistiques dans le cadre de l'enseignement de l'histoire des arts : un modèle complexe, une didactique impossible ? », dans *Tréma* [En ligne], n°45, 2016. [URL : http://journals.openedition.org/trema/3501]

CHABIN, Marie-Anne, *Le Management de l'archive*, Paris, Hermès, 2000.

CHARBONNEAU, Normand & ROBERT, Mario (dirs), *La gestion des archives photographiques*, Québec, Presses universitaires de l'Université du Québec, coll. « Gestion de l'information », 2001.

CHARTIER, Jean-François, CHARTRAND, Louis, PULIZZOTTO, Davide & MEUNIER, Jean-Guy, « A Data-Driven Computational Semiotics », dans *Semiotica*, n°230, 2018, pp. 16-69.

CHASTEL, André, *Fables, formes, figures I & II*, Paris, Flammarion, coll. « Idées et recherches », 1978.

—, *L'image dans le miroir*, Paris, Gallimard, 1980.

CHÂTEAU, Dominique, *Sémiotique et esthétique de l'image : théorie de l'iconicité*, Paris, L'Harmattan, 2007.

CHÂTEAU, Dominique & DARRAS, Bernard (dirs), *Arts et multimédia : L'œuvre d'art et sa reproduction à l'ère des médias interactifs*, Paris, Publications de la Sorbonne, 2000.

CHAUMIER, Jacques & GICQUEL, Florence (coll.), *Les techniques documentaires au fil de l'histoire : 1950-2000*, Paris, ADBS Éditions, coll. « Sciences de l'information », 2003.

CHAUMIER, Jacques, *Document et numérisation. Enjeux techniques, économiques, culturels et sociaux*, Paris, ADBS, 2006.

CHEN, Chun-Houh, HARDLE, Wolfgang & UNWIN, Antony (dirs), *Handbook of Data Visualization*, Berlin, Heidelberg, Springer-Verlag, coll. « Springer Handbooks of Computational Statistics », 2008.

CHEN, Chen, REN, Yuzhuo & KUO, C.-C. Jay, *Big Visual Data Analysis. Scene Classification and Geometry Labeling*, Singapour, Springer Singapore, 2016.

CHIROLLET, Jean-Claude, *L'interprétation photographique des arts. Histoire, technologies, esthétique*, Paris, L'Harmattan, coll. « Histoire et idées des Arts », 2013.

CLAVERT, Frédéric & MULLER, Caroline (dirs), *Le goût de l'archive à l'ère numérique*, dans *La Gazette des archives*, n°253, 2019.

CLOUGH, Paul, HILL, Timothy, PARAMITA, Monica Lestari & GOODALE, Paula, « Europeana: What Users Search For and Why », dans *Research and Advanced Technology for Digital Libraries. Theory and Practice of Digital Libraries*, 2017, pp. 207-219.

CODD, Edgar Frank, « A Relational Model of Data for Large Share Data Banks », dans *Communications of ACM*, vol. 13, n°6, 1970, pp. 377-387.

CŒURÉ, Sophie & DUCLERT, Vincent, *Les Archives*, Paris, La Découverte, 2011.

COHN, Danièle, « Une logique des images », dans *Critique*, n°740-741, 2009, pp. 39-48.

COLAS-BLAISE, Marion, « Mode d'existence, modalité et modalisation : les apports de la sémiotique », dans *SIGNATA*, n°13, 2022. [URL : https://doi.org/10.4000/signata.3554]

COLAS-BLAISE, Marion, PERRIN, Laurent & TORE, Gian Maria (dirs), *L'énonciation aujourd'hui. Un concept clé des sciences du langage*, Limoges, Lambert-Lucas, 2016.

COMPAGNO, Dario (éd.), *Quantitative Semiotic Analysis*, Berlin, Springer, 2018.

COQUET, Jean-Claude, « L'École de Paris », dans COQUET *et alii* (dirs), *Sémiotique : l'École de Paris*, Paris, Hachette Université, coll. « Langue, Linguistique, Communication », 1982.

COUCHOT, Edmond, *Images, de l'optique au numérique : les arts visuels et l'évolution des technologies*, Paris, Londres, Lausanne, Hermès, 1988.

COUCHOT, Edmond & HILLAIRE, Norbert, *L'art numérique : comment la technologie vient au monde de l'art*, Paris, Flammarion, coll. « Champs arts », 2003.

COUÉGNAS, Nicolas, « L'épistémologie critique de la sémiotique », dans *SIGNATA*, n°13, 2022. [URL : https://journals.openedition.org/signata/3859#tocto2n2]

COUTURE, Carol & ROUSSEAU, Jean-Yves, *Les archives du XIXe siècle : une réponse aux besoins de l'administration et de la recherche*, Montréal, Université de Montréal, 1982.

COUTURE, Carol (dir.), *Les fonctions de l'archivistique contemporaine*, Sainte-Foy, Presses de l'Université du Québec, coll. « Gestion de l'information », 1999.

CROWLEY, Elliot J. & ZISSERMAN, Andrew, « The State of the Art: Object Retrieval in Paintings using Discriminative Regions », dans VALSTAR, FRENCH & PRIDMORE (éds), *Proceedings British Machine Vision Conference 2014*, BMVA Press, 2014.

D'ANGELO, Mario, « Les langues dans la glocalisation. Le cas du français », dans *Cahiers du CEREN*, n°52, 2019, pp. 69-76.

DACOS, Marin & MOUNIER, Pierre, *Humanités numériques : État des lieux et positionnement de la recherche française dans le contexte international*, Institut Français, 2014. [URL : http://issuu.com/institut_francais/docs/if_humanites-numeriques]

DAHLGREN, Anna, HANSSON, Karin, REICHERT, Ramon & WASIELEWSKI, Amanda, « The Politics of Metadata », dans *Digital Culture and Society*, vol. 6, n°2, 2021.

DAMISCH, Hubert, *Théorie du nuage. Pour une histoire de la peinture*, Paris, Seuil, 1972.

—, *La Dénivelée. À l'épreuve de la photographie*, Paris, Seuil, coll. « Fiction & Cie », 2001.

DAVALLON, Jean, *L'Exposition à l'œuvre : Stratégies de communication et médiation symbolique*, Paris, L'Harmattan, 1999.

—, *Le don du patrimoine : une approche communicationnelle de la patrimonialisation*, Paris, Hermès-Lavoisier, 2006.

—, « Penser le patrimoine selon une perspective communicationnelle », dans *Sciences de la société*, n°99, 2016, pp. 15-29.

DAVIES, Roy, *Machine Vision – Theory, Applications & Practicalities*, Burlington, Morgan-Kaufmann, 2005.

DEBRAY, Régis & MERZEAU, Louise, « Médiasphère », dans *Médium*, n°4, 2005, pp. 162-169.

DEBRAY, Régis, *Cours de médiologie générale*, Paris, Gallimard, coll. « Bibliothèque des idées », 1991.

—, *Vie et mort de l'image. Une histoire du regard en Occident*, Paris, Gallimard, coll. « Folio essais », 1992.

DEJOB, Agnès & MOSER, Chloé, « Métamorphoses des archives, constance de l'archiviste ? », dans *La Gazette des archives*, n° 245, 2017, pp. 97-117.

DELEDALLE, Gérard, *Théorie et pratique du signe : introduction à la sémiotique de C. S. Peirce*, Paris, Payot, 1979.

DELOCHE, Bernard, *Le musée virtuel*, Paris, Presses universitaires de France, coll. « Questions », 2001.

DELSALLE, Paul, *Une histoire de l'archivistique*, Sainte-Foy, Presses de l'université du Québec, 1998.

DEMEULENAERE-DOUYÈRE, Christiane, PLOUVIER, Martine & SOUCHON, Cécile (dirs), *Des images et des mots. Les documents figurés dans les archives*, Paris, CTHS, coll. « Orientations et Méthodes », 2010.

DENG, Jia, DONG, Wei, SOCHER, Richard, LI, Li-Jia, LI, Kai & FEI-FEI, Li « ImageNet: A large-scale hierarchical image database », dans *CVPR*, 2009, pp. 2-9.

DENG, Jia, RUSSAKOVSKY, Olag, KRAUSE, Jonathan, BERNSTEIN, Michael, BERG, Alec & FEI-FEI, Li, « Scalable multi-label annotation », dans *ACM conference on human factors in computing* (CHI), 2014. [URL : http://ai.stanford.edu/~olga/papers/chi2014-MultiLabel.pdf]

DENTON, Emily, HANNA, Alex, AMIRONESEI, Razvan, SMART, Andrew & NICOLE, Hilary, « On the genealogy of machine learning datasets: A critical history of ImageNet », dans *Big Data & Society*, vol. 8/2, 2021. [URL : https://journals.sagepub.com/doi/epub/10.1177/20539517211035955]

DESPRÉS-LONNET, Marie, « L'écriture numérique du patrimoine, de l'inventaire à l'exposition : les parcours de la base Joconde », dans *Culture & Musées*, n°14, 2009, pp. 19-38.

—, « Le patrimoine culturel numérique : entre compilation et computation », dans *Culture & Musées*, n°22, 2013, pp. 71-88.

DIDEROT, Denis (dir.), *Encyclopédie ou dictionnaire raisonné des sciences, des arts et des métiers*, Paris, Gallimard-Flammarion, 1986.

DIDI-HUBERMAN, Georges, *Devant l'image. Questions posées aux fins d'une histoire l'art*, Paris, Éditions de Minuit, coll. « Critique », 1990.

—, *Ce que nous voyons, ce qui nous regarde*, Paris, Éditions de Minuit, coll. « Critique », 1992.

—, *Devant le temps. Histoire de l'art et anachronisme des images*, Paris, Éditions de Minuit, coll. « Critique », 2000.

—, « Aby Warburg et l'archive des intensité », dans *Études photographiques*, n°10, 2001, pp. 144-168.

—, *L'image ouverte. Motifs de l'incarnation dans les arts visuels*, Paris, Gallimard, coll. « Le Temps des images », 2007.

DIDI-HUBERMAN, Georges & STIEGLER, Bernd, *« Iconic Turn » et réflexion sociétale*, dans *Trivium*, n°1, 2008.

DOERR, Martin, GRADMANN, Stefan, HENNICKE, Steffen, ISAAC, Antoine, MEGHINI, Carlo & VAN DE SOMPEL, Herbert, « The Europeana Data Model (EDM) », dans *World library and information congress: 76th ifla general conference and assembly* 10-15 August 2010, Gothenburg, Sweden, 2010. [URL : https://cdn.ifla.org/past-wlic/2010/149-doerr-en.pdf]

DONAHUE, Jeff, JIA, Yangqing, VINYALS, Oriol, HOFFMAN, Judy, ZHANG, Ninf, TZENG, Eric & DARRELL, Trevor, « DeCAF: a deep convolutional activation feature for generic visual recognition », dans *International Conference on Machine Learning*, 2014, pp. 647–655.

DONDERO, Maria Giulia, « Sémiotique de l'image scientifique », dans *SIGNATA*, n°1, 2010. [URL : https://journals.openedition.org/signata/291]

—, « L'énonciation énoncée dans l'image », dans COLAS-BLAISE, TORE & PERRIN (éds), *L'énonciation aujourd'hui*, Limoges, Lambert Lucas, 2015, pp. 241-258.

—, « Du texte à la pratique : pour une sémiotique expérimentale », dans *Semiotica*, n°219, 2017a, pp. 335-356.

—, « Énonciation et modes d'existence », dans *Actes sémiotiques* [En ligne], n°120, 2017b. [URL : https://www.unilim.fr/actes-semiotiques/5871]

—, « Bruno Latour et la sémiotique de l'énonciation : fondements et évolutions », dans *Symposium : Canadian Journal of Continental Philosophy*, n°22/1, 2018a, pp. 22-40.

—, « La remédiation d'archives visuelles en vue de nouvelles iconographies : le cas de la "Media Visualiszation" de Lev Manovich », dans *INTERIN*, vol. 23 n°1, 2018b, pp. 85-107.

— (dir.), *SIGNATA*, n°10, *Image et connaissance*, Liège, Presses universitaires de Liège, 2019a. [URL : https://journals.openedition.org/signata/1966]

—, « Le geste machinique dans l'analyse et le visionnage de larges collections d'images », dans *MEI. Médiation & Information*, n°49, 2019b. [URL : https://hdl.handle.net/2268/238007]

—, *Les langages de l'image. De la peinture aux Big Visual Data*, Paris, Hermann, coll. « Cultures numériques », 2020.

—, « Composition and Decomposition in Artistic Portraits, Scientific Photography, and Deep Fake Videos », dans *Lexia*, n°38-39, 2021, pp. 439-454.

DONDERO, Maria Giulia & SONESSON, Göran, « Le Groupe µ. Quarante ans de rhétorique – Trente-trois ans de sémiotique visuelle : introduction », dans *Actes sémiotiques* [En ligne], 2010. [URL : https://www.unilim.fr/actes-semiotiques/3070]

DONDERO, Maria Giulia & FONTANILLE, Jacques, *Des images à problèmes. Le sens du visuel à l'épreuve de l'image scientifique*, Limoges, Presses universitaires de Limoges, 2012.

DONDERO, Maria Giulia & REYES, Everardo, « Les supports de l'image : de la photographie à l'image numérique », dans *Revue Française des Sciences de l'Information et de la Communication* [En ligne], n°9, 2016. [URL : https://journals.openedition.org/rfsic/2124]

DONDERO, Maria Giulia, BEYAERT-GESLIN, Anne & MOUTAT, Audrey (dirs), *Les plis du visuel. Réflexivité et énonciation dans l'image*, Limoges, Lambert-Lucas, 2017.

DOULKARIDOU, Elli, « Reframing Art History », dans *DAHJ*, n°1, 2015.

DRUCKER, Johanna, « Humanities Approaches to Interface Theory », dans *Culture Machine*, vol. 12, 2011.

—, « Is There a "Digital" Art History? », dans *Visual Resources: An International Journal of Documentation*, vol. 29, n°1-2, 2013, pp. 5-13.

—, « Performative Materiality and Theoretical Approaches to Interface », dans *Digital Humanities Quarterly*, vol. 7, n°1, 2013.

—, « Information Visualization and/as Enunciation », dans *Journal of Documentation*, vol. 73, n°5, 2017, pp. 903-916.

—, *Visualisation. L'interprétation modélisante*, tr. fr. de Marie-Mathilde BORTOLOTTI, Paris, Éditions B42, coll. « Esthétique des données », 2020.

DUBOIS, Philippe, *L'acte photographique*, Paris, Nathan, 1983.

DUGENIE, Pascal, FREIRE, Nuno & BROEDER, Daan, « Building new knowledge from distribueted scientific corpus : HERBADROP & EUROPEANA : Two concrete cas studies for exploring big archival data », dans *2017 IEEE International Conference on Big Data*, 2017.

DUVIGNAUD, Jean, *Sociologie de l'art*, Paris, Presses universitaires de France, 1972.

ECO, Umberto, *L'œuvre ouverte*, Paris, Seuil, coll. « Points », 2015 [1965].

—, *La structure absente*, Paris, Mercure de France, 1972 [1968].

—, *Lector in fabula. Le rôle du lecteur ou la coopération interprétative dans les textes narratifs*, Paris, Librairie générale française, coll. « Le livre de poche, biblio essais », 1989 [1979].

—, *Sémiotique et philosophie du langage*, Paris, Presses universitaires de France, coll. « Quadrige », 1988.

—, *Le signe. Histoire et analyse d'un concept*, Paris, Librairie générale française, coll. « Le livre de poche, biblio essais », 1992 [1988].

—, *Vertige de la liste*, Paris, Flammarion, 2009.

EDWARDS, Elizabeth & MORTON, Christopher (éds), *Photographs, Museums, Collections. Between Art and Information*, London, New-York, Bloomsbury, 2015.

EMERLING, Jae, *Photography: History and theory*, Abingdon, Routledge, 2012.

EZHILRAMAN, S. Vahini & SRINIVASAN, Sujatha, « State of the art in image processing & big data analytics: issues and challenges », dans *International Journal of Engineering & Technology*, n°7, 2018, pp. 195-199.

FABRE, Maxime, « Pour une sémiosis du devenir photographique : du régime chimique au régime numérique », dans *SIGNATA*, n°10, 2019.

—, *Photographie de presse. Régimes de croyances*, Louvain-la-Neuve, Academia, coll. « Extensions sémiotiques », 2020.

FAURE, Élie, *L'Esprit des formes. Histoire de l'art, tome 1*, Paris, Gallimard, coll. « Folio essais », 1992a.

—, *L'Esprit des formes. Histoire de l'art, tome 2*, Paris, Gallimard, coll. « Folio essais », 1992b.

FAVIER, Jean, *Les Archives*, Paris, Presses universitaires de France, coll. « Que sais-je ? », 1958.

FAYET-SCRIBE, Sylvie, « Histoire des outils de médiation de l'information », dans *Histoire de la documentation en France*, CNRS Éditions, 2000, Annexe 1, pp. 229-245.

FIELDING, Nigel G. & LEE, Raymond M., *Computer analysis and qualitative research*, Thausend Oaks, SAGE publications, 1998.

FLETCHER, Pamela, « Reflections on Digital Art History », dans *Re-Views: Field Editors' Reflections* (*College Art Association reviews*) Juin 2015. [URL : http://www.caareviews.org/reviews/2726#.Y2O8Py3pNz8]

FLOCH, Jean-Marie, *Sémiotique plastique et langage publicitaire : analyse d'une annonce de la campagne de lancement de la cigarette News*, Paris, Groupe de recherches sémio-linguistiques, 1981.

—, *Sémiotique, marketing et communication. Sous les signes, les stratégies*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2002 [1990].

FLUSSER, Vilém, *La civilisation des médias*, tr. fr. de Claude MAILLARD, Belval, Circé, 2006.

FOCILLON, Henri, *La vie des formes*, Paris, Presses universitaires de France, coll. « Quadrige », 2013 [1934].

FONTANILLE, Jacques, *Les espaces subjectifs. Introduction à la sémiotique de l'observateur : discours-peinture-cinéma*, Paris, Hachette, coll. « Langue, Linguistique, Communication », 1989.

—, *Sémiotique du visible : des mondes de lumières*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 1995.

—, *Sémiotique et littérature*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 1999.

—, « La sémiotique est-elle générative ? », dans *Linx*, n°44, 2001, pp. 107-132.

—, *Sémiotique du discours*, Limoges, Presses universitaires de Limoges, 2003 [1998].

—, *Soma et Séma. Figures du corps*, Paris, Maisonneuve & Larose, coll. « Dynamique du sens », 2004.

—, « Écritures : du support matériel au support formel », dans KLOCK-FONTANILLE & ARABYAN (dirs), *L'écriture entre support et surface*, Paris, L'Harmattan, 2005.

—, « Textes, objets, situations et formes de vie. Les niveaux de pertinence du plan de l'expression dans une sémiotique des cultures », dans BERTRAND & CONSTANTINI (dirs), *Transversalité du Sens*, Paris, Presses universitaires de Vincennes, 2007a. [URL : https://www.unilim.fr/pages_perso/jacques.fontanille/textes-pdf/CSituations_pratiques2004_06.pdf]

—, « Affichages : de la sémiotique des objets à la sémiotique des situations », dans *Actes sémiotiques* [En ligne], 2007b. [URL : https://www.unilim.fr/actes-semiotiques/1113]

—, *Pratiques sémiotiques*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2008.

—, *Corps et sens*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2011.

—, « L'énonciation pratique : exploration, schématisation, transposition », dans LECLERCQ (éd.), *Common' 14 : Communication multimodale et collaboration instrumentale*, Université de Liège, ARC COMMON, 2014.

—, *Formes de vie*, Liège, Presses universitaires de Liège, coll. « Sigilla », 2015.

—, « Archéologie et anthropologie de la dimension plastique des sémiotiques visuelles », dans *Estudos Semióticos*, vol. 19, n°2, 2023, pp. 87-120.

FONTANILLE, Jacques & ZILBERBERG, Claude, *Tension et signification*, Liège, Mardaga, 1998.

FORSYTH, David A. & PONCE, Jean , *Computer Vision – A Modern Approach*, Upper Saddle River, Prentice Hall, 2003.

FOUCAULT, Michel, *Les mots et les choses*, Gallimard, coll. « Tel », 1966.
—, *L'archéologie du savoir*, Gallimard, coll. « Tel », 1969.

FRANCASTEL, Pierre, *Art et technique aux XIX^e^ et XX^e^ siècle*, Paris, Gallimard, coll. « Tel », 1956.
—, *La Figure et le lieu : l'ordre visuel du Quattrocento*, Gallimard, 1967.

FULCHIGNONI, Enrico, *La civilisation de l'image*, Paris, Payot, 1969.

GAGNON-ARGUIN, Louise, *Typologie des documents des organisations : de la création à la conservation*, Sainte-Foy, Presses de l'université du Québec, 1998.

GALLAND, Bruno, *Les archives*, Paris, Presses universitaires de France, coll. « Que sais-je », 2016.

GALLOWAY, Alex, *The Interface Effect*, Cambridge, Polity Press, 2012.

GARNIER, François, *Thésaurus iconographique : système descriptif des représentations*, Paris, Le Léopard d'Or, 1984.

GENETTE, Gérard, *L'Œuvre de l'art*, Paris, Seuil, 2010.

GERVEREAU, Laurent, *Voir, comprendre, analyser les images*, La Découverte, coll. « Repères », 1994.
—, *Histoire du visuel au XX^e^ siècle*, Paris, Seuil, coll. « Points », 2003.

GOLD, Matthew K., *Debates in the Digital Humanities*, Minneapolis, London, University of Minnesota Press, 2012.

GOMBRICH, Ernst, *L'Art et l'illusion : psychologie de la représentation picturale*, Paris, Gallimard, 1996.

GOODFELLOW, Ian, BENGIO, Yoshua & COURVILLE, Aaron, *L'apprentissage profond*, Paris, Éditions Massot, Quantmetry, 2018.

GOODMAN, Nelson, *Langages de l'art*, tr. fr. de Jacques MORIZOT, Paris, Fayard, coll. « Pluriel », 2011 [1968].

GOODY, Jack, *La raison graphique : la domestication de la pensée sauvage*, tr. fr. de Jean BAZIN & Alban BENSA, Éditions de Minuit, coll. « Le sens commun », 1979.

GORGELS, Peter, « Rijksstudio: Make Your Own Masterpiece! », dans PROCTOR & CHERRY (éds), *Museums and the Web 2013*, 2013. [URL : https://mw2013.museumsandtheweb.com/paper/rijksstudio-make-your-own masterpiece/index.html]

GREIMAS, Algirdas Julien, *Sémantique structurale : recherche de méthode*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2002 [1966].
—, *Du sens, essais sémiotiques*, Paris, Éditions du Seuil, 1970.
—, « Sémiotique figurative et sémiotique plastique », dans *Actes Sémiotiques*, VI, n°60, 1984.

GREIMAS, Algirdas Julien & COURTES, Joseph, *Sémiotique : dictionnaire raisonné de la théorie du langage*, Paris, Hachette Université, coll. « Langue, Linguistique, Communication », 1993.

GROS, Patrick, « Description et indexation automatiques des documents multimédias : du fantasme à la réalité », dans *Documentaliste-Sciences de l'Information*, n°42/6, 2005, pp. 383-391.

GROUPE μ, *Rhétorique générale*, Paris, Seuil, coll. « Points », 1982 [1970].

—, *Traité du signe visuel. Pour une rhétorique de l'image*, Paris, Seuil, coll. « La couleur des idées », 1992.

—, *Principia Semiotica. Aux sources du sens*, Bruxelles, Les impressions nouvelles, 2015.

GUILLAUME, Paul, *Psychologie de la forme*, Paris, Flammarion, coll. « Nouvelle bibliothèque scientifique », 2000 [1929].

GUNNING, Tom, « La retouche numérique à l'index. Pour une phénoménologie de la photographie », dans *Études photographiques*, n°19, 2006, pp. 96-119.

HAGELSTEIN, Maud, « *Mnemosyne* et le *Denkraum* renaissant. Pratique du document visuel chez Aby Warburg », dans *MethIS* [En ligne], vol. 2, 2009, pp. 87-111.

—, « Iconic turn et critique du paradigme langagier », dans *Bulletin d'analyse phénoménologique*, XII/2, 2016, pp. 340-356.

—, « Henri Focillon. 1881-1943 », dans TALON-HUGON (dir.), *Les théoriciens de l'art*, Paris, Presses universitaires de France, 2017, pp. 226-227.

HALBWACHS, Maurice, *La mémoire collective*, Paris, Presses universitaires de France, coll. « Bibliothèque de sociologie contemporaine », 1968.

HAMARD, Damien, « L'archiviste et le syndrome de l'araignée : état des réseaux de l'archiviste en 2010 », dans *La Gazette des archives*, n°222, 2012, pp. 161-174.

HANSEN-LØVE, Laurence (dir.), *La philosophie de A à Z*, Paris, Hatier, coll. « De A à Z », 2007.

HARRIS, Roy, *La sémiologie de l'écriture*, Paris, CNRS Éditions, 1993.

HAVELANGE, Carl, *De l'œil du monde, une histoire du regard au seuil de la modernité*, Paris, Fayard, 1998.

—, « Le visuel entre l'oral et l'écrit (Brèves réflexions pour introduire le vers dans le fruit) », dans *Autour de l'oralité et de l'écriture* (Actes du colloque international), El Jadida, Université Chouaib Doukkali, Faculté des Lettres et des Sciences Humaines, 2007.

—, « Voir/savoir. Le partage du sensible », conférence présentée à la journée d'étude « Voir/savoir : images et visualisations scientifiques », Paris, Centre Alexander Koyré, 2012.

HE, Kaiming, ZHANG, Xiangyu, REN, Shaoqing & SUN, Jian, « Deep Residual Learning for Image Recognition », dans *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

HEBERT, Louis, *Dispositifs pour l'Analyse des Textes et des Images*, Limoges, Presses universitaires de Limoges, 2007.

—, *Dictionnaire de sémiotique*, version 15.1, dans *Dictionnaire de sémiotique* [En ligne], Rimouski (Québec), 2021. [URL : https://semiotique.org/wp-content/uploads/2021/07/Dictionnaire-sémiotique.pdf]

HEINICH, Nathalie, *Le Triple Jeu de l'art contemporain : sociologie des arts plastiques*, Paris, Éditions de Minuit, coll. « Paradoxe », 1998.

HÉNAULT, Anne & BEYAERT-GESLIN, Anne (dirs), *Ateliers de sémiotique visuelle*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2004.

HÉNAULT, Anne, *Histoire de la sémiotique*, Paris, Presses universitaires de France, coll. « Que sais-je », 1997 [1992].

HINTON, Geoffrey & VAN DER MAATEN, Laurens, « Visualizing High-Dimensional Data Using t-SNE », dans *Journal of Machine Learning Research*, vol. 9, 2008, pp. 2579–2605.

HJEMSLEV, Louis, *Prolégomènes à une théorie du langage* [suivi de] *La Structure fondamentale du langage*, tr. fr. de Una CANGER, Annick WEWER & Anne-Marie LÉONARD, Paris, Éditions de Minuit, coll. « Argument », 1971.

HOCHMAN, Nadav & MANOVICH, Lev, « A View from Above: Exploratory Visualizations of the Thomas Walther Collection », dans ABBASPOUR, DAFFNER & MORRIS HAMBOURG (éds), *Object:Photo. Modern Photographs: The Thomas Walther Collection 1909–1949*, New York, The Museum of Modern Art, 2014.

HONBA HONBA, Cédric, « Objet et pratiques sémiotiques : quelques enjeux médiatiques du support photographique », dans *SIGNATA*, n°9, 2018, pp. 551-574.

—, « Médiation, design et gestualité numériques », dans *MEI. Médiation & Information*, n°47, 2019, pp. 175-190.

—, « Sémiotique et archive d'images. Communication et médiatisation des documents d'archives photographiques », dans OUVRARD & FARGE (dir.), *Corpus audiovisuels. Quelles approches ? Quels usages ?*, Paris, Éditions des archives contemporaines, coll. « Plidam », France, 2022, pp. 125-146.

HOOLAND, Seth van, GILLET, Florence, HENGCHEN, Simon & DE WILDE, Max, *Introduction aux humanités numériques : méthodes et pratiques*, Bruxelles, De Boeck Université, 2016.

HOWELLS, Richard, *Visual Culture*, Malden, Blackwell, 2003.

HRISTOVA, Stefka, « Images as Data: Cultural Analytics and Aby Warburg's MnemosynePeer », dans *DAHJ*, n°2, 2016. [URL : https://journals.openedition.org/communication/2128#tocto1n3]

IMPETT, Leonardo & MORETTI, Franco, « Totentanz. Operationalizing Aby Warburg's Pathosformeln », dans *Literary Lab Pamphlet*, n°16, 2017.

IVINS, William M. Jr, *Prints and Visual Communication*, Cambridge, MIT Press, 1969.

JACOBI, Daniel, « Sémiotique du discours de vulgarisation scientifique », dans *Semen*, n°2, 1985, mis en ligne le 21 août 2007. [URL : http://journals.openedition.org/semen/4291]

—, *La communication scientifique. Discours, figures, modèles*, Grenoble, Presses universitaires de Grenoble, coll. « Communication, médias et sociétés », 1999.

JEANNERET, Yves & SOUCHIER, Emmanuel, « Pour une poétique de "l'écrit d'écran" », dans *Xoana*, n°6, 1999, pp. 97-107.

JEANNERET, Yves, « La prétention sémiotique dans la communication. Du stigmate au paradoxe », dans *Semen,* n°23, 2007.

JIMINEZ, Marc, *La querelle de l'art contemporain*, Paris, Gallimard, coll. « Folio essais inédit », 2005.

JOHNSON, Steve, *Interface Culture*, New York, Harper Edge, 1997.

JOLY, Martine & VANOYE, Francis, *Introduction à l'analyse de l'image*, Paris, Armand Colin, coll. « Focus. Cinéma », 3e édition, 2015.

JUANALS, Brigitte & MINEL, Jean-Luc, « La construction d'un espace patrimonial partagé dans le Web de données ouvert. Entre interopérabilité et stratégies institionnelles », dans *Communication*, vol. 34 n°1, 2016. [URL : http://journals.openedition.org/communication/6650]

JURAFSKY, Daniel & MARTIN, James H., *Speech and Language Processing*, Upper Saddle River, Pearson Prentice Hall, 2008.

KARAYEV, Sergen, TRENTACOSTE, Matthew, HAN, Helen, AGARWALA, Aseem, DARRELL, Trevor, HERTZMANN, Aaron & WINNEMOELLER, Holger, « Recognizing image style », dans *ECCV*, 2014, pp. 1-20.

KATTNIG, Cécile, *Gestion et diffusion d'un fond d'image*, Paris, Armand Colin, ADBS, coll. « 128 : information et documentation », 2005.

KITTLER, Freidrich, *Médias optiques. Cours Berlinois 1999*, tr. fr. de Anaïs CARVALHO, Tamara EBLE, Ève VAYSSIÈRE, Slaven WAELTI & Audrey RIEBER, Paris, L'Harmattan, coll. « Esthétiques allemandes », 2015.

KLINKE, Harald & SURKEMPER, Liska (éds), *International Journal for Digital Art History : Visualizing Big Image Data*, n°2, Munich, Graphentis Verlag, 2016.

KLINKE, Harald, « Big Image Data within the Big Picture of Art History », dans *International Journal for Digital Art History*, n°2, Munich, Graphentis Verlag, 2016.

KLINKENBERG, Jean-Marie, *Précis de sémiotique générale*, Paris, Seuil, coll. « Points », 1996.

—, « Entre dépendance et autonomie. Pour une définition de l'écriture dans les sciences du langage et du sens », dans *SIGNATA*, n°9, 2018, pp. 103-129.

KLINKENBERG, Jean-Marie & POLIS, Stéphane, « De la scripturologie », dans *SIGNATA*, n°9, 2018, pp. 9-56.

KÖLHER, Wolfgang, *Psychologie de la forme*, tr. fr. de Serge BRICIANER, Paris, Gallimard, coll. « Folio essai », 2000 [1929].

KUHN, Virginia, CRAIG, Alan, SIMEONE, Michael, SATHEESAN, Sandeep Puthanveetil & MARINI, Luigi, « The VAT. Enhanced Video Analysis », dans *XSEDE '15*, 2015.

KUHN, Virginia, CRAIG, Alan, FRANKLIN, Kevin, SIMEONE, Michael, ARORA, Ritu, BOCK, Dave & MARINI, Luigi, « Large Scale Video Analytics. On-demand, iterative inquiry for moving image research », dans *2012 IEEE 8th International Conference on E-Science, e-Science 2012*, 2012.

KURDI, Mohamed Zakaria, *Traitement automatique des langues et linguistique informatique 1 : parole, morphologie et syntaxe*, London, ISTE Editions, coll. « Sciences cognitives », 2017.

—, *Traitement automatique des langues et linguistique informatique 2 : sémantique, discours et applications*, London, ISTE Editions, coll. « Sciences cognitives », 2018.

LAFORET, Anne, *Le Net Art au Musée. Stratégies de conservation des œuvres en ligne*, Paris, Vrin, coll. « Questions théoriques », 2011.

LAGNEL, Jean-Marie, *Manuel de datavisualisation*, Paris, Dunod, 2021.

LATOUR, Bruno & FABBRI, Paolo, « La rhétorique de la science », dans *Actes de recherches en sciences sociales*, n°13, 1977, pp. 81-95.

LAUREL, Brenda (dir.), *Art of human-computer interface design*, Reading, Addison Wesley, 1990.

LAVALLEYE, Jacques, *Les Quatre époques de l'art flamand. Trésor artistique de Belgique : roman, gothique, renaissance, baroque*, Paris, Richard-Masse Editeur, 1947.

LAVÉDRINE, Bertrand, *Connaître et conserver les photographies anciennes*, Paris, CTHS, 2016, 3e édition.

LECLERCQ, C., « Le projet *E.A.T. Datascape* ou ce que le design numérique peut faire à l'histoire sociale de l'art », dans FÉTRO & RITZ-GUILBERT (éds), *Collecta. Des pratiques antiquaires aux humanités numérique*, Paris, École du Louvre, 2017, pp. 44-66.

LECUN, Yann, BENGIO, Yoshua & HINTON, Geoffrey, « Deep Learning », dans *Nature*, n°521, 2015.

LÉON, Jacqueline, « TAL et linguistique : application, expérimentation, instrumentatisation », dans *Éla. Études de linguistique appliquée*, n°190, 2018, pp. 195-203.

LEONE, Massimo, « Préface », dans *Lexia*, n°38-39, 2021, pp. 9-25.

LEOPOLD, Edda, « On Semantic Spaces », dans *LDV Forum*, vol. 20, 2005, pp. 63-86.

LEROI-GOURHAN, André, *Milieu et Techniques*, Paris, Albin Michel, coll. « Sciences d'aujourd'hui », 1945.

—, *Le geste et la parole. Technique et langage*, Paris, Albin Michel, coll. « Sciences d'aujourd'hui », 1964.

—, *Le geste et la parole. La mémoire et les rythmes*, Paris, Albin Michel, coll. « Sciences d'aujourd'hui », 1965.

LÉVY, Pierre, « L'hypertexte, instrument et métaphore de la communication », dans *Réseaux*, vol. 9, n°46-47, 1991, pp. 59-68.

LINDEKENS, René, *Dans l'espace de l'image : essai*, Paris, Aux amateurs de livres, 1986.

—, *Texte, image et société*, Paris, Aux amateurs de livres, 1991.

LIPOVETSKY, Gilles & SERROY, Jean, *L'écran global. Culture-médias et cinéma à l'âge hypermoderne*, Paris, Seuil, coll. « La couleur des idées », 2007.

LIU, Alan, « The Meaning of Digital Humanities », dans *Publications of the Modern Language Association*, n°128/2, 2013, pp. 409-423.

LOWE, David G., « Distinctive image features from scale-invariant keypoints », dans *International Journal of Computer Vision*, n°60/2, 2004, pp. 91–110.

LTTR13, « Figures de l'énonciation : les gestes discursifs du savoir », dans BIGLARI & SALVAN (éds), *Figures en discours*, Louvain-la-Neuve, Academia-L'Harmattan, 2016, pp. 93-116. [URL : https://orbi.uliege.be/handle/2268/194724]

LUGON, Olivier, *Le style documentaire : d'August Sander à Walker Evans, 1920-1945*, Paris, Macula, coll. « Le champ de l'image », 2001.

MAINGUENEAU, Dominique, *L'analyse du discours : introduction aux lectures de l'archive*, Paris, Hachette Université, 1991.

MALRAUX, André, *Le musée imaginaire*, Paris, Gallimard, 1965.

MANOVICH, Lev, *Le langage des nouveaux médias*, tr. fr. de Richard CREVIER, Dijon, Les presses du réel, coll. « Perceptions », 2015a [2001].

—, « Media visualization : Visual techniques for exploring large media collections », dans GATES (éd.), *The International Encyclopedia of Media Studies Vol. VI: Media Studies Futures*, Blackwell, 2012.

—, « Museum Without Walls, Art History Without Names : Visualization Methods for Humanities and Media Studies », dans RICHARDSON, VERNALLIS & HERZOG (éds), *The Oxford Handbook of Sound and Image in Digital Media*, Oxford University Press, 2013a.

—, *Software Takes Command*, New York, Bloomsbury Academic, 2013b.
—, « Data Science and Digital Art History », dans *Digital Art History Journal*, n°1, 13-35, 2015b.
—, « The Science of Culture ? Social Computing, Digital Humanities and Cultural Analytics », dans SCHÄFER & VAN ES (éds), *The Datafied Society. Studying Culture through Data*, AUP, 2017a.
—, « Cultural Data. Possibilities and limitations of the digital data universes », dans GRAU, COONES & RÜHSE (éds), *Museum and Archive on the Move. Changing Cultural Institutions in the Digital Era*, Berlin, Boston, De Gruyter, 2017b, pp. 259-276.

MARIN, Louis, « La description de l'image », dans *Communications*, n°15, 1970, pp. 186-209.
—, *Études sémiologiques. Écritures, peintures*, Paris, Klincksieck, 1971.
—, « Sémiologie de l'art », dans *Encyclopaedia Universalis*, vol. 14, Paris, 1972, pp. 863-865.
—, « L'œuvre d'art et les sciences humaines », dans *Encyclopaedia Universalis*, vol. 17, Paris, 1973, pp. 135-152.
—, « Introduction : les trois formules », dans *Le portrait du roi*, Paris, Éditions de Minuit, coll. « Le sens commun », 1981, pp. 7-22.
—, *Des pouvoirs de l'images. Gloses*, Paris, Seuil, coll. « L'ordre philosophique », 1993.
—, *De la représentation*, Paris, Seuil, Gallimard, coll. « Hautes Études », 1994.

MARTIN, Robert, *Comprendre la linguistique. Épistémologie élémentaire d'une discipline*, Paris, Presses universitaires de France, coll. « Quadrige manuels », 2018 [2002].

MARTINEZ, Ruth (dir.), *Des clichés et des clics : le poids de la photographie dans la société de l'information*, Paris, ADBS, coll. « Sciences et techniques de l'information », 2007.

MARTY, Robert, *L'Algèbre des signes. Essai de sémiotique scientifique d'après C. S. Peirce*, Amsterdam, John Benjamins, coll. « Foundations of Semiotics », 1990.

MATTELART, Armand & MATTELART, Michèle, *Histoire des théories de la communication*, Paris, La Découverte, coll. « Repères », 2018 [1995].

MATTELART, Armand, *Histoire de la société de l'information*, Paris, La Découverte, coll. « Repères », 2018 [2001].

MAYEUR, Ingrid, « Médiations informatisées et discours de savoir », dans *Cahiers de praxématique* [En ligne], 71, 2018. [URL : http://journals.openedition.org/praxematique/5186]
—, « Quand le discours de savoir se fait technodiscours. Hypertextualité, commentaires et unité textuelle du billet scientifique », dans *Corela* [En ligne], HS-33, 2020. [URL : http://journals.openedition.org/corela/11876]

MC LUHAN, Marshall, *Pour comprendre les média*, tr. fr. de Jean PARÉ, Paris, Seuil, coll. « Points », 1968.

MEHLER, Alexander, « Methodological Aspects of Computational Semiotics », dans *SEED Journal*, n°3/3, 2003, pp. 71-80.

MELOT, Michel, *Une brève histoire de l'image*, Paris, Jean-Claude Béhar Éditions, coll. « Une brève histoire », 2007.

MERLEAU-PONTY, Maurice, *L'Œil et l'Esprit*, Paris, Gallimard, 1960.
—, *Œuvres*, Paris, Gallimard, coll. « Quarto », 2010.

MERZEAU, Louise, « Ceci ne tuera pas cela », dans *Les Cahiers de médiologie*, n°6, 1998, pp. 27-39.

METZ, Christian, *Essai sur la signification au cinéma I*, Paris, Klincksieck, 1968.
—, *Langage et cinéma*, Paris, Albatros, coll. « Ça cinéma », 1992 [1971].

—, *Essai sur la signification au cinéma II*, Paris, Klincksieck, 1973.
—, *Essais sémiotiques*, Paris, Klincksieck,1984 [1977].
—, *L'énonciation impersonnelle ou le site du film*, Paris, Klincksieck, 1991.

MEUNIER, Jean-Guy, « Humanités numériques ou computationnelles : enjeux herméneutiques », dans *Sens public*, 2014.

MINARD, Charles Joseph, *Des tableaux graphiques et des cartes figuratives*, Paris, impr. de Thunot, 1862. [URL : https://gallica.bnf.fr/ark:/12148/btv1b53074870p]

MITCHEL, W. J. T., *Picture Theory. Essays on Verbal and Visual Representation*, Chicago, University of Chicago Press, 1994.

MITKOV, Ruslan (éd.), *The Oxford Handbook of Computational Linguistics*, Oxford, Oxford University Press, 2003.

MOLES, Abraham A. & VALLANCIEN, Bernard (dirs), *Communications et langages*, Paris, Éditions Gauthiers-Villars, coll. « Information et cybernétique », 1963.

MOLES, Abraham A., *L'image : communication fonctionnelle*, Tournai, Casterman, 1981.

MORETTI, Franco, *Distant Reading*, Londres, New-York, Verso, 2013.

MORRIS, Tim, *Image Processing & Computer Vision*, Londres, Palgrave Macmillan, 2004.

MOUNIER, Pierre (éd.), *Read/Write Book 2 : Une introduction aux humanités numériques*. Marseille, OpenEdition Press, 2012. [URL : http://books.openedition.org/oep/226]

MOUNIN, Georges, « Pour une sémiologie de l'image », dans *Communication & langages*, n°22, 1974, pp. 48-55.

NEUMAN, Yair, COHEN, Yochai & ASSAF, Dan, « How Do We Understand the Meaning of Connotations? A Cognitive Computational Model », dans *Semiotica*, n°205, 2015, pp. 1-16.

NGUYEN, Giang Phuong & WORRING, Marcel, « Interactive access to large image collections using similarity-based visualization », dans *Journal of Visual Languages & Computing*, n°19/2, 2008, pp. 203-224.

NOGIER, Jean-François, *Ergonomie du logiciel et design Web. Le manuel des interfaces utilisateurs*, Paris, Dunod, 2008.

NOWICKI, Joanna & OUSTINOFF, Michaël, « Introduction », dans *Hermès, La Revue*, n°49/3, 2007, pp. 9-17.

ODIN, Roger, « La question du public. Approche sémio-pragmatique », dans *Réseaux*, n°99, Paris, 2000, pp. 49-72.
—, *Les espaces de communication*, Grenoble, Presses universitaires de Grenoble, coll. « La communication en plus », 2011.

OQUAB, Maxime, BOTTOU, Leon, LAPTEV, Ivan, & SIVIC, Josef, « Learning and transferring mid-level image representations using convolutional neural networks », dans *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1717–1724. [URL : https://www.cv-foundation.org/openaccess/content_cvpr_2014/papers/Oquab_Learning_and_Transferring_2014_CVPR_paper.pdf]

PANOFSKY, Erwin, *La perspective comme forme symbolique*, Paris, Éditions de Minuit, coll. « Le sens commun », tr. fr. de Guy BALLANGÉ, 1975.

—, *L'œuvre d'art et ses significations. Essais sur les « arts visuels »*, Paris, Gallimard, coll. « Folio essais », tr. fr. de Bernard & Marthe TEYSSÈDRE,1969 [1955].

—, *Essais d'iconologie : thèmes humanistes dans l'art de la renaissance*, Paris, Gallimard, coll. « Tel », tr. fr. de Claude HERBETTE & Bernard TEYSSÈDRE, 2021 [1939].

—, *Idea. Contributions à l'histoire du concept de l'ancienne théorie de l'art*, tr. fr. de Henri JOLY, Paris, Gallimard, coll. « Tel », 1989.

PAOLUCCI, Claudio, « Quelques réflexions sur les énoncés : textes, pratiques et culture », dans *Actes Sémiotiques* [En ligne], n°113, 2010.

PASTOUREAU, Michel, *Bleu. Histoire d'une couleur*, Paris, Seuil, 2000.

—, *Noir. Histoire d'une couleur*, Paris, Seuil, 2008.

—, *Vert. Histoire d'une couleur*, Paris, Seuil, 2013.

—, *Rouge. Histoire d'une couleur*, Paris, Seuil, 2016.

—, *Jaune. Histoire d'une couleur*, Paris, Seuil, 2019.

—, *Blanc. Histoire d'une couleur*, Paris, Seuil, 2022.

PAUL, Christiane, *Digital Art*, Londres, Thames & Hudson, coll. « World of Art », 2015.

PÉDAUQUE, Roger T., *Document : Forme, signe et médium. Les reformulations du numérique*, Paris, STIC-CNRS, 2003.

PEIRCE, Charles Sanders, *Collected Papers of Charles Sanders Peirce*, 8 vol., Cambridge, Harvard University Press, 1931-1958.

PERELMAN, Chaim & OLBRECHTS-TYTECA, Lucie, *Traité de l'argumentation. La nouvelle rhétorique*, Bruxelles, Éditions de l'Université libre de Bruxelles, coll. « UBlire Fondamentaux », 2008 [1958].

PETIT, Roger, VAN OVERSTRAETEN, Daniel & COPPENS, Herman, *Terminologie archivistique en usage aux Archives de l'État en Belgique. 1 : Gestion des archives*, Bruxelles, Archives générales du royaume, coll. « Miscellanea archivistica. Manuale 16 », 1994.

PICARD, David, GOSSELIN, Philippe-Henri & GASPARD, Marie-Claude, « Challenges in Content-Based Image Indexing of Cultural Heritage Collections », dans *IEEE Signal Processing Magazine*, n°32/4, 2015, pp. 95-102.

PIGNIER, Nicole & DROUILLAT, Benoît, *Le Webdesign : Sociale expérience des interfaces web*, Paris, Hermès-Lavoisier, 2008.

PIGNIER, Nicole, « Sémiotique du Webdesign : quand la pratique appelle une sémiotique ouverte », dans *Communication & langages*, n°159, 2009, pp. 91-110. [URL : https://www.cairn.info/revue-communication-et-langages1-2009-1-page-91.htm]

PINAULT, Madeleine, *L'Encyclopédie*, Paris, PUF, coll. « Que sais-je ? », 1993.

PLAYFAIR, William, *The Commercial and Political Atlas : Representing, by Means of Stained Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure and Debts of England during the Whole of the Eighteenth Century*, 1786. [URL : https://archive.org/details/PLAYFAIRWilliam1801TheCommercialandPoliticalAtlas]

RANCIÈRE, Jacques, *Le destin des images*, Paris, La fabrique, 2003.

—, *Le spectateur émancipé*, Paris, La fabrique, 2008.

RASTIER, François, *Sémantique interprétative*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2009 [1987].

—, *Sens et textualité*, Paris, Hachette, 1989.

—, « Sur l'immanentisme en sémantique », dans *Cahiers de Linguistique Française*, 15, 1994, pp. 325-335.

—, « Entretien de François Rastier avec les étudiants du séminaire Sémiotique narrative et discursive, octobre 1998 », dans *Débats sémiotique*, vol. 6, 2000, pp. 5-15. [URL : http://www.revue-texto.net/1996-2007/Dialogues/Rastier_Quebec.html]

—, « Éléments de théorie des genres », dans *Texto !* [En ligne], 2001. [URL : http://www.revue-texto.net/Inedits/Rastier/Rastier_Elements.html]

—, *Arts et sciences du texte*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2001.

—, « Sémiotique et sciences de la culture. Une introduction », dans *Texto !* [En ligne] 2006. [URL : http://www.revue-texto.net/1996-2007/Reperes/Themes/Rastier/Rastier_Intro-Semiotique.pdf]

—, « Sémiotique et linguistique des corpus », dans *SIGNATA*, n°1, 2010, pp. 13-38.

—, *La mesure et le grain. Sémantique de corpus*, Paris, Honoré Champion, coll. « Lettres numériques », 2011.

—, « De la sémantique structurale à la sémiotique des culture », dans *Actes sémiotiques* [En ligne], 120, 2017. [URL : https://www.unilim.fr/actes-semiotiques/5734]

RASTIER, François, CAVAZZA, Marc & ABEILLÉ, Anne, *Sémantique pour l'analyse : De la linguistique à l'informatique*, Paris, Masson, coll. « Sciences cognitives », 1994.

RAZAVIAN, Ali Sharif, SULLIVAN, Josephine, MAKI, Atsuto & CARLSSON, Stefan, « A baseline for visual instance retrieval with deep convolutional networks », dans *arXiv e-prints*, 2014. [URL : https://arxiv.org/abs/1412.6574]

REBOUL-TOURÉ, Sandrine, « À la recherche de nouvelles catégories pour l'analyse du discours – quand la vulgarisation scientifique passe par les blogs », dans *SHS Web of Conferences*, n°78, 2020. [URL : https://www.shs-conferences.org/articles/shsconf/abs/2020/06/shsconf_cmlf2020_01026/shsconf_cmlf2020_01026.html]

RÉGIMBEAU, Gérard, « Clés d'accès aux images fixes : indexation et perspectives pédagogiques à partir des ressources d'Internet », dans *Spirale*, n° 28, 2001, pp. 233-251.

REY, Alain, « Encyclopédie », dans *Encylopædia Universalis* [En ligne], Paris.

REYES-GARCIA, Everardo, *The Image-Interface. Graphical Supports for Visual Information*, London, ISTE-Wiley, 2017.

REYES, Everardo & SONESSON, Göran, « New approaches to plastic language: Prolegomena to a computer-aided approach to pictorial semiotics », dans *Semiotica*, n°230, 2019, pp. 71-95.

RICŒUR, Paul, *Le conflit des interprétations : essais d'herméneutique*, Paris, Seuil, coll. « L'ordre philosophique », 1969.

—, *La métaphore vive*, Paris, Seuil, coll. « L'ordre philosophique », 1975.

RINCK, Fanny, « L'analyse linguistique des enjeux de connaissance dans le discours scientifique », dans *Revue d'anthropologie des connaissances*, vol. 4, n°3, 2010, 427-450.

ROMASCU, Alina-Elena, « Le signe multimédia. Contribution à une herméneutique des sites Web », dans *Communication*, vol. 28, n°1, 2010, pp. 226-240.

RONNEBERGER, Olaf, FISCHER, Philipp & BROX, Thomas, « U-Net: Convolutional Networks for Biomedical Image Segmentation », dans *arXiv e-prints*, 2015. [URL : https://arxiv.org/pdf/1505.04597.pdf]

ROUILLÉ, André, *La Photographie, entre document et art contemporain*, Paris, Gallimard, 2005.

ROUSTAN, Mélanie (dir.), *La recherche dans les institutions patrimoniales : sources matérielles et ressources numériques*, Villeurbanne, Presses de l'ENSSIB, coll. « Papiers, État de l'art », 2016.

SAHARIA, Chitwan, CHAN, William, SAXENA, Saurabh, LI, Lala, WHANG, Jay, DENTON, Emily, GHASEMIPOUR, Seyed Kamyar Seyed, AYAN, Burcu Karagol, MAHDAVI, S. Sara, LOPES, Rapha Gontijo, SALIMANS, Tim, « Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding », dans *arXiv e-prints*, 2022. [URL : https://arxiv.org/abs/2205.11487]

SAHLGREN, Magnus, « The Word-Space Model: Using Distributional Analysis to Represent Syntagmatic and Paradigmatic Relations between Words in High-Dimensional Vector Spaces », Stockholm, 2006. [URL : https://api.semanticscholar.org/CorpusID:11917163]

SALAH, Albert Ali & GEVERS, Theo, *Computer Analysis of Human Behavior*, London, Springer-Verlag, 2011.

SALAÜN, Jean-Michel & VANDENDORPE, Christian (dirs), *Les défis de la publication sur le web : hyperlectures, cybertextes et méta-éditions*, Lyon, Presses de l'Enssib (École nationale supérieure des sciences de l'information et des bibliothèques), coll. « Référence », 2004.

SAOUTER, Catherine, « Discours de la photographie et discours de la médiation », dans TARDY (dir.), *Les médiations documentaires des patrimoines*, Paris, L'Harmattan, coll. « Nouvelles études anthropologiques », 2014.

SARTRE, Jean-Paul, *L'imaginaire*, Paris, Gallimard, coll. « Folio essais », 1940.

SAUSSURE, Ferdinand de, *Cours de linguistique générale*, Paris, Payot, coll. « Bibliothèque scientifique Payot », 1995 [1916].

SCHAEFER, Gerald, « Interactive Navigation of Image Collections », dans KIM *et al*. (éds), *Future Generation Information Technology. FGIT 2011. Lecture Notes in Computer Science*, vol. 7105, Springer, Berlin, Heidelberg, 2011, pp. 30-36.

SCHAPIRO, Meyer, *Les mots et les images*, Paris, Macula, 2000 [1973].

—, *Style, artiste et société*, Paris, Gallimard, coll. « Tel », 1982.

SCHELBERT, Georg, « Art History in the World of Digital Humanities. Aspects of a Difficult Relationship », dans *kunsttext.de – e-Journal für Kunst- und Bildgeschichte*, n°4, 2017.

SCHELLENBERG R., Theodore, *Modern Archives: Principles and Techniques*, Chicago, Chicago University Press, 1956.

SCHMIDT, Ben, « Vector Space Models for the Digital Humanities », dans *Ben Bookworm Blog*, 2015. [URL : http://bookworm.benschmidt.org/posts/2015-10-25-Word-Embeddings.html]

SEGUIN, Benoit, STRIOLO, Carlotta, DILENARDO, Isabella & KAPLAN, Frederic, « Visual Link Retrieval in a Database of Paintings », dans HUA & JÉGOU (éds), *Computer Vision - ECCV Workshops*, 2016, pp. 753-767.

SEGUIN, Benoît, *Making large art historical photo archives searchable*, École Polytechnique Fédérale de Lausanne, thèse de doctorat, 2018.

SERIS, Jean-Pierre, *La technique*, Paris, Presses universitaires de France, coll. « Quadrige », 1994.

SHNEIDERMAN, Ben, PLAISANT, Catherine, COHEN, Maxine S., JACOBS, Steven M. & ELMQVIST, Niklas, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Londres, Pearson, Sixth Edition, 2016.

SIMON, Phil, *The Visual Organization. Data Visualization, Big Data, and the Quest for Better Decisions*, New Jersey, Wiley, 2014.

SIMONDON, Gilbert, *Communication et information*, Paris, Presses universitaires de France, 2015 [2010].

—, *Sur la technique*, Paris, Presses universitaires de France, 2014.

SIMONYAN, Karen & ZISSERMAN, Andrew, « Very deep convolutional networks for large scale image recognition », dans *arXiv e-prints*, 2014, pp. 1–10. [URL : https://arxiv.org/abs/1409.1556]

SOHIER, Estelle, LUGON, Olivier & LACOSTE, Anne (dirs), *Musées des photographies documentaires*, Paris, Genève, Macula-Transbordeur, 2017.

SOLOMON, Chris & BRECKON, Toby, *Fundamentals of Digital Image Processing. A Practical Approach with Examples in Mathlab*, Chichester, Wiley-Blackwell, 2011.

SOUMAGNAC-COLIN, Karel, « Construire l'interopérabilité culturelle autour de la documentation et de l'information dans une communauté de pratique. Le cas du domaine de l'écoconstruction », dans *Communication*, vol. 34, n°1, 2016. [URL : https://journals.openedition.org/communication/6670]

SOURIAU, Etienne & SOURIAU, Anne (dirs), *Vocabulaire d'esthétique*, Paris, Presses universitaires de France, coll. « Quadrige », 2004 (1990).

SOUTET, Olivier, *Linguistique*, Paris, Presses universitaires de France, coll. « Quadrige Manuel », 2017 [1995].

SPENCER, Lori, *Reading Writing Interfaces*, Minneapolis, University for Minnesota Press, 2014.

STIEGLER, Bernd, « "Iconic Turn" et réflexion sociétale », dans *Trivium* [En ligne], n°1, 2008. [URL : http://journals.openedition.org/trivium/308]

STOCKINGER, Peter, *Les nouveaux produits d'information : conception et sémiotique du document*, Paris, Hermes, 1999.

— (dir.), *Les archives audiovisuelles. Description, indexation et publication*, Paris, Hermes-Lavoisier, coll. « Traité des sciences et techniques de l'information », 2011.

—, *Analyse des contenus audiovisuels : Métalangage et modèle de description*, Hermès Sciences Lavoisier, coll. « Traitement de l'information », 2012.

—, « The semiotic turn in digital archives and livraries », dans *Les Cahiers du numérique*, n°11/1, 2015.

—, *Le site web comme institution numérique : une approche sémiotique et fonctionnelle*, dans *HAL* [cf. hal-01737212], 2018.

STOICHITA, Victor, *L'instauration du tableau. Métapeinture à l'aube des temps modernes*, Paris, Méridiens, Klincksieck, 1993.

STRICKLAND, Eliza, « DALL-E 2's Failures Are the Most Interesting Thing About It », dans *IEEE Spectrum*, 2022. [URL : https://spectrum.ieee.org/openai-dall-e-2]

SUH, Sang C. & ANTHONY, Thomas (éds), *Big Data and Visual Analytics*, Cham, Springer International Publishing, 2017.

SVENSSON, Patrik, « Humanities Computing as Digital Humanities », dans *Digital Humanities Quarterly*, n°3/3, 2009. http://digitalhumanities.org/dhq/vol/3/3/000065/000065.html.

TANAKA-ISHII, Kumiko, *Semiotic of Programming*, New-York, Cambridge University Press, 2010.

TARDY, Cécile, « La photographie documentaire à l'épreuve du dess(e)in informatisé : mise en contexte sur les sites Web de musée », dans *Documentation et bibliothèques*, n°56/1, 2010, pp. 11-23.

TARDY, Cécile (dir.), *Les médiations documentaires des patrimoines*, Paris, L'Harmattan, coll. « Nouvelles études anthropologiques », 2014.

TARDY, Cécile & JEANNERET, Yves (dirs), *L'écriture des médias informatisés : espaces de pratiques*, Paris, Hermès Sciences Publications, coll. « Systèmes d'information et organisations documentaires », 2007.

THOM, René & NOËL, Émile, « Pour une théorie de la morphogenèse », dans AA. VV., *Les sciences de la forme aujourd'hui*, Paris, Seuils, coll. « Points Sciences », 1994, pp. 177-188.

THÜRLEMANN, Felix, « *Blumen-Mythos* (1918) de Paul Klee », dans HÉNAULT & BEYAERT, (dirs), *Ateliers de sémiotique visuelle*, Paris, Presses universitaires de France, coll. « Formes sémiotiques », 2004 [1982].

TORE, Gian Maria, « Fin de la théorie du cinéma ? Considérations sur l'épistémologie actuelle des études en cinéma », dans BEYLOT, LE CORFF & MARIE, (éds), *Images en question. Cinéma, télévision, nouvelles images : les voies de la recherche*, 2011.

—, « La réflexivité : Une question unique, des approches et des phénomènes différents », dans *SIGNATA*, n°4, 2013. [URL : https://journals.openedition.org/signata/548].

—, « Pour une pragmatique des images – Le 'sens' de deux séquences filmiques : figuration et énonciation », dans BEYAERT-GESLIN, DONDERO & MOUTAT (éds), *Les plis du visuel – Réflexivité et énonciation dans l'image*, 2017.

TRELEANI, Matteo, *Qu'est-ce que le patrimoine numérique ? Une sémiologie de la circulation des archives*, Lormont, Éditions Bord de l'eau, coll. « UDPN », 2017.

TRICOT, André, LEMARIÉ, Julie & SAHUT, Gilles, *Le document : communication et mémoire*, Louvain-La Neuve, De Boeck Supérieur, 2016.

TSIPRAS, Dimitris, SANTURKAR, Shibani, ENGSTROM, Logan, ILYAS, Andrew & MADRY, Aleksander, « From ImageNet to Image Classification: Contextualizing Progress on Benchmarks », 2020. [URL : https://arxiv.org/abs/2005.11295]

TUFTE, Edward, *The Visual Display of Quantitative Information*, Graphics Press USA, 2001 [1983].

UPJOHN, Everard M., WINGERT, Paul S. & MALHER, Jane Gaston, *Histoire mondiale de l'art*, 6 tomes, Verviers, André Gérard, 1975.

VALLÉ, Andrea & MAZZEI Alessandro, « Sapir–Worf *vs* Boas–Jakobson. Enunciation and the Semiotics of Programming Languages », dans *Lexia*, n°27-28, pp. 505-525.

VANDENDORPE, Christian, *Du papyrus à l'hypertexte. Essai sur les mutations du texte et de la lecture*, Montréal, Paris, La Découverte, 1999.

—, « L'hypertexte : vers un nouveau rapport à la lecture et au monde », dans ANGÉ (dir.), *Les objets hypertextuels. Pratiques et usages hypermédiatiques*, Londres, ISTE editions, 2015, pp. 29-51.

VAN DEN HERIK, Jaap & POSTMA, Eric O., « Discovering the Visual Signature of Painters », dans KASABOV (éd.), *Future Directions for Intelligent Systems and Information Sciences. Studies in Fuzziness and Soft Computing*, vol. 45, 2000, pp. 129-147.

VANHOUTTE, Edward, NYHAN, Julianne & TERRAS, Melissa (éds), *Defining Digital Humanities: A Reader*, Farnham, Ashgate Publishing Ltd, 2014.

VANMALDEREN, Luc, « Un Glossaire de la sémiologie de l'image », dans *Communication et langages*, n°54, 1982, pp. 10-24.

WARBURG, Aby, *Essais florentins*, tr. fr. de Sibylle MULLER, Paris, Klincksieck, coll. « L'esprit et les formes », 2003.

WIENER, Norbert, *Cybernetics, or Control and Communication in the Animal and the Machine*, Cambridge, MIT Press, 1948.

WILDER, Kelley, *Photography and Science*, London, Reaktion Books, 2009.

WINKIN, Yves (éd.), *La nouvelle communication*, Paris, Seuil, coll. « Points », 1981.

WÖLFFLIN, Heinrich, *Réflexions sur l'histoire de l'art*, tr. fr. de Rainer ROCHLITZ, Paris, Flammarion, coll. « Champs arts », 1997.

—, *Principes fondamentaux de l'histoire de l'art*, tr. fr. de Claire & Marcel RAYMOND, Marseille, Parenthèses, coll. « eupalinos », 2017.

YANG, Jing, FAN, Jianping, HUBBAL, Daniel, GAO, Yuli, LUO Hangzai & RIBARSKY, William, « Semantic Image Browser : Bridging Information Visualization with Automated Intelligent Image Analysis », dans WONG & KEIM (éds), *2006 IEEE Symposium On Visual Analytics Science And Technology*, 2006, pp. 191-198.

YAU, Nathan, *Data visualisation : de l'extraction des données à leurs représentation graphique*, Paris, Eyrolles, tr. fr. de Xavier GUESNU, 2013.

YOUNG, Forrest W., VALERO-MORA, Pedro M. & FRIENDLY, Michael, *Visual Statistics: Seeing Data with Dynamic Interactive Graphics*, New Jersey, Wiley, 2011.

ZIMMER, Michael, « Renvois of the past, present and future: hyperlinks and the structuring of knowledge from the Encyclopédie to Web 2.0 », dans *new media & society*, n°11, 2009, 107-108.

ZWEIG, Benjamin, « Forgotten Genealogies: Brief Reflections on the History of Digital Art History », dans *DAHJ*, 1, 2015.

TABLES DES ILLUSTRATIONS

Figure 1 : Exemple d'un interprète permettant d'exécuter un programme écrit dans un langage informatique (dans cet exemple, il s'agit de Python) 177
Figure 2 : Représentation schématique d'un « élément structurant » ; à gauche un « damier de pixels » contenant une forme visuelle, à droite la représentation numérique de cette forme. 187
Figure 3 : Exemple de scatter plot (« nuage de points ») à deux variables. 198
Figure 4 : Exemple de scatter plot (3D) ; cf. https://commons.wikimedia.org/wiki/File:Scatter_plot.jpg 199
Figure 5 : Exemple d'histogramme (histogram) à deux variables ; cf. https://commons.wikimedia.org/wiki/File:Housingprice.png 199
Figure 6 : Exemple de network (« visualisation en réseaux ») composé de différents points et liaisons ; cf. https://commons.wikimedia.org/wiki/File:Social_Network_Analysis_Visualization.png 200
Figure 7 : Schéma représentant les principales opérations utilisées dans le traitement d'images à l'aide de ConvNet ; cf. https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/ 203
Figure 8 : Exemple de recherche par similitude visuelle (Google Images) ; à gauche, l'image utilisée pour la recherche, et à droite les résultats du moteur de recherche. 211
Figure 9 : Autre exemple de recherche par similitude visuelle (Google Images) 211
Figure 10 : Autre exemple de recherche par similitude visuelle (Google Images) 211
Figure 11 : Exemple de recherche en fonction des proportions de couleurs dans des images (Multicolr). 212
Figure 12 : Page d'accueil de la plateforme ouverte du patrimoine ; au milieu, le « champ de recherche », en dessous, des « hyperliens » renvoyant aux différentes bases de données regroupées sur cette plateforme. 228
Figure 13 : Suite de la page d'accueil de la plateforme ouverte du patrimoine ; différents « accès thématiques » en fonction de catégories générales (photographie, peinture et dessin, mode). 228
Figure 14 : Page d'accueil de Belgian Art Links and Tools ; cette page présente des exemples différents de ressources à chaque actualisation de la page. 229
Figure 15 : Page d'accueil de Memobase. 229
Figure 16 : Page d'accueil du Getty Search Gateway ; la page est composée d'un champ de recherche, de divers « accès thématiques » (by source, by type, highlights) et d'informations verbales sur le service Web. 230
Figure 17 : Détail de la page d'accueil de Belgian Art Links and Tools ; le logo de l'institution fonctionne comme un hyperlien qui renvoie vers le site de l'Institut royal du patrimoine artistique. 230
Figure 18 : Détail de la page d'accueil du Getty Search Gateway ; le logo fonctionne aussi comme un hyperlien qui renvoie vers le site Web du Getty. 230
Figure 19 : Rubrique « à propos » de Memobase 231
Figure 20 : Schéma typique du fonctionnement du base d'images numériques (en traits pointillés, les opérations dissimulées, et en traits pleins, les opérations visibles sur la page-écran) ; en suivant ce

schéma, l'exploration d'un large ensemble d'images est conditionné par l'utilisation de mots-clés (langage verbal). 238
Figure 21 : Présentation des résultats d'une recherche sur la plateforme ouverte du patrimoine ; la zone A comprend le « champ de recherche », la zone B regroupe les « filtres » pour « affiner » la recherche, et la zone C présente les résultats à travers une « liste ». 246
Figure 22 : Présentation des résultats d'une recherche sur le Getty Search Gateway ; la zone A comprend le champ de recherche, la zone B regroupe les filtres, et la zone C présente les résultats à travers une liste. 246
Figure 23 : Champ de recherche avancée de Memobase (zone A). 247
Figure 24 : Filtres (zone B) et liste (zone C) des résultats d'une recherche sur Memobase. 247
Figure 25 : Champ de recherche avancée dans Belgian Art Links and Tools. 248
Figure 26 : Filtres et résultats d'une recherche sur Belgian Art Links and Tools. 248
Figure 27 : Présentation d'une ressource numérique sur Belgian Art Links and Tools. 249
Figure 28 : Présentation d'une ressource numérique sur la plateforme ouverte du patrimoine. 249
Figure 29 : Détail de la présentation d'une ressource sur Memobase. 250
Figure 30 : Outil de cartographie en ligne de la plateforme ouverte du patrimoine ; par rapport aux ressources numériques disponibles sur cette plateforme, cet outil indique la quantité de ressources documentaires et les lieux où ces documents originaux sont conservés (lorsque l'utilisateur agrandit la carte, la disposition des ressources (points bleus) est redistribuée. 250
Figure 31 : Idem que figure 33. 270
Figure 33 : Page d'accueil de WikimediaCommons. 269
Figure 32 : Détail de la page d'accueil (sommaire regroupant les « catégories principales » du service). 270
Figure 34 : Exemple d'une page de catégorie (pictures of fruit). 270
Figure 35 : Cadre (situé en bas de la page de WikimediaCommons) regroupant les « images remarquables » (featured pictures) en fonction de diverses catégories générales. 271
Figure 36 : page de la catégorie Images ; dans le service WikimediaCommons, les catégories sont complétées par des « sous-catégories » (images par date, par format, par support, par sujet, etc.). 271
Figure 37 : Media Viewer ; cet outil de présentation est utilisé dans la plupart des projets de la Wikimedia Foundation, notamment Wikipédia. 272
Figure 38 : Page de la catégorie Dogs avec différentes Galleries correspondant à des sous-catégories. . 272
Figure 39 : Page de la catégorie Felis silvestris catus avec différentes Galleries correspondant à des sous-catégories 273
Figure 40 : Liste des sous-catégories de la catégorie Dogs ; à droite, quelques informations générales sur cette catégorie. 273
Figure 41 : Page d'accueil de WikiArt. 279
Figure 42 : Présentation de quelques ressources disponibles sur WikiArt. 279
Figure 43 : liste de catégories sur WikiArt ; en haut, les catégories concernent des styles et des mouvements artistiques, en bas, les catégories concernent des artistes. 280

Figure 44 : Schéma du processus et des normes de l'International Image Interoperability ; cf. https://iiif.io/get-started/how-iiif-works/. 283
Figure 45 : Information en bas de la page d'Europeana avec des informations générales sur la mission et la proposition de personnaliser la langue du site Web..... 291
Figure 46 : Exemple de présentation des résultats d'une recherche sur Europeana (« vue liste »). 291
Figure 47 : Exemple de présentation des résultats d'une recherche sur Europeana (« vue grille »)..... 292
Figure 48 : Exemple de présentation des résultats d'une recherche sur Europeana (« vue mosaïque »). 292
Figure 49 : Détail de la page de présentation d'une ressource d'Europeana avec toutes ses métadonnées descriptives..... 293
Figure 50 : Page d'accueil de la rubrique « histoires » (Stories) regroupant différentes « expositions ». 293
Figure 51 : Page d'accueil de la rubrique « collections ». 294
Figure 52 : Cette exposition comprend une rubrique qui renvoie vers de nombreuses ressources sur l'Art Nouveau. 297
Figure 53 : Exemple d'une « exposition » sur l'Art Nouveau (Europeana)..... 297
Figure 54 : Voir fig. 55..... 298
Figure 55 : Même si elle présente quelques images, l'exposition s'appuie surtout sur le discours verbal. 298
Figure 56 : Page d'accueil de Paris Musées ; remarquons l'intégration, à droite, d'un diagramme qui représente la répartition quantitative des ressources disponibles en ligne. 312
Figure 57 : Moteur de recherche avancée de Paris Musées. 312
Figure 58 : Exemple d'une exposition virtuelle proposé sur le site Paris musées. À droite, le texte de présentation de l'exposition..... 313
Figure 59 : Même exemple de parcours thématique proposé sur le site Paris musées. Cette partie reprend les musées de Paris concernés par cette exposition, quelques mots-clés qui caractérisent le sujet, la période, le thème, le domaine et les personnalités relatives à cette exposition. 313
Figure 60 : L'exposition virtuelle (Paris musées) est composée de plusieurs étapes (scénarisation). 314
Figure 61 : Page d'accueil des collections numérisées de la bibliothèque de l'Institut national d'histoire de l'art..... 314
Figure 62 : Exemple d'une exposition virtuelle (Institut national d'histoire de l'art). 315
Figure 63 : Page de présentation des collections en ligne de l'Institut national d'histoire de l'art..... 315
Figure 64 : Banque d'images de la Bibliothèque nationale de France..... 316
Figure 65 : Page d'accueil des expositions virtuelles proposées par le service Web de la Bibliothèque nationale de France. 316
Figure 66 : Page d'accueil de la Galerie de Photographie de la Bibliothèque nationale de France..... 317
Figure 67 : Exemple d'exposition virtuelle proposée par la Bibliothèque nationale de France. 317
Figure 68 : Idem que figure 67. 318
Figure 69 : Autre exemple d'exposition virtuelle (Paysages français) de la Bibliothèque nationale de France..... 318
Figure 70 : Idem que la fig. 69..... 319

Figure 71 : Idem que la fig. 69 ; dans ce cas, l'introduction est manifestée par le moyen d'une capsule vidéo. 319
Figure 72 : Page d'accueil du Rijksmuseum..... 327
Figure 73 : Moteur de recherche (search the collection) du Rijksmuseum. 327
Figure 74 : Sélection d'images sur le Rijksstudio (sélection réalisée par le Rijksmuseum). 328
Figure 75 : Sélection d'images sur le Rijksstudio (sélection réalisée par une utilisatrice). 328
Figure 76 : Page d'accueil des stories (Rijksmuseum). 329
Figure 77 : Story sur le Cygne menacé de Jan Asselyn, ca. 1640-1650, sous la forme d'une vidéo explicative (Rijksmuseum). 329
Figure 78 : Application then and now (Rijksmuseum). 330
Figure 79 : Application then and now (Rijksmuseum). 330
Figure 80 : Page d'accueil des Rijksmuseum walks..... 331
Figure 81 : Exemple de ballade virtuelle (Rijksmuseum walks). 331
Figure 82 : Exemple de « ballade virtuelle » (Rijksmuseum walks). 332
Figure 83 : Interface Web du projet 12 Sunset qui présente un double point de vue sur Sunset Boulevard en fonction de deux paramètres : la position sur la carte et l'année (ici 1976). 336
Figure 84 : Interface Web du projet 12 Sunset qui présente un double point de vue sur Sunset Boulevard en fonction de deux paramètres : la position sur la carte et l'année (ici 2007) 337
Figure 85 : Présentation d'une ressource unique dans le projet 12 Sunset. 337
Figure 86 : Présentation d'une sélection aléatoire d'images ; en haut, un filtre permet de réduire les résultats en fonction d'une période temporelle, de tags visuels ou d'éléments textuels dans l'image..... 338
Figure 87 : Page « acteur » ; à droite, une liste d'acteurs (related actors) avec lesquels un acteur particulier a été le plus souvent associé, ce qui permet de rendre compte de la coprésence de certains acteurs par rapport à une même activité..... 354
Figure 88 : Exemple de page « activité » (E.A.T. Datascape). 354
Figure 89 : Vue d'ensemble (E.A.T. Datascape) ; dans ce cas, il s'agit d'un diagramme qui représente la quantité d'activité en fonction de la durée (années). 355
Figure 90 : Graph visualization (E.A.T. Datascape) ; cette visualisation en réseau distingue plusieurs événements (les gros points avec un titre) et met en évidence les relations entre les acteurs et ces événements. L'utilisateur peut agrandir la visualisation afin d'observer et de suivre ces relations. 355
Figure 91 : Application mapping photographs ; à gauche, une rubrique informative, à droite, l'outil de cartographie en ligne couplé à un dispositif de visualisation, et en bas, la timeline (« ligne du temps »). 361
Figure 92 : Application mapping artists' lives (à gauche, une légende, à droite, le dispositif de visualisation, en bas, la ligne du temps). 362
Figure 93 : Application comparing photographs (l'outil permet de comparer la quantité de photographies produites en fonction de différentes périodes temporelles). 362
Figure 94 : Application connecting artists (cet outil permet de visualiser des relations entre un artiste et d'autres artistes, ou encore entre cet artiste et des lieux, des institutions, des expositions). 363

Figure 95 : Quatre exemples de « signatures visuelles » ; dans chaque cas, les images sont organisées en fonction de paramètres colorimétriques ou de luminosité, donc des caractéristiques plastiques des images. 366
Figure 96 : Exemple de montage d'images (Phototrails)..... 368
Figure 97 : Exemple de montage d'images (Phototrails ; vue rapprochée). 368
Figure 98 : Exemple de diagramme d'images correspondant à scatter plot ou « nuage de points », où les points sont remplacés par des reproductions miniatures des images(Phototrails). 369
Figure 99 : Exemple de diagramme d'images (vue rapprochée)..... 370
Figure 100 : Application Vy (toutes les peintures sont classées en fonction des « motifs » détectés dans les images)..... 372
Figure 101 : Application Vy (toutes les peintures sont classées en fonction du « style » des images).... 373
Figure 102 : Schéma de Seguin représentant les différentes opérations du pipeline permettant d'extraire les éléments visuels et textuels des documents numérisés (cf. Seguin, 2018 : 28)..... 387
Figure 103 : Présentation schématique détaillée du projet Replica (cf. Seguin, 2018 : 6-8) ; en plus de convertir des archives photographiques en ressources numériques et d'indexer ces ressources, nous pouvons observer que le but est de proposer une interface de recherche, et surtout d'organiser ces ressources en fonction de caractéristiques visuelles qui sont à la fois mises en évidence par la machine (détection automatique) et par les utilisateurs (corrections et modifications). 393
Figure 104 : Exemple de présentation d'une ressource (Replica ; à gauche, une reproduction numérique de l'image, à droite, les principales métadonnées descriptives)..... 395
Figure 105 : Exemple d'une recherche à partir d'une image et de la présentation des résultats par le moyen d'une grille d'images (grid) (Replica) ; le système informatique sélectionne les images qui correspondent à la requête de l'utilisateur, qui peut éventuellement filtrer les résultats en fonction d'une période temporelle. 396
Figure 106 : Sélection de 60 résultats par le moyen d'une visualisation (map visualization interface ; les images sont situées dans l'espace de visualisation en fonction de la similarité visuelle établie par le traitement informatique des images). 396
Figure 107 : Sélection de 600 résultats (map visualization interface). 397
Figure 108 : Présentation d'une ressource de la Fondazione Giorgi Cini ; à droite, quelques métadonnées descriptives..... 400
Figure 109 : Résultat d'une recherche par correspondance visuelle en fonction d'une période donnée, dans ce cas à peu près cinq siècles (plus de 360 000 résultats ; présentation grid). 400
Figure 110 : Résultat d'une recherche par correspondance visuelle en fonction d'une période donnée, dans ce cas à peu près un siècle (plus de 14 000 résultats ; présentation grid). 401
Figure 111 : Le moteur de recherche Replica permet aussi de filtrer les résultats en excluant certaines images (cf. à gauche). 401
Figure 112 : Sélection de 60 images qui résulte d'une recherche par correspondance visuelle en fonction d'une période donnée, dans ce cas à peu près trois siècles (plus de 360 000 résultats ; présentation map visualization interface). 402
Figure 113 : La Madone à la prairie de Raphael (disponible sur WikimediaCommons)..... 402

Figure 114 : Résultat d'une recherche sur Replica à partir d'une « image extérieure » (WikimediaCommons). 403
Figure 115 : Madonna d'Edvard Munch (disponible sur WikimediaCommons). 403
Figure 116 : Résultat d'une recherche sur Replica à partir d'une « image extérieure » (WikimediaCommons). 404
Figure 117 : Visualisation en réseaux qui regroupe les images en fonction de leurs similarités visuelles (PixPlot). 406
Figure 118 : Visualisation qui se fonde sur le regroupement des images en fonction de leurs métadonnées (PixPlot). 407
Figure 119 : Visualisation qui classe les images en fonction du nom du fichier, c'est-à-dire par ordre alphabétique (PixPlot). 407
Figure 120 : Autre exemple de visualisation (NewsPlot). 408
Figure 121 : Même exemple de visualisation (NewsPlot ; vue rapprochée) 409
Figure 122 : Autre exemple de visualisation (NewsPlot). 409
Figure 123 : Autre exemple de visualisation (NewsPlot) 410
Figure 124 : Autre exemple de visualisation (PixPlot) 410
Figure 125 : Autre exemple de visualisation (PixPlot) 411
Figure 126 : Exemple de visualisation en fonction de la couleur du plan (The Sensory Moving Image Archive). 412

INDEX

1

12 Sunset 334

A

accessibilité à distance 90
algorithme 135, 153, 178, 183, 192, 204, 210, 372, 383, 393, 406, 412
Alloa, E. **123**
analyse d'images..... 35, 39, 116, 132, 136, 138, 142, 146, 154, 167, 186, 188, 194, 201, 368, 378
analyse des formes visuelles..... 160
analyse informatique d'images 144, 171, 347, 374, 378, 380, 398
Arasse, D. **129**
archivage, pratiques d'archivage .. 31, 32, 38, 53, 54, 57, 58, 77, 80, 81, 84, 85, 89, 103, 104, 105, 115, 134, 158, 181, 232, 233, 241, 243, 245, 265, 286, 305
Arnheim, R. **164**
automatisme 66, 72, 102, 132, 151, 240, 309, 357, 441

B

Bachimont, B. **39**, **58**, **93**, **103**, **155**
Barthes, R. **36**, **137**
base d'images107, 223, 226, 227, 239, 241, 251, 255, 280, 281, 295, 299, 300, 301, 309, 382
base de données, système de gestion de base de données.....98, 207, 224, 257, 268, 290, 307, 321, 391
Belgian Art Links and Tools 224, 227, 232, 239, 243
Belting, H. **82**, **128**, **163**
Benveniste, É. **45**, **101**, **274**
Bildwissenschaft 127, 128
Bordron, J.-F..... **71**, **85**, **148**
Brunet, F..... **85**

C

Cambone, M. **92**
Chastel, A. **30**, **117**, **166**
classification d'images ...34, 177, 189, 193, 202, 204, 206, 275, 364, 369, 372, 381
codage numérique, langage de programmation 96, 107, 175, 178
communauté discursive..... 38, 41, 53, 157, 158, 231, 276, 301, 333
computer vision 193, 194, 202, 367, 370, 382, 390, 391, 406
conditions de visibilité, style de visibilité 71, 77, 86, 102, 105, 106, 129, 130, 139, 149, 173, 237, 241, 251, 263, 275, 287, 306, 309, 333, 347, 363, 364, 397, 408
contenu (inscription sémiotique) 59
Couchot, E. **173**
culture visuelle 81, 121, 233, 349, 364

D

data visualization, *media visualization* . 195, 364
Debray, R. **29**, **82**
deep learning..... 201, 371, 383, 390, 405
Deloche, B. **83**
diagramme ... 101, 196, 200, 309, 360, 366, 367, 371
Didi-Huberman, G. **119**, **125**, **129**
digital art history..... 131
digital humanities 78
dimension plastique, sémiotique plastique ... 142
discours, énoncé-discours 38, 46, 105, 147
dispositif de lecture 78, 103, 107, 192, 219, 226, 237, 238, 244, 251, 282, 289, 298, 299, 300, 305, 306, 308, 309, 320, 323, 324, 334, 335, 357
Dondero, M. G. **88**, **147**, **154**, **200**, **366**
Drucker, J..... **101**, **131**

E

Eco, U. **43**, **68**, **107**
énonciation..... 45, 51, 147
espace de communication 68
Europeana..... 284, 310
Experiment in Art and Technology Datascape 350
expositions virtuelles 306

F

figures de la lecture 70
figures, figures sémiotiques, figures visuelles, sémiotique figurative. 52, 116, 125, 135, 140, 142, 161, 163, 166
Focillon, H. **161**
Fontanille, J. **49**, **143**, **148**
Foucault, M. **37**, **42**

Francastel, P. **163**

G

Gervereau, L. **29**, **121**
Getty Search Gateway 225, 227, 232, 239
Greimas, A. J. **43**, **45**, **142**
Groupe µ **67**, **144**

H

Hagelstein, M. **124**
Havelange, C. **122**
histoire de l'art 116, 160, 350, 374, 379
Hjemslev, L. **43**
Howells, R. **81**, **92**, **121**
hypertextes 97, 227, 252, 261, 262, 266, 267, 278, 296, 300, 310, 335

I

iconic turn 124
iconologie 117, 121, 123, 160
image processing 185
indexation d'images 155
interface 94

K

Klinkenberg, J.-M. **43**, **67**, **144**

L

langage, sciences du langage, théorie du langage 42
Lavédrine, B. **87**

M

Maingueneau, D. **38**
Manovich, L. **28**, **107**, **178**, **183**, **365**
Marin, L. **127**, **139**, **244**, **245**
matrice de pixels, image-matrice 174
médiasphère 29
Memobase 222, 227, 232, 233, 239, 243
métalangage 151, 158, 206
Moles, A. **28**, **33**

O

Object Photo 358
Odin, R. **68**
ontologies 155, 274

P

page-écran 57, 94
Panofsky, E. **35**, **117**, **123**, **162**
pattern recognition 193
Phototrails 365
Pignier, N. **57**, **70**, **307**
Plateforme ouverte du patrimoine 222, 227, 232, 239, 243

Q

quantification de la surface des images 154, 172, 178, 185, 205

R

Rastier, F. **40**, **44**, **48**
Replica 379
Reyes, E. **148**, **154**, **195**
rhétorique 67
Rijksstudio 321
Romescu, A.-E. **91**

S

Saemmer, A. **70**, **86**
Schapiro, M. **162**
sémio-pragmatique 68
sémiotique de la communication en ligne 57
service Web 94
Stockinger, P. **91**, **157**
stratégies de communication en ligne 70, 78, 92, 106, 134, 151, 220, 225, 299, 301
stratégies, stratégies d'énonciation 54

T

Tanaka-Ishii, K. **64**
textes, textes-énoncés 48, 52
théorie du support 58
Treleani, M. **90**

V

Vy 371

W

Warburg, A. **125**, **135**, **160**, **340**
WikimediaCommons 263

TABLE DES MATIÈRES

INTRODUCTION 13

PREMIÈRE PARTIE. SÉMIOTIQUE DE LA COMMUNICATION EN LIGNE : ÉNONCIATION ET STRATÉGIES MULTIMÉDIAS 25

CHAPITRE 1 POUR UNE APPROCHE SÉMIO-RHÉTORIQUE DES ARCHIVES ET DES COLLECTIONS D'IMAGES NUMÉRIQUES 27

1.1. Accroissement des archives et des collections d'images 28

1.2. Traitement documentaire des images : conserver, décrire, classer 32

1.3. Point de vue sémiotique sur l'archive 36
1.3.1. L'archive en tant que pratique énonciative 37
1.3.1.1. Théorie du langage (1) : langage et analyse des signes 41
1.3.1.2. Théorie du langage (2) : énonciation, discours et textes 45
1.3.2. Sémiotique du discours et praxis énonciative 49
1.3.3. Sémiotique de la communication en ligne : objets-supports, énonciation et stratégies de l'archivage numérique 53

1.4. Approche théorique du support numérique 57

1.5. Rhétorique et pragmatique de la communication en ligne 67

1.6. Conclusions 73

CHAPITRE 2 STRATÉGIES DE COMMUNICATION EN LIGNE D'IMAGES ET DÉMATÉRIALISATION DE LA CULTURE VISUELLE 77

2.1. Humanités numériques : une réaction transdisciplinaire 78

2.2. Archives d'images numériques et musée virtuel 81
2.2.1. Contrecoup de la dématérialisation et de la communication en ligne des archives d'images 86
2.2.2. Accessibilité à distance et médiation multimédia 90

2.3. Patrimoine numérique : tensions entre technique et interprétation 93
2.3.1. Conception des services et des interfaces Web 94
2.3.2. Bases de données et interaction humain-machine 98
2.3.3. Point de rencontre entre rhétorique et numérique : le dispositif de lecture 102

2.4. Conclusions 109

DEUXIÈME PARTIE. ANALYSES D'IMAGES ET MÉTHODES ASSISTÉES PAR UN ORDINATEUR 113

CHAPITRE 3 PROGRÈS DES MÉTHODES D'ANALYSE D'IMAGES EN HISTOIRE DE L'ART ET EN SCIENCES DU LANGAGE 115

3.1. De l'histoire de l'art aux sciences des images 116
3.1.1. Iconologie et interprétation des arts visuels 117
3.1.2. Culture visuelle, *iconic turn* et réflexivité des images 121
3.1.3. Nouvelles conditions de visibilité des images : regarder et penser autrement 127
3.1.4. *Digital art history* 130

3.2. Sciences du langage et interprétation d'images 137
3.2.1. Sémiologie et analyse d'images : message, support, représentation 137
3.2.2. Sémiotique visuelle et rhétorique des signes visuels 141
3.2.3. Analyse énonciative des images 147
3.2.4. *Computational semiotics* 152

3.3. Indexation d'images et limites du principe de métalangage de description 155

3.4. Revenir à l'analyse des formes visuelles 160

3.5. Conclusions 167

CHAPITRE 4 IMAGES NUMÉRIQUES, NOUVEAUX MÉDIAS ET ANALYSE INFORMATIQUE D'IMAGES 171

4.1. De l'analyse d'images à l'analyse d'images numériques : quelle différence ? 172
4.1.1. Quantification et affichage des images numériques 173
4.1.2. Le langage des nouveaux médias 178
4.1.3. *Image Processing*, détection des formes et classification d'images 185

4.2. *Computer vision*, *data visualization* et *deep learning* 193
4.2.1. Visualisation de données 195
4.2.2. La voie ouverte par le *deep learning* 201
4.2.3. *Datasets* et classification d'images : persistance du langage ? 205
4.2.4. Recherche d'images en fonction de la similitude visuelle 209

4.3. Conclusions 213

TROISIÈME PARTIE. ANALYSE DU CORPUS (1) : BASES D'IMAGES NUMÉRIQUES ET EXPOSITIONS VIRTUELLES 217

CHAPITRE 5 PREMIÈRE TENDANCE STRATÉGIQUE : LES BASES D'IMAGES NUMÉRIQUES DES INSTITUTIONS PATRIMONIALES 219

5.1. Des archives et des collections d'images aux bases d'images numériques 220
5.1.1. *Plateforme ouverte du patrimoine* 221
5.1.2. *Memobase* 222
5.1.3. *Belgian Art Links and Tools* 223
5.1.4. *Getty Search Gateway* 225

5.2. L'interface Web en tant qu'espace graphique et multimédia 226

5.3. Fonctionnalités des moteurs de recherche et présentation des résultats 238

5.4. Conclusions 252

CHAPITRE 6 RENOUVELLEMENT DES BASES D'IMAGES NUMÉRIQUES (1) : PARTICIPATION ET COOPÉRATION 255

6.1. Médiathèques en ligne sur le modèle de l'encyclopédie 256
6.1.1. *WikimediaCommons* : une stratégie participative 263
6.1.2. Neutralité de la page-écran, hyperlien et classification par catégories 266

6.1.3. *WikiArt* : une encyclopédie en ligne des arts visuels 277

6.2. Agglomérer et standardiser les bases d'images numériques 281
6.2.1. *Europeana project* : le choix stratégique de la coopération 284
6.2.2. Modélisation et standardisation des données 286
6.2.3. De l'exploration à la vulgarisation : écrits d'écran et rhétorique multimédia 289

6.3. Conclusions 302

CHAPITRE 7 RENOUVELLEMENT DES BASES D'IMAGES NUMÉRIQUES (2) : STRATÉGIES DE L'INTERACTIVITÉ 305

7.1. Des bases d'images numériques aux expositions virtuelles 306

7.2. Renouvellement d'un service Web à travers les applications interactives 320
7.2.1. La stratégie du *Rijksmuseum* 321
7.2.2. De la vulgarisation à l'observation : multiplication des points de vue dans l'interface 323

7.3. Autres exemples de communication en ligne dans une perspective interactive 334
7.3.1. *12 Sunset : Exploring Ed Ruscha's Archive* 334
7.3.2. *VISUALL-tek : Ikonik Analysis Toolkit* 339
7.3.3. *Rekall* 341

7.4. Conclusions 342

QUATRIÈME PARTIE. ANALYSE DU CORPUS (2) : CULTURE VISUELLE ET *BIG VISUAL DATA* 345

CHAPITRE 8 TROISIÈME TENDANCE STRATÉGIQUE : COMMUNICATION EN LIGNE ET VISUALISATION D'IMAGES 347

8.1. La *data visualization* au service des recherches en sciences humaines 349
8.1.1. Conception d'un service Web pour la recherche en histoire de l'art : *E.A.T. Datascape* 350
8.1.2. *Object : Photo* 358

8.2. De l'observation à l'expérimentation : *computer vision* et *media visualization* 364
8.2.1. Montrer les performances de l'analyse informatique d'images (photographies, peintures) 365
8.2.2. *Phototrails* 365
8.2.3. *Vy* 371

8.3. Conclusions 375

CHAPITRE 9 DE LA COMMUNICATION EN LIGNE D'IMAGES À L'ANALYSE INFORMATIQUE DES FORMES VISUELLES 377

9.1. *Media visualization* et analyse des formes visuelles 379
9.1.1. *Replica Project* 380
9.1.2. Première étape : définir un cadre de référence 382
9.1.3. Deuxième étape : de l'archive physique à l'archive numérique 384
9.1.4. Troisième étape : organiser l'information visuelle 388
9.1.5. Quatrième étape : le moteur de recherche *Replica* 391

9.2. Autres projets qui se basent sur la *media visualization* et sur les algorithmes 405

9.3. Conclusions 415

CONCLUSION 419

RÉFÉRENCES BIBLIOGRAPHIQUES ET WEBOGRAPHIE 443

TABLES DES ILLUSTRATIONS 469

INDEX 475

TABLE DES MATIÈRES 477