

## *m Partager des corpus et leurs analyses à l'heure du Web 2.0*

Aurélien Bénel et Christophe Lejeune

Ce numéro spécial étant consacré à l'Association pour l'autobiographie (APA), il est important de noter que dans la définition même de ses statuts, l'APA est engagée dans la création, l'archivage et la mise à disposition de corpus autobiographiques. Or, le partage, l'exploration, l'analyse et, au travers elles, l'indexation de corpus de textes est au cœur de deux champs de pratiques nous intéressant spécifiquement : l'analyse collaborative de corpus à des fins scientifiques (en sciences humaines et sociales) et l'enrichissement par le lecteur de documents partagés sur Internet (le Web 2.0).

Ces pratiques rappellent l'herméneutique philosophique, la théorie de l'interprétation inspirée de l'herméneutique biblique et choisie par Dilthey (XIXe s.) comme épistémologie et méthodologie des sciences de l'homme et de la société.

L'herméneutique peut nous en apprendre sur les documents, leur interprétation et la question de l'intersubjectivité. Pour chacune de ces dimensions, nous présenterons les outils que nous avons développés et illustrerons leur usage en archéologie et en sociologie.

### *Web 2.0 et documents*

Alors que les ordinateurs furent d'abord créés pour calculer des nombres et que l'informatique d'entreprise fut principalement vouée à la gestion et au partage de données, il est surprenant que les applications dites du « Web 2.0 » soient conçues pour le partage par les utilisateurs de leurs vidéos (YouTube, Dailymotion), photographies (Flickr), transparents (Slideshare) et textes (blogs, wikis).

Néanmoins, le partage de documents numériques au lieu de « données » ou de « faits » n'est pas réellement nouveau. Par exemple, en 1993, l'un des premiers sites web français fut celui de l'Association des Bibliophiles Universels (ABU) qui proposait à des scribes volontaires

### INTRODUCTION

### INVESTIGATION

À BASE DE DOCUMENTS

de déposer pour les partager des textes du domaine public (Cubaud & Girard, 1998). D'autres projets similaires furent même antérieurs au Web : ArXiv, un dépôt en libre accès de prépublications en physique, fut créé en 1991 grâce à une liste de diffusion et à un serveur FTP; le projet Gutenberg permit dès 1971 le partage de textes culturels du domaine public sur un "mainframe" connecté à un réseau qui deviendrait un jour l'Internet.

Avec l'adoption massive de technologies de partage de documents, quelques projets récents préfigurent ce que ces technologies pourraient apporter aux Sciences de l'homme et de la société. L'un des meilleurs exemples en est NietzscheSource (ex-HyperNietzsche), offrant à la fois ses fac-similés en ligne de sources primaires comme les manuscrits, les épreuves typographiques, les lettres, mais aussi sa littérature secondaire comme les transcriptions, les traductions, les parcours génétiques... (D'Iorio, 2007).

### *Herméneutique et documents*

Les documents constituent le socle des études en sciences de l'homme et de la société. À la fin du XIXe siècle, en réaction au positivisme, Dilthey établit que de telles études sont possibles malgré la singularité des sujets humains, parce que la vie de ces sujets est fixée dans des « ensembles structurés » (ex : des textes) compréhensibles par d'autres (Ricœur, 1986, p.92-93). Il adopte alors comme épistémologie une herméneutique philosophique, obtenue par généralisation de l'herméneutique biblique (théorie et méthode d'interprétation de la Bible).

Le nom même « d'Ancien (ou Nouveau) Testament » atteste du statut herméneutique archétypal de la Bible. En latin, « *testamentum* » désigne ce qui est témoigné, attesté. Il fait preuve, non comme une preuve mathématique, mais plutôt comme une pièce à conviction : c'est un témoignage à conserver, un document à produire. Les Évangiles en particulier sont écrits par des témoins, Jésus comme Socrate étant un orateur et non un écrivain (Ricœur, 1986, p.138-140). Ces textes constituent de plus une première couche d'interprétation, une première profession de foi de la communauté chrétienne naissante (Ricœur, 1969, p.378). À l'image du reste de la Bible, ces textes entremêlent récits et professions de foi (Ricœur, 1986, p.134-137). Ainsi, les Écritures qui font le récit de personnes qui interprètent des signes

doivent elles-mêmes être interprétées comme des signes.

À la différence d'une sémantique référentielle, l'herméneutique philosophique envisage également que le sens est médiatisé par les textes. Par exemple, la philologie et l'histoire qualifient les manuscrits et les inscriptions de « *testimonia* », un terme dont l'étymologie est similaire à celle de « *testamentum* ». Après des décennies de succès du modèle structuraliste et des grandes théories du social, les sciences sociales témoignent actuellement d'un regain d'intérêt pour ces questions et insistent sur le besoin d'un ancrage empirique des interprétations. En effet, les scientifiques savent que les phénomènes humains et sociaux ne sont pas directement accessibles. Pour accéder au statut de matériel empirique, être conservés et utilisés comme preuves, ces phénomènes sont « traduits » en une série de documents (Callon, 1986).

### *Illustration en archéologie*

Le lecteur pourrait s'étonner que nous traitions de l'importance des documents en archéologie, alors même que cette discipline se distingue de l'histoire par l'étude des vestiges matériels au lieu de celle des textes. Qui plus est, les musées et les dépôts sont remplis de ces vestiges. Pour autant, il ne faudrait pas négliger le statut très particulier des archives en archéologie. Par exemple, à l'École française d'Athènes, chaque plan, chaque photographie produite lors d'une mission subventionnée doit être déposée aux archives institutionnelles, même si son accès peut être restreint. Bien que la plupart de ces documents intermédiaires ne soient jamais publiés, il est nécessaire de disposer de témoins à travers l'espace et le temps. En effet, la distance et la difficulté d'accès des sites ou des vestiges entrent en tension avec la nécessité de les étudier (en bibliothèque ou en laboratoire) et de les montrer à des pairs. En outre, les vestiges s'altèrent avec le temps, voire sont détruits par les éléments naturels, les pillages, les promoteurs immobiliers ou... les fouilles archéologiques qui, par nature, détruisent le contexte des vestiges. Les archéologues prennent donc grand soin d'enregistrer ce contexte à travers ces documents. D'une certaine manière, la reproductibilité d'une fouille archéologie ne peut être garantie que par la fouille virtuelle des documents produits à son sujet (Iacovella *et al.*, 2005).

Plusieurs raisons préviennent l'assimilation des documents inter-

médiaires à des « données ». Tout d'abord, les mesures, les cartes et les photographies procèdent d'une perception sélective dépendant de la problématique, des hypothèses et des méthodes de leur auteur. C'est pour cette raison qu'elles sont documentées d'indices situationnels (souvent appelés « métadonnées ») de sorte que leur lecteur puisse les situer dans leur histoire. Ensuite, si les vestiges archéologiques, étudiés pour certains depuis des siècles, voient parfois leur numéro d'inventaire modifiés au cours du temps, les documents visuels restent pour leur part toujours identifiables. D'une certaine manière, ces documents constituent dès lors la seule référence pérenne. En troisième lieu, selon certains chercheurs, il n'existe pas en archéologie de différence épistémologique fondamentale entre ce qu'il est communément admis de qualifier de sources primaires (les textes anciens), de sources secondaires (les commentaires) et les données produites aujourd'hui. Partant du principe que « les objets n'ont pas d'attributs par eux-mêmes », mais à travers des « discours » d'auteurs anciens ou modernes, on peut considérer tous ces documents comme des « sources » (Desfarges & Helly, 1991).

Ces dernières années, nos prototypes ont été utilisés sur différents type de sources, comme des séries de monographies d'un centre de recherche, les chroniques annuelles des fouilles d'un pays donné, des rapports de fouille d'une nécropole ou des photographies de vases décrivant un sujet donné et provenant d'une ville donnée. La figure 1 illustre comment notre plateforme logicielle assiste la documentation d'une photographie émanant d'une archive scientifique.

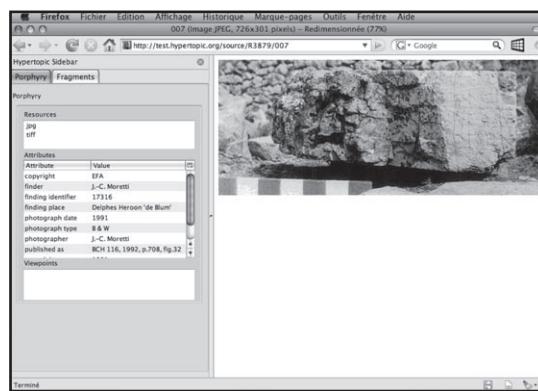


Figure 1: Documentation de la photographie d'une inscription, extraite d'une archive scientifique

### Illustration en sociologie

On pourrait nous objecter que la sociologie s'occupe de phénomènes sociaux (comme les représentations, les croyances ou l'expérience) et non de documents. Cependant, les chercheurs ne peuvent accéder directement à ces phénomènes : il leur faut passer par des entretiens menés par le chercheur ou par l'enregistrement (audio ou vidéo) d'interactions sociales dans ce qui constitue leur milieu naturel : les activités domestiques ou professionnelles de tous les jours. Les documents résultants sont donc de natures variées. Par exemple, dans une étude exploratoire sur la notion de société civile, Gautier Pirotte a rassemblé un corpus de déclarations publiques, d'articles de presse et de rapports annuels d'organisation non gouvernementales. Traditionnellement, les chercheurs en sciences humaines et sociales ne se contentent pas de rassembler des documents, ils participent à leur production : Sébastien Brunet, Pierre Delvenne et Catherine Fallon ont ainsi conduit une série de groupes focalisés (focus groups) auprès de comités de quartier à propos des risques électromagnétiques (notamment liés aux pylônes de mobilophonie) ; le corpus de recherche intègre les transcriptions de ces discussions (voir Figure 2) ainsi que des articles de loi.

The screenshot shows the Stealite software interface with the title "Stealite - Highlights - Iceweasel". The main window displays a transcript of a focus group discussion. The text reads:

vous pousse continuellement à courir de plus en plus de risques ? C'est un gros problème aussi bien philosophique que politique ou économique. Oui parce que c'est l'économie qui pousse la charrette, les gens aiment consommer, le GSM est un bel exemple. A partir du moment où on invente quelque chose de neuf, qui est pratique, qui apporte un confort, les gens foncent dedans. Les gens foncent, c'est normal, et alors la technologie elle invente toujours pour faire plus de pognon et c'est un cercle vicieux. Une spirale vicieuse et ils rendent les appareils indispensables.

Monsieur 1 : chez les adolescents il y a 38 pourcents d'accoutumance au GSM, il y a des études qui ont été faites. C'est énorme.

Monsieur 2 : un autre aspect est que les jeunes sont de plus en plus embarqués dans cette galère alors que ce sont eux qui devraient trouver les moyens pour trouver un autre système pour vivre. Or ils sont conditionnés. Nous leur avons montré l'exemple, et ils continuent, ils sont encore plus consommateurs que nous. Quand vous voyez le succès des GSM auprès des jeunes c'est effrayant, mais ce sont vraiment des besoins artificiels, qu'est-ce qu'on a besoin d'une télévision sur son GSM.

The interface includes a "Porphyry Sidebar" with a QR code and a "clap" icon. On the left, there are sections for "ITEM" (with a thumbnail of a person), "RESOURCES" (listing "source"), and "ATTRIBUTES" (a table showing frequent, repeated, and specific words). The "ATTRIBUTES" table has columns for "Attribut" and "Value".

Attribut	Value
répété	logement social
répété	nouveau besoin
répété	parlement européen
répété	rayonnement électromagnétique
répété	Région wallonne
spécifique	charrette
spécifique	dehors
spécifique	engrenage

Figure 2 : Segments fréquents, répétés et spécifiques d'un *focus group* sur les risques électromagnétiques

Pour chaque recherche, un corpus particulier est constitué afin de tester les hypothèses de la recherche. Plus que son format, le critère

de sélection d'un document est donc sa pertinence par rapport au sujet de la recherche. Un corpus est donc toujours patiemment défini, construit, élaboré. Les phénomènes sociaux n'y sont donc pas enregistrés comme des données, ils sont traduits en documents intermédiaires. Une discussion classique concerne précisément la fidélité de cette traduction : étant donné que la conduite d'un entretien biographique (ou l'observation de terrain) varie en fonction des hypothèses et des options théoriques du chercheur, seules des règles de méthodes précises (et néanmoins raisonnables) peuvent garantir que le document intermédiaire est un représentant fidèle du phénomène social investigué (Bourdieu, 1994).

## INTERPRETATION

### *Web 2.0 et interprétation*

L'histoire de l'informatique est intimement liée à celle de l'intelligence artificielle. Ceci est particulièrement visible dans la figure de son fondateur : Alan Turing. La « machine de Turing », d'une part, préfigure l'ordinateur en établissant la première démonstration de la faisabilité d'un automate programmable capable de calculer tout ce qui est calculable (Arsac, 2003). Le « test de Turing », d'autre part, constitue la première tentative de définition de ce que serait une intelligence machinique. Aussi, est-il assez surprenant de constater que le Web a évolué vers un Web 2.0 où « chaque usager peut contribuer au contenu, encourageant ainsi une intelligence collective » (O'Reilly, 2005), plutôt que vers un Web sémantique « doté d'un contenu compréhensible par les ordinateurs » (Berners-Lee *et al.*, 2001). En implémentant les indexations folksonomiques (Del.icio.us, Flickr, Connotea...) et les « trackbacks » (blogs), les développeurs de sites Web 2.0 prennent le parti de l'interprétation humaine plutôt que celui de la déduction automatique.

Au-delà de leur effet de nouveauté, on peut toutefois rattacher « folksonomies » et « trackbacks » à l'hypertexte tel que Theodor Nelson le conçut en 1965. En effet, en tant que diplômé en philosophie et en sociologie, il entendait transformer les « calculateurs » en « machines littéraires » destinées aux problèmes textuels. Elles devaient en particulier permettre la « transclusion » définie comme la combinaison virtuelle par le lecteur de documents existants en de nouveaux documents, en tissant les liens de citation correspondants (Nelson, 1999).

### *Herméneutique et interprétation*

La place du lecteur constitue un des nœuds de l'herméneutique. Ce nœud est désigné par l'expression « *liber et speculum* », traduisant un double rapport entre le livre et le “miroir” : parce que « c'est notre expérience qui explicite le mystère contenu dans le texte » mais aussi que « nous nous découvrons dans le miroir des mots » (Ricœur, 1969, p. 377). Cela est d'ailleurs vrai de tous les grands textes de l'humanité (de l'Odyssée au Coran) dont l'actualité est sans cesse renouvelée. Dès lors, Paul Ricœur se demande pourquoi l'herméneutique fut d'abord créée explicitement pour la Bible et en particulier pour la Bible chrétienne, et identifie les éléments qui suivent (Ricœur, 1969, p.373-380). Tout d'abord, la tradition canonique a intentionnellement conservé quatre évangiles et quatre récits de la création (Gen.1, Gen.2, Is.40-55, Jn.1), ce qui introduit des contradictions d'un point de vue strictement logique. En second lieu, certains passages invitent à lire « allégoriquement » d'autres (par exemple, la lettre de Paul aux Galates dit qu'il faut comprendre la descendance des épouses d'Abraham comme deux Alliances). Enfin, le « Nouveau Testament » se réfère et interprète constamment « l'Ancien Testament ». La conscience du problème herméneutique de la Bible peut donc se comprendre en vertu de la pluralité interne au texte. Paul Ricœur note d'ailleurs que les tentatives d'évitement ou de réduction de cette pluralité furent déclarées hérétiques (comme par exemple le marcionisme, au IIe siècle, qui rejettait l'Ancien Testament et ne conservait qu'un seul évangile).

En herméneutique philosophique, Dilthey reprend l'idée du « *liber et speculum* » dans la notion de « compréhension ». Pour lui, cette interprétation subjective joue le même rôle en sciences de l'homme que « l'explication » en sciences de la nature. Toutefois, depuis le XIXe siècle, la ligne de démarcation s'est atténuée. Du côté des sciences de la nature, la subjectivité du chercheur est aujourd'hui assumée tout au moins dans les phases de la pratique scientifique, comme l'invention des hypothèses, où elle n'a pas d'impact méthodologique (Popper, 1935). Du côté des sciences de l'homme et de la société, on ne peut plus limiter le travail du texte à l'étude de la « puissance de l'œuvre à se projeter hors d'elle-même ». Si l'on veut rendre compte du structuralisme, il faut y adjoindre l'explication comme étude de la « dynamique interne du texte » (Ricœur 1986, p.36).

Ce n'est pas autre chose qu'affirme Max Weber (1971), considéré

comme le fondateur de la sociologie compréhensive, qui ajoute par ailleurs que les problèmes soulevés à propos de la subjectivité sont souvent mal compris. Ces difficultés ne sont pas étrangères aux différentes acceptations de la subjectivité.

A. Depuis des décennies, les théories sociologiques ont largement diffusé les mises en garde contre la subjectivité : le chercheur doit être conscient de sa position dans le monde social ainsi que de ses options (explicites ou implicites) de son cadre méthodologique ou épistémologique, car elles déterminent sa subjectivité. S'il n'en tient pas compte, cette subjectivité peut le conduire à des résultats inattendus (par exemple, en filtrant les matériaux empiriques). C'est pour cette raison que Pierre Bourdieu, citant Gaston Bachelard, met les chercheurs en garde contre leurs propres hypothèses implicites de sens commun (Bourdieu *et al.*, 1968). Dans ce sens, la subjectivité est un obstacle.

B. La subjectivité concernée par la mise en garde précédente est bien celle du chercheur : la subjectivité des acteurs n'est, elle, pas concernée (Lejeune, 2001). En effet, le sens commun des acteurs (leur subjectivité) constitue une riche ressource pour la recherche. La subjectivité des acteurs fournit son objet à la science sociale compréhensive; sans elle, aucune investigation ne serait possible.

C. Par ailleurs, la compréhension du point de vue des acteurs et de leur vécu exige de l'empathie de la part du chercheur. En d'autres termes, le chercheur doit être capable de partager l'arrière-plan culturel des acteurs (Schütz, 1987). La commune humanité rend cela possible : le chercheur mobilise sa subjectivité d'acteur ordinaire comme un outil lui permettant de comprendre pleinement le monde vécu des acteurs. La subjectivité du chercheur est donc également une ressource (en tant qu'outil) pour accéder à la subjectivité des acteurs (entendue comme le réceptacle du phénomène social investigué).

D. Enfin, la subjectivité du chercheur se définit également par une série d'hypothèses, de questions et d'options théoriques déterminées. Dans cette quatrième (et dernière) acceptation, elle garantit la productivité de l'interprétation scientifique. Max Weber illustre cette productivité en mentionnant une controverse autour de l'engagement d'un professeur de droit; il soutient qu'être anarchiste n'est non seulement pas un obstacle à l'exercice d'une telle position mais que, en outre, la perspective qu'amène à adopter de telles convictions est susceptible de renouveler les recherches en droit (Weber, 1992 : 376). Dans cette

optique – orthogonale à celle définissant la subjectivité comme un obstacle –, la subjectivité enrichit l'interprétation et, comme nous allons le montrer, constitue un prérequis du débat scientifique.

La compréhension nécessite donc un arrière plan culturel et la conscience son implication historique (la « Wirkungsgeschichtliches Bewusstsein » de Gadamer). Pour comprendre, il faut prendre en considération le contexte historique du phénomène, le contexte de l'analyse et la différence entre ces deux subjectivités. Étant donnée cette altérité, la subjectivité n'est pas source d'erreur; au contraire, elle constitue un garant de scientificité.

### *Illustration en archéologie*

Le statut interprétatif de l'archéologie a été longuement débattu dans les années 1960 et 1970 entre les partisans d'une « nouvelle archéologie » et les défenseurs d'une « archéologie humaniste ». La thèse d'Anne-Marie Guimier-Sorbets sur « l'analyse et la formalisation de l'ornement géométrique [dans les mosaïques gréco-romaines] en vue du traitement automatique » illustre particulièrement bien cette controverse. À la suite de celle-ci, un des membres du jury, Philippe Bruneau, écrit un article qui donne naissance à une polémique inédite. Il y discute le statut des « descripteurs » dans les banques de données. Tout d'abord, il note que ce sont souvent des descripteurs comme « couleur de fond » ou « couleur de motif » qui sont retenus pour leur universalité et leur indépendance à la chronologie et à la géographie, ce qui est plutôt étonnant dans une science historique. En second lieu, il s'interroge sur la validité de ces deux descripteurs dans le cas de méandres, où chaque frise noire sur blanc a son complémentaire blanc sur noir. La conceptrice de la base de données répond : « nous convenons d'analyser la ligne de postes situées le plus à l'extérieur de la mosaïque [...] ; l'autre partie, complémentaire, est analysée comme fond ». Philippe Bruneau regrette que l'ensemble de la description doive, par convention, aller des bords vers le centre, compte tenu que les mosaïques étaient construites depuis leur centre vers l'extérieur. Enfin, la forme agentive du terme « descripteur » laisserait penser que c'est le dispositif, et non l'archéologue, qui décrit l'artefact. Et oublier le rôle de description de l'archéologue conduit à oublier que le premier à décrire et à classer l'artefact était l'usager ancien lui-même (Bruneau, 1976).

Des travaux récents d'archéologues comme Roland Etienne adoptent une perspective historiographique. En historiographie (l'*histoire de l'histoire*), il n'existe pas un seul sens historique, mais une succession de sens actualisés. L'*histoire* y est de plus envisagée comme le résultat de la confrontation du passé et du présent, entre ce qui organisait alors la vie ou la pensée et ce qui permet aujourd'hui de la penser (de Certeau, 1975).

Nos prototypes informatiques permettent au chercheur de créer librement et de faire évoluer son cadre analytique propre en fonction de ses hypothèses. La figure 3 illustre ainsi l'*analyse qualitative* qu'ont élaborée des étudiants de maîtrise en histoire de l'art à partir des scènes figurées sur des photographies de vases.



Figure 3: Analyse qualitative des scènes illustrées sur des vases

#### *Illustration en sociologie*

Comme nous l'avons discuté précédemment, l'*interprétation* soulève la question de la subjectivité. L'*exemple* déjà mentionné sur la mobilisation autour des risques électromagnétiques nous permet d'illustrer ces questions. Dans une telle enquête, un postulat d'*irrationalité* des

acteurs en matière de perception des risques précède (et est indépendant) de tout examen de documents intermédiaires dans lesquels les acteurs expriment leur point de vue. Une telle posture est susceptible de filtrer les éléments empiriques recueillis et constitue par conséquent un obstacle à une interprétation valide, entravée par la subjectivité du chercheur dans sa première acception. Les documents intermédiaires collectés attestent par ailleurs de la subjectivité des acteurs, regroupant notamment des croyances, des conceptions ou des argumentations à propos des pylônes électriques. En tant qu'objets d'études des sciences sociales, ces phénomènes illustrent notre deuxième acception de la subjectivité.

Le chercheur peut considérer que ces raisonnements sont non seulement des objets d'études, mais également qu'ils recèlent un intérêt. Grâce à un arrière plan culturel partagé (éventuellement acquis sur le terrain), il pourra par exemple suivre le développement des théories de sens commun sur les effets contre-intuitifs et les phénomènes auto-entretenus (Lejeune, 2008). La Figure 4 illustre comment nos prototypes informatiques peuvent l'assister dans cette tâche. Notons que l'option ici adoptée par le chercheur – et consistante avec le programme symétrique de l'anthropologie des sciences et des techniques – emprunte une direction opposée au postulat d'irrationalité mentionné ci-dessus.

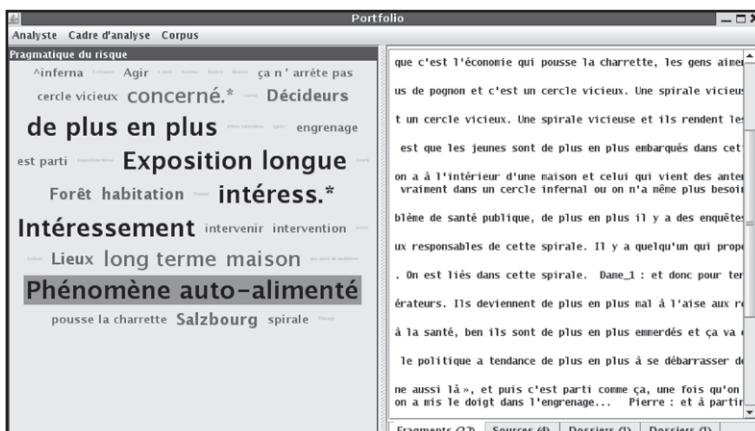


Figure 4: Analyse des phénomènes auto-entretenus dans la perception des risques

En outre, la discipline sociologique reconnaît plusieurs paradigmes et, par voie de conséquence autant d'interprétations. On reconnaît ici l'exercice de la subjectivité du chercheur dans sa quatrième acception. Enfin, les questions de subjectivité sont abordées différemment selon le paradigme. Quand, par exemple, Bourdieu approche les productions artistiques, il se concentre sur les positions et les stratégies de préservation. Becker, pour sa part, adopte une perspective différente : dans sa brillante analyse des musiciens de Jazz, il montre comment les interactions sont localement conduites et régulées (Becker, 1985).

À chaque fois, les chercheurs prennent soin de ne pas projeter leur subjectivité (dans son acception première) sur les documents intermédiaires qu'ils collectent, ils mobilisent plutôt leur subjectivité comme outils (dans sa troisième acception) afin d'accéder à la subjectivité des acteurs (la deuxième acception). Cependant, leurs interprétations diffèrent, en fonction du paradigme dont elles proviennent. Quel que soit le terrain, un grand nombre d'analyses peuvent être réalisées (et le sont effectivement), contribuant partant à autant d'interprétations différentes, voire conflictuelles. D'ailleurs, même au sein d'un paradigme donné, les chercheurs n'interprètent pas nécessairement les documents intermédiaires de la même façon. Cette subjectivité (dans sa quatrième acception) fait la richesse et la productivité du travail interprétatif de l'activité scientifique.

## INTERSUBJECTIVITE

### *Web 2.0 et intersubjectivité*

Depuis le temps des systèmes à temps partagé, les technologies de l'information permettent à des milliers d'utilisateurs de travailler ensemble sur un support numérique. L'innovation du Web 2.0 réside quant à elle dans l'usage massif de systèmes multi-utilisateurs par des volontaires sans requérir une organisation a priori du collectif. Là où l'informatique traditionnelle se concentre sur des structures de données et des procédures de travail consensuelles (schémas de bases de données, ontologies, workflows, business processes), les technologies du Web 2.0 s'emparent du dissensus entre les utilisateurs en montrant les fils de commentaires des usagers, les « nuages » de tags attribués à des documents par différents lecteurs ou l'onglet de discussion sur une page Wiki.

Sans inventer la prise en compte de l'intersubjectivité en informatique, le Web 2.0 rend cependant accessibles à tous des fonctionnalités alors cantonnées à quelques prototypes pour les bibliothèques numériques comme le partage de commentaires et de signets à des fins de filtrage collaboratif (Röscheisen et al., 1995) ou l'expression des points de vue des lecteurs afin d'améliorer l'indexation (Béné et al., 1999, Nanard & Nanard, 2001).

### *Herméneutique et intersubjectivité*

La théorie des sens multiples de l'Écriture (« multiplex intellectus ») constitue également un point crucial de l'herméneutique biblique (Ricœur, p.376-377). Origène (IIIe s. apr. J.-C.) et les exégètes chrétiens qui le suivent distinguent :

- un sens « littéral » à apprendre (parce que les mots ont été écrits par d'autres mains dans d'autres temps),
- un sens « allégorique » à croire (parce qu'ils peuvent être envisagés comme signes d'un autre temps pour aujourd'hui),
- un sens « moral » auquel se conformer (« moral » signifiant « relatif aux manières »),
- un sens « anagogique » ou « mystique » vers lequel tendre (« anagogique » signifiant « élevé » ; « mystique » signifiant « gardé silencieux »).

Ces quatre sens se retrouvent également dans la tradition juive en tant que « Pshat », « Remez », « Drash » et « Sod ». Dans un conte mystique de Moshe Cordovero (XVIIe s.), chacun de ces sens (abréviés « P », « R », « D », « S ») constitue un niveau du même « PaRDeS » (jardin). Il est important de noter que dans ce conte chaque sens est exploré par un rabbin différent.

En herméneutique philosophique, l'altérité entre le monde social et l'analyste se complète d'une seconde altérité entre son point de vue et ceux de ses collègues. Après la subjectivité, l'intersubjectivité devient l'enjeu de l'analyse. En effet, comme mis en évidence par les sociologues des sciences, la confrontation de points de vue conflictuels est au cœur de la dynamique du champ scientifique (Bourdieu, 1976; Merton, 1979). Cependant, l'exercice d'une telle intersubjectivité à travers les frontières disciplinaires, les paradigmes et les méthodes n'en reste pas moins un défi de taille.

### *Illustration en archéologie*

Dès l'établissement des premières « banques de données » archéologiques dans les années 1970, leurs promoteurs s'interrogent sur l'impact possible des différences entre les points de vue des experts. Jean-Claude Gardin, pourtant porteur d'un « programme logiciste » pour les sciences de l'homme et de la société, est obligé de reconnaître qu'il est impossible d'envisager les bases de données archéologiques sans se questionner sur l'accord de l'ensemble de la communauté archéologique sur de telles données (Gardin, 1984). Sans pour autant proposer de réelles solutions, René Ginouvès et Anne-Marie Guimier-Sorbets conseillent aux concepteurs de base de données de viser l'intersubjectivité plutôt que d'espérer un consensus hypothétique assurant l'objectivité des données (Ginouvès & Guimier-Sorbets, 1978).

Une des fonctionnalités les plus innovantes de nos prototypes informatiques réside dans la comparaison de différents cadres d'analyse en tant que points de vue différents sur un même corpus de documents. Sur la figure 5, on peut ainsi voir deux analyses, construites par différentes étudiantes, de photographies de vases géométriques.

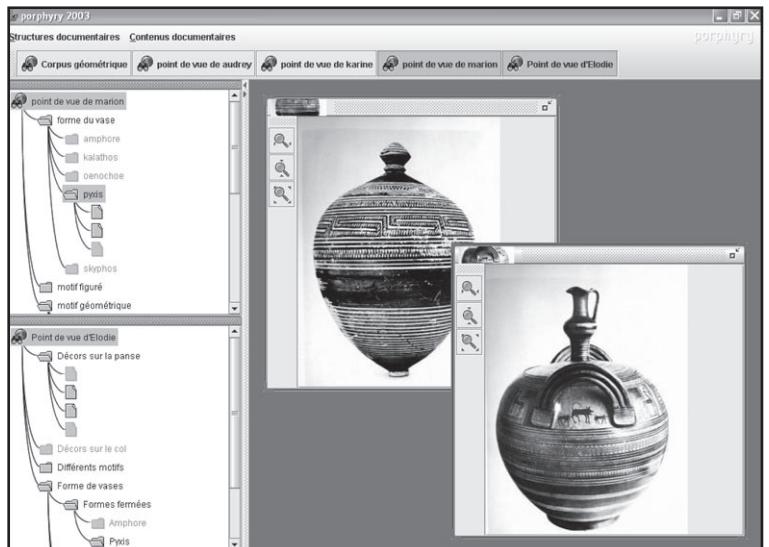


Figure 5: Confrontation de deux analyses des mêmes vases géométriques

Des projets récents en archéologie tiennent compte du fait que les vestiges complexes sont souvent étudiés par différentes disciplines dont les bibliographies se recouvrent peu. Bruno Helly indique par exemple que les tombes et les monuments funéraires sont étudiés indépendamment en fonction de l'origine de la pierre, de leur sculpture, peintures, architecture et inscriptions. Selon lui, un support numérique contribuerait à confronter les points de vue disciplinaires et ainsi à comprendre de tels objets complexes dans leur totalité (Iacovella *et al.*, 2005).

### *Illustration en sociologie*

Il arrive très souvent qu'un même phénomène soit analysé par des sociologues différents. Il est par contre très rare qu'ils partagent leurs matériaux empiriques (Rigot, 2006). De plus, compte tenu de la division de la sociologie en paradigmes, les débats scientifiques se déroulent au sein de chaque paradigme. Et le fait que des sociologues appartenant à différents paradigmes étudient des phénomènes similaires ne débouche pas nécessairement sur un débat, vu que « on adhère à un paradigme, on n'en démontre pas la validité » (Rigaux, 1992, p. 21).

Les débats ne sont pas absents pour autant. Dans les années quatre-vingts, l'anthropologie des sciences s'est attelée à la respécification du paradigme du constructivisme social. Au contraire de la conception réaliste de la science (selon laquelle les faits scientifiques reflètent le monde tel qu'il est), les socio-constructivistes étudient les interactions sociales (entre chercheurs) qui contribuent à constituer une science (Latour & Woolgar, 1988). La perspective de l'accès à la réalité se révèle incapable d'expliquer l'existence des controverses (constitutives de la dynamique de la science). De son côté, le socio-constructivisme ne permet pas de comprendre pourquoi certaines théories l'emportent et d'autres non. Le débat qui eut lieu alors participa à définir une nouvelle version du constructivisme social. Selon celle-ci, la solidité des assertions scientifiques est fonction du nombre d'éléments convergents (Latour, 1989); les faits sont donc solides grâce (et non en dépit) de leur construction (Latour, 1996). La controverse (donc le conflit entre les partisans de différents paradigmes) a donc contribué à l'élaboration d'une perspective plus riche (Simmel, 1995).

Des controverses similaires se déroulent sur nos plateformes. Comme pour l'anthropologie des sciences, ces débats contribuent à

rendre les interprétations plus solides. Dans l'exemple de la recherche sur la perception des risques électromagnétiques, les membres d'une équipe en science politique ont développés deux cadres d'analyses concurrents, sur base des mêmes documents – les groupes focalisés (focus groups) transcrits (Brunet & al., 2008). Pour sa part, le sociologue Gautier Pirotte confronte, grâce à nos outils, deux théories concurrentes – les économies de la grandeur (Boltanski & Thevenot, 1991) et sa propre théorie de la société civile (Pirotte, 2007) – pour étudier les rapports annuels d'organisations non gouvernementales (voir la figure 6).

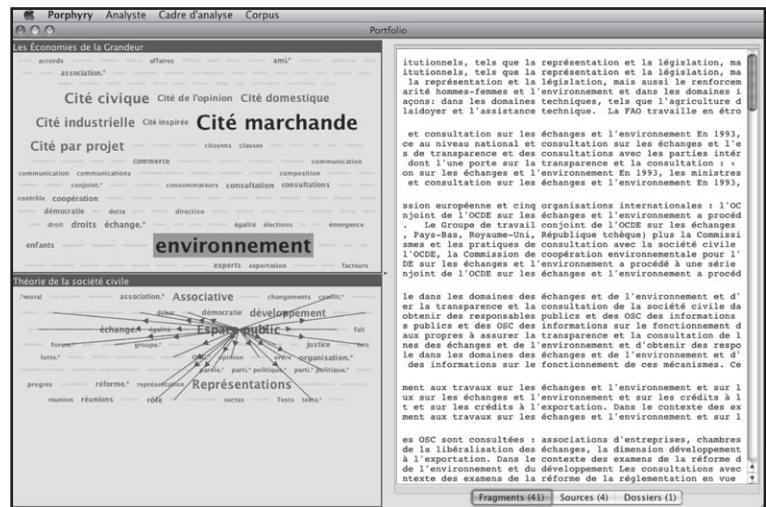


Figure 6 : Rapports annuels d'ONG à travers le prisme des économies de la grandeur et de la théorie de la société civile

## CONCLUSION

Dans cet article nous avons exploré en quoi les technologies du Web 2.0 pouvaient être considérées comme des résurgences de l'herméneutique, alors que l'informatique traditionnelle le serait du positivisme logique.

Nous avons donné des illustrations en archéologie et en sociologie des échos des grandes questions de l'herméneutique philosophique dans les débats méthodologiques de ces disciplines. Nous avons vu comment il était possible de concevoir et de prototyper des logiciels en tenant compte de notions comme les enquêtes à base de documents,

l'interprétation et l'intersubjectivité, non dans le vain espoir de les automatiser, mais dans celui de fournir un support propice aux analyses des chercheurs.

Dès lors, ne pourrait-on pas voir dans le succès et la maturité du Web 2.0 une chance unique d'encourager le développement des « Digital Humanities » et des Sciences de l'homme et de la société elles-mêmes ?

Arsac, Jacques, « L'informatique pose la question du sens », In : Bernard d'Espagnat (éd.), *Implications philosophiques de la science contemporaine*, Tome 3, Paris : PUF, 2003.

Becker, Howard, *Outsiders. Études de sociologie de la déviance*, Paris, Métailié, 1985.

Bénel, Aurélien, Calabretto, Sylvie, Pinon, Jean-Marie, « Indexation "sémantique" de documents archéologiques » in *Deuxième colloque du chapitre français de l'ISKO*, Lyon, 21-22 Octobre 1999. Disponible sur :

[http://benel.tech-cico.fr/publi/benel\\_ISKO\\_99.pdf](http://benel.tech-cico.fr/publi/benel_ISKO_99.pdf)

Berners-Lee, Tim, Hendler, James, Lassila, Ora, *The Semantic Web*, in *Scientific American*, May 2001. Disponible sur :

<http://www.sciam.com/article.cfm?id=the-semantic-web>

Boltanski, Luc, Thevenot, Laurent, *De la justification. Les économies de la grandeur*, Paris, Gallimard, 1991.

Bourdieu, Pierre, « Le champ scientifique », *Actes de la recherche en sciences sociales*, vol. 2, n°2, 1976, pp. 88-104.

Bourdieu, Pierre, *Raisons pratiques. Sur la théorie de l'action*, Paris, Seuil, 1994.

Bourdieu, Pierre, Chamboredon, Jean-Claude, Passeron, Jean-Claude, *Le métier de sociologue*, Paris, Mouton/Bordas, 1968.

Bruneau, Philippe, « Quatre propos sur l'archéologie nouvelle », *Bulletin de Correspondance Hellénique*, n°100, tome 1, 1976, p.103-135. Disponible sur : [http://cefael.efa.gr/detail.php?serie\\_id=BCH&volume\\_number=100&issue\\_number=1&page\\_number=103&page\\_type=1](http://cefael.efa.gr/detail.php?serie_id=BCH&volume_number=100&issue_number=1&page_number=103&page_type=1)

Brunet, Sébastien, Delvenne, Pierre, Fallon, Catherine et Lejeune, Christophe, « La perception du risque en situation de haute incertitude scientifique : le cas des champs électromagnétiques en Belgique », *Politique et sociétés*, 2008, soumis pour publication.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- Callon, Michel, « Éléments pour une sociologie de la traduction. La domestication des coquilles Saint-Jacques et des marins-pêcheurs dans la baie de Saint-Brieuc », *L'Année sociologique*, vol. 36, 1986, pp. 169-208.
- Cubaud, Pierre, Girard, Didier, « ABU : une bibliothèque numérique et son public », *Documents numériques*, vol. 2, Hermès, 1998. Disponible sur : <http://jasmin.cnam.fr:8081/RR9813/>
- Desfarges, Patrick, Helly, Bruno, « L'archéologie, système d'information scientifique », *Aplicaciones Informáticas en Arqueología : Teorías y sistemas*. Saint-Germain-en-Laye, 1991.
- Dosse, François, *L'empire du sens*, Paris: La Découverte, 1995.
- Gardin, Jean-Claude, « Les bases de données dans les sciences de l'antiquité : l'ajustement nécessaire des fins aux moyens », *Banques de données et sciences de l'antiquité*, 1984.
- Ginouvès, René, Guimier-Sorbets, Anne-Marie, *La constitution des données en archéologie classique*. Ed. du CNRS, 1978.
- Iacovella, Andrea, Bénel, Aurélien, Calabretto, Sylvie, Helly, Bruno, « Assistance à l'interprétation dans les bibliothèques numériques pour les sciences historiques », 2005 in *Actes du colloque de bilan du programme interdisciplinaire « Société de l'information » 2001-2005*, p. 167-179. Disponible sur : [http://publi.tech-cico.fr/paper\\_details.php?paperid=41](http://publi.tech-cico.fr/paper_details.php?paperid=41)
- D'Iorio, Paolo, « Nietzsche on New Paths: The HyperNietzsche Project and Open Scholarship on the Web », in Maria Cristina Fornari, Sergio Franzese (éds.), *Friedrich Nietzsche. Edizioni e interpretazioni*, Pisa ETS, 2007. Disponible sur : <http://www.hypernietzsche.org/doc/files/new-paths.pdf>
- Lacour, Philippe, « Portrait de l'intellectuel en DJ : Wikipédia face à l'expertise scientifique », *La vie des idées*, 2008. Disponible sur : <http://www.laviedesidees.fr/Portrait-de-l-intellectuel-en-DJ.html>
- Latour, Bruno, *La Science en action*, Paris, La Découverte, 1989.
- Latour, Bruno, *Petite réflexion sur le culte moderne des dieux Faitiches*, Paris: Les Empêcheurs de penser en rond, 1996.
- Latour, Bruno, Woolgar, Steve, *La vie de laboratoire. La production des faits scientifiques*, Paris, La Découverte, 1988.
- Lejeune, Christophe, « Du mode de définition de deux programmes de recherche en sociologie et en ethnométhodologie », *Carnets de bord*, Vol 2, 2001, p. 56-66.

Lejeune, Christophe, « Au fil de l'interprétation. L'apport des registres aux logiciels d'analyse qualitative », *Revue suisse de sociologie*, 2008, vol 34, Issue 3, pp.55-70.

King Merton, Robert, *The Sociology of Science*, Southern Illinois University Press, 1979.

Nanard, Marc, Nanard, Jocelyne, « Cumulating and sharing end users knowledge to improve video indexing in a video digital library », *Proceedings of the first ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'01)*, ACM Press, 2001 p.282-289. Disponible sur : <http://doi.acm.org/10.1145/379437.379683>

Holm Nelson, Theodor, « Xanalogical Structure Needed Now More Than Ever », *ACM Computing Surveys*, Volume 31, Issue 4, ACM Press, 1999. Disponible sur : [http://www.cs.brown.edu/memex/ACM\\_HypertextTestbed/papers/60.html](http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/60.html)

O'Reilly, Tim, « What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software », September 30, 2005. Disponible sur : <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

Popper, Karl Raimund, *La logique de la découverte scientifique*, Paris, Payot, 1973, 480 p. (Note : Édition originale en allemand publiée en 1935, traduite en anglais en 1959 puis complétée en 1960, 1962, 1965 et 1968).

Pirotte, Gautier, *La notion de société civile*, Paris, La découverte, 2007.

Rigaux, Natalie, *Raison et déraison. Discours médical et démence sénile*. Bruxelles, De Boeck, 1992.

Rigot, Huguette, « (En)-jeux de corpus pour la recherche en SHS. Enoncé, textes et documents », in François Ratier & Michel Ballabriga (dir.), *Corpus en Lettres et Sciences sociales : des documents numériques à l'interprétation, Actes du colloque international d'Albi*, Paris, Texto, 2006. Disponible sur : <http://www.revue-texto.net/1996-2007/Parutions/Livres-E/Albi-2006/Rigot.pdf>

Ricœur, Paul, Le conflit des interprétations : Essais d'herméneutique, Paris, Seuil, 1969.

Ricœur, Paul, Du texte à l'action : Essais d'herméneutique II, Paris, Seuil, 1986.

Röscheisen, Martin, Mogensen, Christian, Winograd, Terry, « Beyond

browsing: Shared comments, soaps, trails, and on-line communities », *The Third International World Wide Web Conference, "Technology, Tools and Applications"*, 1995. Disponible sur : [http://www.igd.fhg.de/archive/1995\\_www95/proceedings/papers/88/TR/WWW95.html](http://www.igd.fhg.de/archive/1995_www95/proceedings/papers/88/TR/WWW95.html)

Schütz, Alfred, *Le chercheur et le quotidien. Phénoménologie des sciences sociales*, Paris: Klincksieck, 1987.

Simmel, Georg, *Le conflit*, Dijon, Circé, 1995.

Weber, Max, Économie et société. Les catégories de la sociologie, Paris, Plon, 1971.

Weber, Max, Essais sur la théorie de la science, Presses Pocket, 1992, [Plon, 1965].

Zacklad, Manuel, Bénel, Aurélien, Cahier, Jean-Pierre, L'Hédi Zaher, Lejeune, Christophe et Zhou, Chao, « Hypertopic : une métasémio-tique et un protocole pour le Web socio-sémantique » in Francky Trichet (dir.), *Actes des 18eme journées francophones d'ingénierie des connaissances*, pp. 217-228, Cépaduès (France), 2007. Disponible sur : [http://publi.tech-cico.fr/paper\\_details.php?paperid=79](http://publi.tech-cico.fr/paper_details.php?paperid=79)