De l'objet connecté aux big data : quelle plus-value pour le renseignement en terme de performance ?

Sandrine Mathen et Patrick Leroy

"Le renseignement, c'est l'érudition de l'Etat."

Alain Dewerpe -
Espion, anthropologie historique du secret d'Etat contemporain
Editions nrf - 1994

« Le défi de l'homme est de reprendre le pouvoir sur les données »

Satya Nadella
Directeur général de Microsoft – octobre 2016

1. Introduction

Comme nous l'avions exprimé récemment[1], les services de renseignement sont confrontés à l'avènement inéluctable de nouveaux champs d'action, fruits de leur adaptation à l'environnement dans lequel ils opèrent, tant par l'évolution de la nature des menaces et de risques, que par l'apparition de nouvelles sources d'informations comme les objets connectés et les données massives (*big data*), entre autres, mais aussi par l'accélération du temps décisionnel (dictature de l'immédiateté) promue bien souvent par l'apparition de nouveaux clients du renseignement mais surtout par la rapidité de la circulation de l'information. Rappelons, que chaque organisation, chaque administration publique a besoin d'être informée de ce qui se passe dans le segment de l'environnement social auquel elle s'adresse[2]. De même façon, les services de renseignement doivent pouvoir maîtriser, comme organisations réductrices d'incertitude, l'environnement opérationnel global et incertain dans lequel ils opèrent au profit de l'Etat et des citoyens, et par conséquent innover, idéalement *a priori,*

[1] Patrick Leroy, « La communauté du renseignement belge : essai de définition ». In *Revue militaire belge*, n° 12, Institut Royal Supérieur de Défense, juin 2016, p. 85.
[2] Erhard Friedberg, *« L'analyse sociologique des organisations ».* In *Pour,* N°28, 1972, (fruit des travaux de Michel Crozier, et du Centre de Sociologie des Organisations), p. 55.

dans leurs méthodes, dans leurs structures et même de façon conceptuelle (certains auteurs parlent de changement de paradigme[3]).

De même que l'adaptation du renseignement à la situation internationale est présentée comme une règle absolue par Olivier Forcade et Sébastien Laurent[4], l'adaptation du renseignement aux nouvelles sources d'informations nous paraît tout aussi fondamentale car elle impacte tant les méthodes de collecte que les méthodes d'évaluation de la pertinence de l'information, mais aussi la relation entre les 2 segments[5] professionnels du renseignement (collecte et analyse) et par conséquent les relations entre l'analyse et les décideurs politiques, économiques et militaires, les consommateurs du renseignement[6]. Les objets connectés en effet se multiplient à en donner le vertige. Chacun d'entre nous sera tôt ou tard pris dans *l'internet des objets*, si ce n'est déjà le cas. Ces objets, par nature, génèrent des données, voire des métadonnées, dont la quantité augmente proportionnellement jusqu'à obtenir des données massives, des *big data*, dont le traitement est bien la clef du problème. Une exploitation qui devrait s'avèrer utile pour les services de renseignement, aux niveaux micro et macro, mais dont les cadres éthiques gagnent à être (re)pensés. Car le potentiel détournement de la finalité des données et le non-respect de la vie privée s'invitent à la table de l'efficacité opérationnelle, voire décisionnelle. *Small data, smart data* ou *big data*, une distinction est à faire. Les premières plus opérationnelles ne permettent pas l'analyse prédictive des secondes ni la détermination de tendances des troisièmes. Le renseignement pourrait tirer parti des différents niveaux. Nous trouvons une manifestation de cet intérêt par la communauté du renseignement américaine lors de l'audition du directeur national du renseignement James Clapper par le Sénat, en février 2016[7] qui a eu cette franchise : *Les services de renseignement pourraient tirer parti de l'Internet des objets pour identifier, surveiller ou localiser des suspects, découvrir des indicateurs potentiels, ou obtenir des mots de passe.* Le profilage (politique) et l'analyse prédictive sont d'autres facettes exploitables de ces *data*. Un rapport publié le 1er février 2016 par le centre de recherche Berkman de l'université Harvard estime même que la quantité de données rassemblées par ces objets en font l'une des pistes

[3] William Lahneman, *"The need for a new intelligence paradigm." In International Journal of intelligence and Counterintelligence*, Vol 23, N°2, Summer 2010, pp 201-225.

[4] Olivier Forcade & Sébastien Laurent, *Secrets d'Etat – pouvoirs et renseignement dans le monde contemporain,* Editions Armand Colin, 2005, p. 193.

[5] Anselm Strauss, « La dynamique des professions », in *La trame de la négociation - Sociologie et interactionnisme,* Editions Lharmattan, 1992, p. 68.

[6] Guy Rapaille, Dirk Peeters & Patrick Leroy, « Ethique et analyse : la relation avec le décideur », in, *Renseignement et éthique : le moindre mal nécessaire,* Groupe Européen de Recherche en Ethique et Renseignement, 2014, pp 267-283.

[7] James Clapper, *Worldwide Threat Assessment of the US Intelligence Community*, *Statement for the Record*, februari 9, 2016. In the future, intelligence services might use the IoT for identification, surveillance, monitoring, location tracking, and targeting for recruitment, or to gain access to networks or user credentials. En ligne sur le site du Comité du Sénat US sur les Services armés https://www.armed-services.senate.gov/imo/media/doc/Clapper_02-09-16.pdf

privilégiées pour que les agences de renseignement puissent contourner les protections mises en place sur de plus en plus de moyens de communication « classiques ».[8] .

Dans cet article, nous traiterons donc des objets connectés et des données, qu'elles soient massives, intelligentes ou plus réduites en nombre, mais en tout cas sources d'informations des services de renseignement. Sont-ce de réelles opportunités ou remettent-elles plutôt en question le traditionnel cycle du renseignement, modèle universel d'exploitation fonctionnel de l'information[9] au service d'un décideur ? Comme le postulat essentiel du cycle du renseignement reste basé sur un principe : *« toute information collectée doit être exploitée avant d'être diffusée aux décideurs sous forme de connaissance opérationnelle, ou élément d'aide à la décision »*[10], nous poserons les questions : les services de renseignement (belges) doivent-ils adapter leur processus (cycle du renseignement) et leur structure organisationnelle pour faire face aux défis que représentent les *big data* et les objets connectés ? Comment doivent-ils entamer ou poursuivre cette évolution ?

Notre hypothèse est que les services de renseignement peuvent tirer avantage des objets connectés qui se multiplient et les données qu'ils transmettent, à différents niveaux. Pour cela, ils doivent prendre en compte des - nouvelles - variables, en renversant parfois certains paradigmes : le temps, la corrélation de données, la valeur de l'information sans que cela n'impacte le processus décisionnel dont le renseignement reste un des moteurs.

Nous abonderons dans le sens de certains spécialistes du renseignement qui ont relevé l'obsolescence du cycle de renseignement traditionnel et qui ont proposé d'autres modèles, que nous considérerons plus adaptés à l'impact des *big data* sur le processus décisionnel.

Nous proposerons une conclusion qui pose le débat de l'avènement des *big data* en terme d'efficacité, d'efficience et d'effectivité pour les services de renseignement de 'taille moyenne'.

[8] Le Monde, *Le directeur du renseignement américain reconnaît s'intéresser aux objets connectés*, 10/02/2016. En ligne sur le site web de Le Monde. http://www.lemonde.fr/pixels/article/2016/02/10/le-directeur-du-renseignement-americain-reconnait-s-interesser-aux-objets-connectes_4862587_4408996.html#9bjRf4mTzPGspPWS.99

[9] Franck Bulinge, « Le cycle du renseignement, analyse critique d'un modèle empirique », in *Market Management,* 2006/3, vol 6, p. 37.

[10] *Id*., p. 38.

2. Big data et objets connectés: de quoi parle-t-on ?

Un "**objet connecté"** est un matériel ayant la technologie adaptée pour communiquer avec un smartphone et lui transmettre des informations via une liaison sans fil. L'intérêt principal de ce genre d'objet est d'être interactif et de transmettre des données et statistiques.[11] Il existe plusieurs types d'objets connectés et plusieurs outils utiles à leur maîtrise : (1) Le vecteur/transmetteur : par exemple des bracelets d'activités (automesures, body connecting), un podomètre, une tablette/smartphone, une carte bancaire, mails, une carte Mobib[12], une voiture, des vêtements, un pilulier[13], un tapis de sommeil[14], toute la domotique ; (2) Le traceur GPS espion[15]: qui vous permet de savoir où se trouvent votre voiture, votre animal, entre autres, à tout moment. Ce boitier traceur GPS, très simple à utiliser, est compatible avec tous les smartphones du marché. Grâce à son microphone intégré, vous pouvez appeler le boitier traceur et écouter ce qu'il se passe autour ; (3) Des sites comme Shodan, un moteur de recherche permettant de prendre le contrôle d'objets connectés non protégés, se sont fait une spécialité d'inventorier les objets vulnérables.[16] "Lorsque Trendnet a commercialisé ses caméras de surveillance *SecureView*, elle a positionné ce produit dans le secteur de la sécurité du domicile, de la surveillance de bébé à distance, etc. Mais le logiciel présentait d'importantes failles de sécurité, permettant, à quiconque en possession de l'adresse IP de la caméra, de voir ou d'écouter ce qu'elle surveillait … et tout cela sur Internet."[17]

Les "petites données" (**small data)** proviennent de la combinaison de l'analyse des données (même avec des volumes limités, avec des outils décisionnels et prédictifs) et des processus de jugement collaboratifs (et itératifs) qui déterminent la prise de bonne décision. On peut même avancer que c'est ce qui fera, à l'avenir, la différence entre les organisations (entreprises ou administrations), dans la mesure où il sera de plus en plus difficile d'intégrer

---

[11] T2M, *Les objets connectés : attention, prêt, connecté !*. En ligne le 9/11/2015 http://www.time2marketing.fr/marketing3.0/objets-connectes.html

[12] Carte rechargeable permettant de se déplacer à Bruxelles en transports en commun.

[13] Boîtier connecté permettant de s'assurer qu'un proche ou un patient prend bien son traitement, le bon, à la bonne heure et en bonne quantité. (Vincent Alzieu, *CES : Medissimo imedipac. Un pilulier connecté, français*. 7/01/2014. En ligne http://www.lesnumeriques.com/capteur-mouvement/ces-medissimo-imedipac-pilulier-connecte-francais-n32530.html)

[14] Petit tapis piqué de capteurs à disposer sous le matelas. (Florence Legrand, *CES : Withings Aura, un assistant connecté pour mieux dormir,* 6/01/2014. En ligne sur le site web Les Numériques http://www.lesnumeriques.com/capteur-activite/ces-withings-aura-assistant-connecte-pour-mieux-dormir-n32523.html)

[15] En ligne le 5/09/2016 http://www.gps-traceur.com/63-tracker-gps-espion-miniature-gtx.html

[16] Le Monde, *INTRUSION 2.0 – Avec Shodan, contrôlez des webcams et imprimez chez les autres*, 10/06/2014. En ligne le 5/09/2016 http://bigbrowser.blog.lemonde.fr/2014/06/10/avec-shodan-controlez-des-webcams-et-imprimez-chez-les-autres/ & Pierre Delort, *Le Big data*, Paris, 2016, p. 27.

[17] Pierre Delort, *Le Big data*, Paris, 2016, p. 27.

des masses énormes d'informations dans les processus de décision – du moins sans disposer des outils adaptés.[18] Toutes n'auront pas les moyens de traiter du *Big data*.

Les "données massives" (**big data)** sont des données dont l'énorme volume d'informations, entraîne des changements dans le stockage, l'analyse et le traitement : Viktor Mayer-Schönberger et Kenneth Cukier expliquent que les *big data* se réfèrent à ce qui peut être fait à grande échelle et pas à une plus petite, avec comme impact la transformation des marchés, des organisations, de la relation entre les citoyens et les gouvernements.[19] La masse de données permet d'accéder à de nouveaux résultats[20] mais elle est parfois également son plus grand défaut dans la mesure où elle peut rendre les données inexploitables de par leur nombre. C'est là la plus-value des *smart data*.

Les "données intelligentes" (**smart data)**: le terme est emprunté au monde du trading, soit de l'exploitation d'opinions diffusées sur les réseaux sociaux en grande quantité qui permet de «capter» le sentiment d'une foule d'investisseurs et de déceler les « signaux faibles » de futurs krachs boursiers. Plus largement, il s'agit au moyen d'algorithmes d'extraire du Big data des données exploitables qui permettront des prédictions. Prenons l'exemple de « l'opinion mining », également appelé « analyse de sentiments ». La méthode consiste à extraire d'une masse de données les sentiments exprimés avant de les analyser et d'en tirer des tendances. La puissance de calcul algorithmique permettrait de suivre toutes les évolutions de l'opinion sur le web en temps réel, rien de moins.[21] On entre ici dans le giron du SOCMINT ou Social Media Intelligence, qui offre au travers de ce qui est exprimé sur les réseaux sociaux, de nouvelles opportunités pour la recherche d'informations et leur compréhension, dans un temps quasi-réel et avec l'ambition de prévoir, par exemple, des crimes et délits.

L' "internet des objets" **(internet of things - IoT)** est « constitué de l'ensemble des objets de la vie de tous les jours qui sont lisibles, reconnaissables, localisables, adressables et/ou contrôlables par Internet, par technologies RFID (*radio frequency identification*), LAN (réseau local) sans-fil, WAN (réseau sur grande distance) ... »[22] Dans cet IoT, des objets

[18] *Big data ou Small Data*, 04/05/2012. En ligne http://www.analysepredictive.fr/marketing-predictif/enjeux-marketing/big-data-ou-small-data
[19] Viktor Mayer-Schönberger & Kenneth Cukier, *Big data*, *A Revolution That Will Transform How We Live, Work, and Think,* Paris, 2014, p. 15.
[20] *Id.*, p. 20.
[21] Dominique Boullier & Audrey Lohard, *Opinion mining et Sentiment analysis*, 30/04/2012. En ligne http://books.openedition.org/oep/202
[22] Pierre Delort, *op. cit.*, p. 16.

parlent entre eux : un smartphone communique avec un réveil, des automobiles "refusent" de se percuter.

3. Renseignement et performance

Le renseignement peut-être abordé sous des angles divers : c'est un produit, une organisation, une aide à la décision (situational awareness) ou encore un processus. Mark Lowenthal définit le renseignement comme : *"Information is anything that can be known, regardless of how it is discovered. Intelligence refers to information that meets the stated or understood needs of policy makers and has been collected, processed and narrowed to meet those needs. Intelligence is a subset of the broader category of information. Intelligence and the entire process by which it is identified, obtained and analyzed respond to the needs of policy makers."*[23] Le processus itératif du renseignement relevé dans la définition est appelé cycle du renseignement et *« Le respect du cycle par les services de renseignement est le fondement essentiel d'une politique de renseignement efficiente »,* a écrit Sherman Kent, initiateur du cycle en 1947.

Comment cette *efficience* dont parle Sherman Kent, cette performance, est-elle assurée dès lors que les agences de renseignement sont amenées à devoir gérer un flux de plus en plus important et rapide d'informations ?

Le terme *performance*, issu de l'ancien français *parformer,* signifie l'accomplissement d'un processus, d'une tâche avec des résultats qui en découlent et le succès qu'on peut y attribuer[24]. Nous différencions la *performance-résultat,* qui se rapproche des objectifs à atteindre (a-t-on réalisé les objectifs ?), la *performance-action* qui distingue plutôt les compétences (est-on à même de réaliser les objectifs ?) et la *performance-succès* qui dépend du niveau d'ambition que l'on s'est fixé.

Si le vocable *performance* nous renvoie inconsciemment aux activités sportives, il convient de pouvoir mesurer, évaluer les prestations des agences de renseignement qui sont des administrations publiques qui animent des politiques publiques, faut-il le rappeler au service du citoyen. Elles sont consubstantielles à l'existence de l'Etat, à la défense de l'intérêt général ce qui sous-entend les valeurs de neutralité, objectivité, probité et responsabilité. En Belgique,

[23] Mark Lowenthal, *Intelligence. From secrets to policy,* CQ press, 2009.
[24] Yvon Pesqueux, « La notion de performance globale », *2003,* in Thierry Le Nedic, *La performance dans le secteur public, outils, acteurs et stratégie*, Ecole des Mines de Paris, 2009, p. 24.

les actions des services sont entreprises dans le respect des lois et rappelons que le Comité permanent de contrôle des services de renseignement – le Comité « R » - a été institué pour conserver *« ... l' équilibre nécessaire entre transparence, efficience, efficacité et légitimité »* dont la tâche est de « ... *confronter l'exécution de sa mission aux normes constitutionnelles et légales »*.[25]

L'*effectivité*[26] d'une administration est analysée par l'évaluation de l'adéquation entre les *outputs* (notre activité en tant que finalité – le produit à destination d'un client) et les impacts réels sur les groupes-cibles et leur degré d'utilisation ou de mise en oeuvre. Les effets (*outcomes*) déficitaires sont souvent imputés à des *outputs* ou des impacts manquants. Les politiques mises en œuvre qui ne produisent pas d'outputs (analyse avec hypothèses décisionnelles) ou les *outputs* sans impacts réels sont qualifiées d'i*neffective*. L'*efficacité* est analysée par l'évaluation de l'adéquation entre les *outcomes* réels (c'est-à-dire les effets de notre politique) et les objectifs visés par un plan directeur. Il s'agit de la prise de mesure des décideurs à la suite de la diffusion des *outputs*. Mais, il importe de ne pas tomber dans le piège de la simplification. Tout ne peut être quantitifé et le temps est une variable importante. La rapidité n'est pas nécessairement un gage de performance dans le renseignement. Il peut arriver que des objectifs soient atteints sans adéquation avec la politique développée ou que des objectifs formulés ne soient d'aucune aide pour mesurer l'efficacité d'une agence de renseignement. L'*efficience allocative* est évaluée par l'analyse du rapport entre les ressources investies et les *outcomes* réels. Deux approches analytiques sont utilisées : l'analyse coûts/bénéfices et l'analyse coûts/utilité. L'*efficience productive* analyse le rapport entre les ressources investies et les *outputs* réels. C'est cette approche de l'efficience qui correspond le mieux aux particularités d'une agence de renseignement : Peut-on produire les mêmes *outputs* avec moins de ressources, ou de meilleurs *outputs* avec autant de ressources, y compris les ressources intangibles ? Ceci nous paraît d'importance lorsqu'on aborde les *big data* et leur gestion : leur exploitation a-t-elle un impact sur les sous-processus qui animent le renseignement ?

[25] Sabine de Bethune, Présidente du Sénat, Préface de l'ouvrage *Regards sur le contrôle : vingt ans de contrôle démocratique sur les services de renseignement*, Wouter Van Laethem & Johan Vanderborght, Ed., Editions Interscientia, 2013, p. 7.

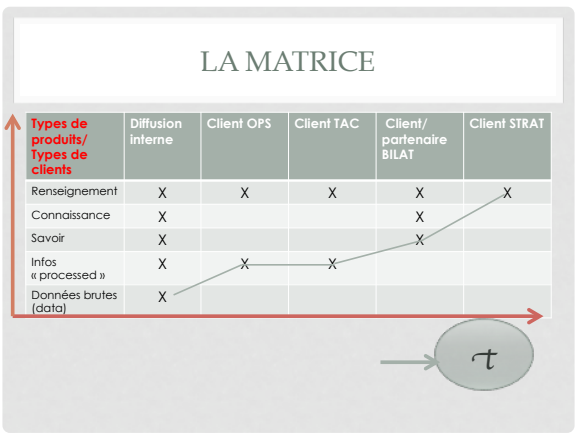
[26] Peter Knoepfel, Frédéric Varone, « Mesurer la performance publique: méfions-nous des terribles simplificateurs », in *Politiques et management public*, Vol 17, n°2, 1999, p. 129.

4. Cycle du renseignement et *big data*

Le cycle du renseignement tel que nous le connaissons a été proposé pour la première fois à la fin de la deuxième guerre mondiale par Sherman Kent. Popularisé ensuite par Harry Howe Ransom[27] et par la commission d'enquête sur les activités de la CIA à la fin des années cinquante, le cycle du renseignement devient un modèle hégémonique au sein des services de renseignement occidentaux. Depuis lors, la question d'une revision du cycle du renseignement se pose. L'environnement dans lequel les services opère évolue tout comme les risques et les menaces qu'ils doivent anticiper. Arthur Hulnick, en 2006, dans un article intitulé « what's wrong with the intelligence cycle ? »[28] écrit que *« ... there were serious problems with the intelligence cycle. It ignores two main parts of intelligence work : counter-intelligence and covert action. »* L'auteur insiste sur l'inapplicabilité du cycle du renseignement traditionnel aux pratiques de contre ingérence. Dans les pratiques quotidiennes du contre-renseignement, l' *intelligence driver* n'est pas le décideur politique (comme le représente le cycle du renseignement traditionnel) mais le responsable du contre-renseignement lui-même, mû par la volonté de *« filling the gaps »* dans les données, par le déclenchement du processus de collecte[29]. Un deuxième point de cette étude et utile à notre réflexion est la nécessité d'appliquer une méthode de travail séquentielle et non cyclique de la coopération entre l'analyse et la collecte, mais qui nécessite une vigilance permanente tant les barrières psychologiques sont encore grandes. Guillaume De Valk, dans une thèse de doctorat soutenue à l'université de Groningen (NDL) en 2005, propose plusieurs modèles de cycle de renseignement. L'auteur[30] exprime, outre le cycle traditionnel, l'idée d'un *modèle matriciel* totalement '*product-oriented'*, est basé sur le principe de la division du travail. Ce modèle répondrait mieux aux réalités quotidiennes des services de renseignement tant leurs complexités ne trouvent de réponses selon lui par le modèle traditionnel cyclique. Selon certains membres de la CIA, ajoute-t-il, cette matrice rendrait la relation, importante entre le renseignement et le politique, beaucoup plus claire. Ce modèle fonctionne sur base de trois piliers. Les deux premiers piliers (collectes et analyse) fonctionnent séparément mais de façon synchronisée. En effet, il existe une différence entre les exigences de la collecte et de l'analyse. Habituellement, la collecte cherche à combler les manquements dans des données

[27] Harry Ransom, *Central Intelligence and national security*, Cambridge, Harvard University Press, 1958. 272p.
[28] Arthur Hulnick, "What's wrong in the intelligence cycle", in *Intelligence and national security,* Vol 1, N°6, Decembre 2006, pp 959-979.
[29] *Id.,* p. 960.
[30] Guillaume Gustav De Valk, *Dutch Intelligence – Towards a qualitative framework for Analysis,* Rijksuniversiteit Groningen, 2005.

existantes, tandis que l'analyse cherche à répondre aux demandes d'un client- consommateur de renseignement. Au cours de la collecte, une certaine forme d'analyse est déjà appliquée, notamment lors d'une première évaluation des informations brutes, ou encore lors de la mise en forme de données dans un format utilisable ou leur mise en perspective. Le troisième pilier est nommé par l'auteur « support and services » et comprend des fonctions telles que bureau de traduction, sécurité physique du personnel et des mesures de contre-espionnage internes, entre autres, qui produisent aussi du renseignement à destination de la direction de l'entreprise. Ce modèle laisse la place aux procédures de feed-back au niveau de la collecte qui, selon l'auteur, ne se produisent pas dans un cycle du renseignement normal. L'avantage serait la différenciation en temps et en contenu de l'information collectée par la collecte en fonction du destinataire.



Partant de l'analyse d'Arthur Hulnick et de Guillaume De Valk, nous posons la question de l'*opérationalité* du cycle du renseignement traditionnel dès lors que le processus et les segments professionnels seraient confrontés aux *big data*. Faut-il adapter le processus du renseignement en fonction des *Big data* et *objets connectés* ?

4.1 La collecte

Si la collecte d'informations par source humaine (HUMINT) reste dans le monde du renseignement la méthode qui, qualitativement, est la plus efficace, il existe des moyens plus techniques comme le MASINT (measurement & signature intelligence), le SIGINT (signal intelligence) et ses subdivisions ELINT (electronic intelligence) et COMINT (communication intelligence). Les sources ouvertes (OSINT) et l'exploitation des données issues des réseaux

sociaux (SOCMINT), sont également devenus des sources d'informations importantes qui représentent en volume 80 % de l'information entrante. Aujourd'hui, la masse d'informations fournie par les *objets connectés*, les *big data* font évoluer les pourcentages tels que des filtres ou des systèmes d'exploitation sont nécessaires au point de voir le processus du renseignement bloqué par un engorgement d'informations. Le renseignement militaire français, par exemple, recrute des experts en ce sens[31]. Les *big data* peuvent en effet s'inscrire dans le descriptif traditionnel des types d'informations traitées par le renseignement :

L'information blanche est celle qui est facilement accessible : " Publique ou réservée, elle est issue de banques de données, publications scientifiques, périodiques, plaquettes d'entreprise, entretiens avec des experts, des fournisseurs, des clients, des partenaires ... Elle est donc libre d'accès et d'exploitation. L'internaute, par exemple, partage en ligne spontanément des données personnelles. Il suffit de les y cueillir (ce que contexte la commission sur la vie privée en Belgique[32]). Avec le développement des technologies de l'information, la masse de données disponibles est devenue énorme. Identifier l'information pertinente dans ce flux sans cesse grandissant nécessite de se doter d'outils informatiques adaptés.[33] La valeur de ces données est également latente et requiert une analyse novatrice pour la faire émerger[34], d'autant plus quand elle se révèle dans le temps, alors que sa première exploitation est déjà bien loin. Nous y reviendrons.

« *L'information grise* est licitement accessible, mais caractérisée par des difficultés dans la connaissance de son existence ou de son accès. »[35] Quelques exemples d'informations grises « exploitées » : (1) En Europe, des services publics installent des compteurs d'électricité « intelligents » qui collectent des données durant toute la journée. Or, le type de consommation varie d'un appareil électrique à l'autre et cette variation constitue la «signature électrique» spécifique à chaque appareil. La consommation d'énergie d'un ménage révèle donc des informations privées, que ce soit le comportement quotidien de l'habitant, son état de santé ou ses activités illégales.[36] ; (2) Twitter a conclu des accords avec deux entreprises pour vendre

[31] Jean-Luc F., *Devenez ingénieur Big data pour le renseignement militaire,* 21/09/2016. En ligne le 15/10/2016 http://www.defense.gouv.fr/ema/interarmees/la-direction-du-renseignement-militaire/servir-a-la-drm/les-metiers-de-la-drm/ingenieur-big-data/devenez-ingenieur-big-data-pour-le-renseignement-militaire
[32] Université d'été du parti socialiste, Sart Tilman, Université de Liège, 1er juillet 2016
[33] Le guide du routard de l'intelligence économique, 2012, Paris.
[34] Viktor Mayer-Schönberger & Kenneth Cukier, *op., cit.*, p. 142.
[35] Association des professionnels de l'information et de la documentation, *Information grise*. En ligne le 26/09/2016, http://www.adbs.fr/information-grise-17424.htm?RH=ACCUEIL
[36] Viktor Mayer-Schönberger & Kenneth Cukier, *op., cit.*, p. 186.

l'accès à ses données (source). De nombreuses entreprises font de l'analyse de tweets, parfois à l'aide de "l'opinion mining", afin de recueillir le retour de groupes de clients ou de juger l'impact de certaines campagnes.[37] ; (3) En France, le sénat entérine la Loi de programmation militaire (LPM) le 18 décembre 2013, laquelle autorise la police, la gendarmerie, ainsi que les services habilités des ministères de la Défense, de l'Économie et du Budget à surveiller les citoyens sur les réseaux informatiques sans l'autorisation d'un juge.[38] Le 29 juin 2013, dans un second article, l'hebdomadaire Der Spiegel révèle que la NSA reçoit quotidiennement les données issues des communications sur les réseaux téléphoniques français (jusqu'à 7 millions par jour en janvier 2013) et allemands.[39]

L'information noire est marquée du sceau de la confidentialité. Elle traite d'informations protégées par le secret (secret militaire, secrets de fabrication, secrets commerciaux) ou relatives à l'organisation (organigramme). Son accès est soumis à des risques de sanctions civiles et pénales (espionnage, vol, débauchage, corruption…) et son exploitation est libre uniquement si elle a préalablement été formellement autorisée. »[40]. Elle devient alors de l'information grise. Quelques illustrations qui fâchent : (1) La National Security Agency (NSA) surveille les communications de non-Américains qui transitent par les services de Google, Facebook, Yahoo ou encore Microsoft.[41] L'acronyme GAFA désigne Google Apple Facebook Amazon, des géants de la collecte de données personnelles. Quand c'est gratuit, c'est que vous êtes le produit. En septembre 2013, les quotidiens The Guardian et The New York Times révèlent que, depuis 2010, la NSA a développé de multiples méthodes de contournement des algorithmes de chiffrement utilisés par les communications sur Internet afin d'avoir accès aux contenus des messages.[42] Ceci a poussé certains services Internet à développer plus encore le chiffrement de leurs communications, ce qui ne fait pas les affaires

---

[37] *Id.,* p. 113.

[38] Philippe Vion-Dury, *Loi de programmation militaire : le scandale qui fait sploutch*, 19/12/2013. En ligne http://rue89.nouvelobs.com/2013/12/19/loi-programmation-militaire-scandale-fait-sploutch-248467 ; Frédéric Bergé, *Loi de programmation militaire : Jacques Attali juge "ahurissant" l'article 20*, 23/12/2013. En ligne http://www.01net.com/actualites/loi-de-programmation-militaire-jacques-attali-juge-ahurissant-larticle-20-610836.html ; La rédaction, *La loi de programmation militaire 2014 à 2019 est déjà promulguée*, 19/12/2013. En ligne http://www.zdnet.fr/actualites/la-loi-de-programmation-militaire-2014-a-2019-est-deja-promulguee-39796465.htm

[39] *The NSA's "Boundless Informant" Program* » [archive], sur http://www.spiegel.de/ [archive], Der Spiegel, 29/06/2013 & Laura Poitras, Marcel Rosenbach, Fidelius Schmid, Holger Stark & Jonathan Stock, *Cover Story: How the NSA Targets Germany and Europe*, 1/07/2013. En ligne http://www.spiegel.de/international/world/secret-documents-nsa-targeted-germany-and-eu-buildings-a-908609.html

[40] Le guide du routard de l'intelligence économique, 2012, Paris.

[41] Sébastian Seibt, *Le renseignement américain puise dans les données de Facebook et Google*, 8/06/2013. En ligne http://www.france24.com/fr/20130607-prism-espionnage-nsa-cybersurveillance-donnees-facebook-google-yahoo-microsoft-scandale-verizon

[42] Le Monde, *Cybersurveillance : la NSA sait déjouer le chiffrement des communications*, 5/09/2013. En ligne http://www.lemonde.fr/technologies/article/2013/09/05/cybersurveillance-la-nsa-a-contourne-les-garde-fous-qui-protegent-les-donnees_3472159_651865.html

de la police ou de la sûreté.[43] ; (2) Sept opérateurs de télécommunications mondiaux d'origine américaine ou britannique collaborent avec l'agence de renseignement électronique britannique, le GCHQ (Government Communications HeadQuarters). Dans le cadre du programme de surveillance britannique Tempora, les sociétés British Telecom, Vodafone Cable, Verizon Business, Global Crossing, Level, Viatel et Interoute ont en effet offert au GCHQ un accès illimité à leurs câbles. Ces câbles transportent une grande part des communications téléphoniques et du trafic internet mondial, couvrant notamment la France, les Pays-Bas ou l'Allemagne.[44]

La valeur de l'information

Toutes les données même les plus simples ou les plus anodines acquièrent une *valeur* dans les *big data,* lit-on chez Viktor Mayer-Schönberger et Kenneth Cukier. Dans le renseignement c'est aussi la cas. La valeur de l'information est immédiate et cachée : les informations brutes peuvent faire l'objet de traitements divers s'échelonnant dans le temps, au gré des besoins. C'est ce qui justifie la création et la gestion de banques de données et les exceptions – contrôlées - à la loi sur la conservation de données à caractère privé. Utiles pour le département « analyse » des services de renseignement, elles enrichissent le traitement des informations et leur transformation en renseignement, en connaissance, qui varient selon l'objectif, déterminé en partie par la nature du destinataire. Ainsi, un ministre n'a pas besoin des mêmes renseignements qu'un agent de terrain, à tout le moins pas diffusés sous la même forme. Il est communément admis dans le monde du renseignement que le produit est diffusé, taillé à la mesure du « client ».

Viktor Mayer-Schönberger et Kenneth Cukier, écrivent : « avec des quantités massives de données, la valeur de l'information ne se trouve plus seulement dans son objectif initial. Il faut compter maintenant avec ses réutilisations, ses nouvelles exploitations. Chose frappante, à l'ère des *big data*, la plupart des utilisations secondaires innovantes n'ont pas été envisagées au moment de la collecte initiale des données.»[45]. Ce qui compte maintenant, c'est la taille des bases de données, précisent-ils, autrement dit il s'agit de détenir de vastes ensembles de données et de pouvoir en recueillir toujours plus sans difficulté. Ainsi, les détenteurs de vastes

[43] Anouch Seydtaghia, *Inviolable, le chiffremetn de WhatsApp irrite le FBI,* 6/04/2016. En ligne https://www.letemps.ch/economie/2016/04/06/inviolable-chiffrement-whatsapp-irrite-fbi

[44] John Goetz & Frederik Obermaier, *Snowden enthüllt Namen der spähenden Telekomfirmen,* 2/08/2013. In Sueddeutsche Zeitung. En ligne http://www.sueddeutsche.de/digital/internet-ueberwachung-snowden-enthuellt-namen-der-spaehenden-telekomfirmen-1.1736791 & Le Monde, *Comment la Grande-Bretagne espionnait avec l'aide d'opérateurs*, 3/08/2013. En ligne http://www.lemonde.fr/technologies/article/2013/08/03/les-compagnies-de-telecom-complices-du-systeme-de-surveillance-britannique_3457203_651865.html

[45] Viktor Mayer-Schönberger & Kenneth Cukier, *op., cit.,* pp 186-187.

volumes de données vont prospérer grâce à la collecte et au stockage d'encore plus grandes quantités de ce qui constitue la matière première de leur activité et qu'ils pourront réutiliser pour créer davantage de valeur.[46] Certes, les capactiés de stockage doivent évoluer, mais aussi et surtout la façon dont il faut les extraire et les exploiter.

Rien de nouveau sous le soleil cependant : dans le renseignement, les données récoltées par les différentes méthodes de collectes, sur base du plan direceur approuvé par le ministre bien entendu, ont toujours été stockées pour des usages ultérieurs éventuels. L'adage en vigueur dans les services , *« un service de renseignement sans mémoire n'est pas un service de renseignement »* garde tout son sens. Il est rare que le renseignement collecte des données en « ratissant large ». L'utilisation de méthodes de collectes traditionnelles sont beaucoup plus ciblées, soit sur un individu, soit sur une organisation ou un phénomène. On s'éloigne de l'allégorie du filet de poissons. Si les mailles sont très fines, les poissons remontés seront de tailles diverses. Les petits poissons comme les gros s'y feront prendre, aucun filtre ne sera opéré. Il s'agira ensuite de trier selon des critères de taille, de poids, de nature, etc. Ce qui demande un temps supplémentaire. La valeur des données résidera dans les résultats finaux, obtenus en les utilisant de toutes les manières possibles. Une liste apparemment infinie dans laquelle il faudra faire des choix. La somme de ces choix donnera leur importance aux données. Mine de diamants productive bien après que sa valeur essentielle a été exploitée.[47]

Le traitement des données : corrélation *vs* pertinence ?

La plus-value en terme de performance proviendra, s'il apparaît que les *big data* soient (ou deviennent) une méthode de collecte utile au renseignement, de la façon dont elles seront traitées. *Corréler*, c'est étudier l'intensité de la liaison entre deux ou plusieurs variables nous dit le dictionnaire, tout en mettant en garde le lecteur : si les variables sont fortement corrélées, cela ne signifie pas pour autant qu'il y ait une relation de causalité. Si l'on récolte les données massivement, sans filtre, sans algorithme, on rassemble une quantité de données brutes mais encore faut-il leur appliquer ensuite un traitement efficace. Celui qui permettrait d'en tirer toute la valeur utile, celle d'aujourd'hui mais aussi celle de demain. Parce qu'une donnée n'a pas qu'une valeur mais bien un nombre indéfini de valeurs qui se révéleront dans le temps au gré des traitements appliqués et de l'inventivité humaine. Les pratiques usuelles du renseignement sont présentées aujourd'hui comme un coup de « génie » de Google, celui d'avoir fait le pari qu'il était utile de conserver les données après épuisement de leur utilité,

[46] *id.*, p. 178.
[47] *Ibid.,*p. 128.

malgré le coût que cela représente ! Google a fait le pari que les exaoctets de requêtes passées, agrégées et anonymisées, pourraient avoir d'autres utilités.[48] *MapReduce,* produit Google, est né du besoin de stockage. *MapReduce* est en fait le cœur de *Hadoop®,* une plate-forme de stockage évolutive conçue pour traiter de très grands ensembles de données.[49] *MapReduce* a permis d'indexer le Web en coordonnant jusqu'à des centaines de milliers de serveurs.[50] Ce qui inquiète le plus les adeptes des méthodes de renseignement traditionnelles c'est la fiabilité de l'information, la pertinence de l'information : produit-on encore du renseignement avec des *big data* ? La réponse fournie par les spécialistes single : Ce que nous perdons en terme de précision à un niveau micro, nous le gagnons en compréhension au niveau macro. Ainsi, si des "petites données" (micro) sont inexactes, les *big data* (macro) effacent ces inexactitudes dès qu'un certain nombre est dépassé.[51] Avec les *big data,* les chiffres doivent être davantage abordés en terme de probabilité que de précision.[52] L'hyper-volume de données fait-il donc la qualité du renseignement ?

Les prédictions basées sur des corrélations sont au cœur des *big data*.[53] Le traitement de gros volumes de données permet de dégager des corrélations, des tendances cherchant à prédire le futur. Mais il y a un hiatus important à souligner. L'homme a tendance à penser que l'évolution de deux phénomènes de concert prouve un lien de type cause à effet entre eux. Or, répétons-le, corrélation ne veut pas dire causalité. Le traitement des *big data* permet de dégager de telles corrélations, corrélations dont on peut très bien se suffire sans pour autant s'interroger sur le pourquoi. Partant du constat d'une corrélation, une action peut être entreprise. C'est un peu l'optique actuelle. Ainsi, les systèmes innovants de recommandations mis au point par Amazon ont fait ressortir des corrélations sous-jacentes. Savoir quoi, et non pourquoi, est largement suffisant.[54] Les méthodes non causales sont imbattables, écrit Pierre Delort, président de l'Association nationale des directeurs des systèmes d'information. La réflexion rapide et intuitive fait ses preuves dans de plus en plus de cas. Cette réflexion vaut mieux que de longues et coûteuses expériences soigneusement contrôlées. En outre, il existe

[48] Pierre Delort, *Le Big data*, Paris, 2016, p. 40.
[49] IBM, *What is Hadoop ?* En ligne le 5/09/2016
http://www.ibm.com/analytics/us/en/technology/hadoop/#hadooparchitecture
[50] Pierre Delort, *op. cit.*, p. 57.
[51] Viktor Mayer-Schönberger & Kenneth Cukier, *op. cit.*, p. 23.
[52] *Id.,* p. 48.
[53] *Ibid.,* p. 71.
[54] *Ibidem.,* p. 68.

aussi un risque à tirer des causalités à partir de données qui reflètent par nature le passé: rien ne garantit que le futur soit semblable au passé.[55]

4.2 L'analyse

« Savoir quoi, et non pourquoi, est largement suffisant », lit-on chez Viktor Mayer-Schönberg et Kenneth Cukier. Certes pas dans le renseignement ! L'analyse, comme sous-processus dans le cycle reste d'importance. Il s'agit de l'ensemble des activités de validation de traitement, de transformation de l'information en connaissance. Si nous nous en tenons aux finalités du renseignement, la phase de collectes permet de répondre aux questions « quoi ? », « quand ? » et « où ? », en réponse à un problème, une crise par exemple. L'analyse qui produit le renseignement y ajoute les réponses « pourquoi et comment ? ». C'est le cœur de métier du renseignement dont l'objectif recherché est de fournir un produit utile au décideur, qui doit prendre … la bonne décision.

Rappelons que la *donnée (data)* n'est que le résultat d'un processus d'acquisition dont le contenu explicatif élémentaire est disponible dans un espace de communication. Cela vaut pour les *big data.* Et la donnée n'est inséparable de celui qui la recueille (identification, sélection, validation, hiérarchisation) ce qui entraîne une part de subjectivité. L'*information* est ce qui est donné à connaître. Elle est brute, subjective ou objective. Elle est ce qu'on en fait (processus d'agrégation d'éléments et de transformation en connaissance). La *connaissance*[56] est un état socio-cognitif résultant de la mise en cohérence d'informations et de la représentation mentale (individu) ou sociale (groupe) à un moment donné. C'est la remise en cause permanente des acquis (savoir) et elle est par nature stratégique. Franck Bulinge précise que *l'intelligence informationnelle*, ou la capacité à identifier quelle information est nécessaire, de la localiser et d'évaluer sa valeur, repose sur la prise en compte des espaces cryptiques dans le cadre du respect des informations protégées.

L'analyse, pilier du cycle du renseignement, s'en trouve donc concernée.

Mais une science des données (*data science*) émerge aujourd'hui, caractérisée par une collecte massive et variée de données, associée à des méthodes de traitements pour en extraire des

[55] Pierre Delort, *op., cit.*, 2016.
[56] Franck Bulinge, *Maîtriser l'information stratégique,* Editions Deboeck, 2014, p. 36.

connaissances nouvelles.[57] C'est grâce à de bons algorithmes que les données collectées pourront prendre toute leur valeur, immédiate et à plus long terme, que ce soit à partir de ce que l'on appelle maintenant, à côté des *big data* les *small data* et les *smart data*. Les *small data* plus opérationnelles ne permettent pas l'analyse prédictive des *smart data* ni la détermination de tendances des troisièmes. Le *big data* consiste à créer en exploratoire et par induction, voire intuition, et sur des masses de données à faible densité en information des modèles à capacité prédictive.[58] – Avec les *big data*, il s'agit d'appliquer des modèles mathématiques pour inférer des probabilités: sommeil, localisation, parcours sportifs[59], santé (physique et psychique), données bancaires, achats en ligne, contenu de correspondances, etc.

	stratégique	**prédictif**	**opérationnel**
BIG	**X**		
SMART		**X**	
SMALL			**X**

Le traitement des *small data* s'inscrit dans le présent et permet d'obtenir des résultats directement exploitables, rapidement opérationnels. Le traitement des *big data* s'inscrit dans un futur plus lointain et permet de dégager des corrélations dont l'analyse guidera la stratégie. Leur masse rend toutefois leur traitement infaisable pour la majorité des entreprises. Quant au traitement de *smart data* s'inscrit dans un futur proche et vise à obtenir des prédictions sur base de données pré-sélectionnées dans le *big data*.

Ce sont les connaissances nouvelles qui émergeraient de la gestion des *big data*, *smart data* et *small data*, qui, en complément des méthodes de collectes traditionnelles, qui apporterait une plus-value au cycle du renseignement ?

---

[57] Khalid Benabdeslem, Christophe Biernacki & Mustapha Lebbah, *Les trois défis du Big data - Éléments de réflexion*, juin 2015. En ligne sur le site web de la Société française de statistique http://statistique-et-societe.fr/ojs/index.php/stat_soc/article/view/445/419

[58] Pierre Delort, *op., cit.,* p. 42.

[59] Le secteur du sport : se motiver et partager ses données grâce à une montre connectée. Ce genre de produit séduit de plus en plus de personnes. Connaitre le nombre de kilomètres exacts courus, à quelle vitesse et quelle fréquence cardiaque. De plus, il est maintenant possible de partager ses performances avec une plate-forme de coaching en ligne. En ligne http://www.time2marketing.fr/marketing3.0/objets-connectes.html

Pour cela, nous devons analyser l'impact de l'inclusion des *big data* sur le processus décisionnel nourri par le cycle du renseignement.

4.3 La prise de décision

La variable *temps* - nous nous référons à l'article de Guy Goemanne dans le présent ouvrage – est aujourd'hui omniprésente dans le processus de décision, y compris dans le renseignement. A côté de l'*infobésité* qui impacte le cycle, si elle n'est pas maîtrisée, la dictature de l'immédiateté, *l'accélération du temps* imposent à notre sens une nouvelle lecture du processus. Nous renvoyons au point 4 de cet article. Les services de renseignement doivent en effet adapter leur organisation, leurs procédures en fonction des entrées massives d'informations qui nécessitent un traitement efficace, plus rapide, grâce, nous l'avons vu à des filtres, des algortihmes (*pre-programmed datamining engines*)[60] qui pré-sélectionnent celles qui doivent être accessibles plus rapidement par les analystes des services. En outre, il est aujourd'hui évident que les collecteurs et les analystes doivent être co-localisés afin de raccourcir certaines boucles du cycle du renseignement et pousser ainsi le dialogue en temps réel entre ces deux segments professionnels du renseignement. La prise de décision ne doit plus être centralisée par une structure hiérarchique verticale. Elle doit être répartie dans une structure en réseau : accroître l'accessibilité des renseignements c'est raccourcir le *'temps-décisionnel'*. C'est tout le cycle du renseignement qui s'en trouve impacté.

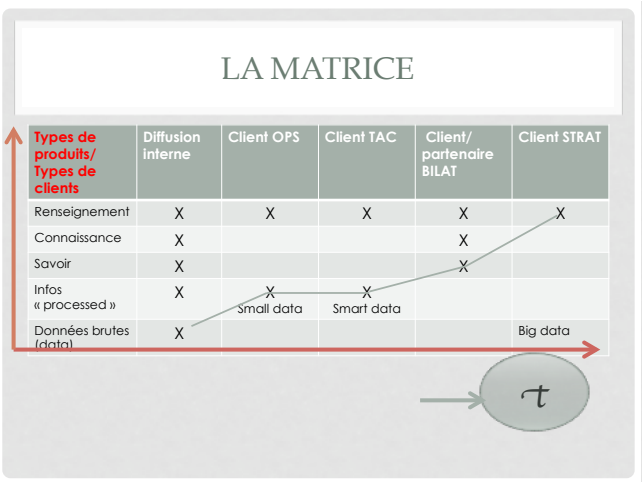
La rapidité du processus décisionnel, le modèle dominant de l'*intelligence for action* ne gomment pas l'absolue nécessité de maintenir une analyse stratégique sur ce qui se passe dans le monde afin d'assurer des jugements prédicitfs (*predictive judgments)*[61] mais aussi d'inciter le développement d'études de prospective, comme le relevait le *livre blanc* de la Défense en France (2013). Le chapitre 6 insistait sur la connaissance et l'anticipation et consacrait la proximité de la prospective et du renseignement, en mettant en exergue le besoin d'anticipation stratégique[62]. L'anticipation reste bien le cœur de métier du renseignement.

Une décision dans la sphère du renseignement peut-elle être prise, basée uniquement sur des *big data* ? La plupart des décideurs appuient leurs décisions sur une combinaison de fait, de réflexion et d'hypothèses. L'analyse des *big data* peut confirmer certaines hypothèses émises

[60] David Omand, *Securing the State,* Editions C. Hurst & Co. Ltd, 2011, p292
[61] *Id.,* p294
[62] Le livre blanc, défense et sécurité nationale, Direction de l'information légale et administrative, Paris, 2013,

entre autres par le choix d'une modélisation d'analyse prédictive. L'importante notion du *temps* y est prise en compte : Les *small data*, données opérationnelles, permettent de répondre rapidement à une demande de terrain, hic et nunc, tandis que les *big data*, données massives, permettent de dégager des tendances, de formuler des recommandations et s'inscrivent dans un processus décisionnel plus stratégique. Les *smart data* quant à elles cherchent à prédire et en cela s'inscrivent dans le futur. Ainsi, il est question de présent, de futur proche et de futur plus lointain.

Remarque importante cependant : si les *big data* sont surtout utiles pour des prévisions, elles ne le sont pas pour déclarer la culpabilité de quelqu'un, écrivent Viktor Mayer-Schönberger et Kenneth Cukier[63]. La corrélation est inadéquate pour juger les actes, précisent-ils.



Si nous croisons les données reprises sur la matrice du renseignement au point 4 et le tableau repris au point 4.2, les *big, smart* et *small data* s'inscrivent en complément des types de produits traditionnels diffusés aux clients des services, en respectant la variable *temps*, mais aussi la notion de renseignement. Le débat mené dans les services sur le fait de pouvoir diffuser des produits qui ne seraient pas le fruit de l'ensemble du cycle du renseignement, c'est-à-dire diffuser de l'information brute ou 'processed' est intense. Il en va de même par la force des choses avec les *big, smart* et *small data.*

[63] Viktor Mayer-Schönberger & Kenneth Cukier, *op., cit.,* p 198

Le renseignement ne peut faire l'impasse de nouvelles sources d'informations pour autant qu'elles soient exploitables en fonction de leur valeur et de leur masse. Accroissent-elles la performance des serives de renseignement et des risques sont-ils envisageables ?

5. *Big data* et performance

L'efficacité, l'efficience, d'un service de renseignement relève de sa faculté et de sa volonté de s'adapter à l'environnement dans lequel il opère, adaptation idéalement entreprise *a priori*. Cela concerne bien entendu les nouvelles sources d'informations. Ce fût le cas par exemple pour la collecte d'informations sur ou par les réseaux sociaux. La plupart des services ont créé en leur sein des département SOCMINT (Social Media Intelligence). La CIA, par exemple, considère l'avènement des *big data* comme un 'big deal'[64], une opportunité pour faire la différence. Un nouveau directorat est apparu dans la structure de l'Agence : le Directorate in Digital Innovation – DDI[65]. Les services de renseignement de taille moyenne ou réduite aujourd'hui n'ont pas la capacité de collecter les données en masse, et s'ils en ont la capacité, ils ne possèdent pas nécessairement les moyens techniques de les exploiter. En revanche ils disposent de nombreuses « petites données » plus facilement exploitables qui leur permettent déjà d'accroître leur connaissance sur les phénomènes dont ils ont la charge. Finalement, *«Ce n'est pas la taille qui compte,* écrit Rufus Pollock , *c'est d'avoir des données —peu importe leur taille— qui nous permettent de résoudre notre problème ou de répondre à la question que nous nous posons».*[66]

Des adaptations sont-elles nécessaires ?

5.1 Au niveau de la structure organisationnelle

Pour que les *big data* soient une opportunité, il convient de confirmer la structure en réseau des services de renseignement (ou de l'implanter si ce n'est pas encore le cas) afin d'accroître la circulation rapide de certaines données et de diluer le processus de décision vers des

[64] https://www.cia.gov/news-information/featured-story-archive/2012-featured-story-archive/big-data-at-the-cia.html, consulté le 18 novembre 2016
[65] https://www.cia.gov/news-information/featured-story-archive/2012-featured-story-archive/big-data-at-the-cia.html, consulté le 18 novembre 2016
[66] Jean-Laurent Cassely, *Big data: ce n'est pas la taille qui compte*, 13/05/2013. En ligne le 7/09/2016 sur http://www.slate.fr/economie/72361/big-data-small-data

sphères plus agiles, plus fluides. Les structures hiérarchiques, à communication verticale sont à ce titre inefficaces. La proximité de la collecte et de l'analyse, structure déjà appliquée dans certains services, notamment en Belgique où la co-localisation des deux segments professionnels existe pour certaines matières traitées, nous paraît nécessaire car elle réduit la lenteur du flux de l'information en leur sein. Cette proximité était déjà prônée par David Omand en 2011[67].

5.2 Au niveau de l'analyse : l'analyse prédictive

« The ultimate goal of the new Directorate, the collaboration between the Directorates, the 10 Missions and the greater efforts of the CIA as an organization is to provide what is called "Anticipatory Intelligence," or the ability to interpret data to anticipate future events and, if needed, take action. », peut-on lire au sujet du DDI[68]

L'anticipation reste le cœur du métier du renseignement. Anticiper, c'est prévoir, supposer ce qui va arriver et adapter sa conduite à cette supposition, lit-on dans le Larousse. Empruntée au monde du trading, l'analyse prédictive consiste en l'exploitation d'opinions diffusées sur les réseaux sociaux en grande quantité, ce qui permettrait de «capter» le sentiment d'une foule d'investisseurs et de déceler les « signaux faibles » de futurs krachs boursiers. On comprend ici tout l'intérêt du renseignement pour l'analyse de ce type d'informations et l'intérêt d'appliquer cette technique pour aider à circonscrire, par exemple, le risque terroriste.[69] L'idée serait de capter sur la toile des termes, des expressions-clefs qui permettraient de débusquer un discours signalant un risque de passage à l'acte. Creusons l'idée.

Certains fonds spéculatifs prédisent la performance du marché boursier en faisant l'analyse des tweets.[70] *MarketPsych* en Californie a commencé à analyser le contenu de tweets mis en données et à les interpréter comme indicateurs de placement en bourse. En association avec Reuters, elle propose 18.864 indices boursiers distincts dans 119 pays, sur des états émotionnels tels l'optimisme, la morosité, la joie, la peur, la colère et même des thèmes comme l'innovation, les litiges et les conflits.[71] Dans un autre registre, le *web bot project* ne

[67] David Omand, *op., cit.,* p291
[68] http://www.smartdatacollective.com/xanderscho/364797/how-cia-reinventing-case-big-data, consulté le 18 novembre 2016
[69] François Assémat, *Traders et terroristes : les risques extrêmes en 2016,* 6/01/2016. In La Tribune. En ligne http://www.latribune.fr/opinions/tribunes/traders-et-terroristes-les-risques-extremes-en-2016-540285.html
[70] Viktor Mayer-Schönberger & Kenneth Cukier, *op. ,cit.*, p. 24.
[71] Viktor Mayer-Schönberger & Kenneth Cukier, *op., cit.*, p. 114.

procède pas autrement sur le fond: Clif High[72] et George Ure prétendent avoir développé un logiciel capable de relever les changements de termes sur Internet et d'en tirer des tendances menant à la prédiction d'événements tels que la fin du monde de 2012, le tsunami, etc. Donc ils disent partir de *big data* et opérer un traitement soulignant des changements dans une sorte d'inconscient collectif virtuel. Dans une émission radio du 21 juillet 2009, George Ure et Clif High présentent leur *web bot technology* qui prédit nombre d'événements négatifs pour les quelques années à venir.[73]

Le *GDELT Project*, pour *Global Database of Events, Language, and Tone*, consiste en une collection d'un quart de milliard d'enregistrements d'événements du monde entier, de 1979 à nos jours. Elle est accompagnée d'un diagramme représentant un réseau gigantesque mettant en connection chaque personne, organisation, lieu et thème avec cette base de données d'événements. Sa vision est de tirer parti de ces données pour construire un catalogue de comportement et de croyances de la société à l'échelle humaine et dans tous les pays du monde. Le but est de capturer ce qui se passe dans le monde, quel est son contexte, qui est impliqué et comment le monde se sent à ce sujet, chaque jour.[74] Le but est de dessiner une carte répertoriant toute manifestation dans l'idée de prédire de nouveaux foyers probables de conflits.[75]

Le film *Minority Report* aurait-il été visionnaire ? La "predictive policing" utilise une approche mathématique et analytique dans le but de détecter les délits et acteurs de terrorisme imminents, soit avant qu'ils ne soient commis. [76] La ville de Santa Cruz a été la première à se doter, en juillet 2011, du programme *PredPol*. Grâce à une base de données recensant les infractions passées, la formule mathématique, l'algorithme, permet d'aiguiller les forces de l'ordre.[77] *Predpol* s'est depuis invité en France où la gendarmerie nationale, au travers de son service central de renseignement criminel, teste un programme d'analyses prédictives en

[72] Vinny Eastwood, *How To Predict The Future With Linguistics And Technology, Cliff High.* En ligne le 13/04/2015 https://youtu.be/VLnon-t9fVQ?t=1m42s
[73] George Noory, *Show with Clif High and George Ure*, 21/07/2009. En ligne http://www.coasttocoastam.com/show/2009/07/21
[74] http://gdeltproject.org/about.html#creation
[75] Mathilde Sagaire, *Peut-on prédire les conflits ?*, 4/12/2013. In Libération. En ligne http://www.liberation.fr/planete/2013/12/04/des-cartes-pour-predire-les-futures-zones-de-conflit_962635
[76] Ronald Meeus Longread, *Comment iPolice peut sécuriser la nation*, 24/06/2016. En ligne http://datanews.levif.be/ict/actualite/longread-comment-ipolice-peut-securiser-la-nation/article-longread-516395.html?utm_campaign=Echobox&utm_medium=social&utm_source=Facebook
[77] Louise Couvelaire, *Le logiciel qui prédit les délits*, 4/01/2013. In *Le Monde.* En ligne http://www.lemonde.fr/ameriques/article/2013/01/04/le-logiciel-qui-predit-les-delits_1812195_3222.html

matière de délinquance et vol de voitures.[78] Le *big data* représente une réelle opportunité pour augmenter l'efficacité dans la lutte contre le terrorisme. L'analyse promet d'identifier des réseaux terroristes et leurs associations en utilisant l'open source intelligence (OSINT) combinées avec des sources de renseignement traditionnelles. Elle peut également travailler rapidement à l'identification de sources radicales au sein de communautés en ligne en utilisant des fonctionnalités telles que l'analyse sociale des médias (SOCMINT) pour identifier les réseaux informels, les sujets émergents, les influenceurs, les liens entre les individus, les groupes, ou des concepts, et l'opinion mining.[79] L'espoir est de pouvoir également définir un profil du délinquant, voire du terroriste. Le *Future Attribute Screening Technology* (FAST) est précisément un outil qui cherche à identifier ces terroristes. En surveillant des constantes vitales, des attitudes et expressions non-verbales, la police espère repérer les personnes susceptibles de passer à l'acte.[80] Toutefois, il faut rester prudent et parler ici de probabilités qu'un comportement donné se présente effectivement. Si la probabilité est de 75%, il restera toujours 25% de probabilité que la prédiction ne se réalise pas. Ainsi, même s'il faut exploiter ce type de programme algorithmique, l'analyse humaine reste définitivement au coeur du système. Il faut des personnes pour réaliser l'analyse finale de toutes ces données. Car si les ordinateurs sont bons dans l'attribution d'un score de suspicion, un esprit analytique, humain, reste nécessaire.[81]

La loi française du 24 juillet 2015 sur le renseignement dispose que des boitiers pourront être disposés par les fournisseurs d'accès Internet afin d'enregistrer les métadonnées accompagnant les communications. Adéquatement traitées, cette masse de métadonnées pourraient livrer des informations utiles à la détection de comportements suspects. L'algorithme est en cours de finalisation. Cette technique de recueil de renseignements ne pourra toutefois être appliquée que dans le cadre de la lutte contre le terrorisme.[82]

---

[78] Institut national des Hautes Etudes de la Sécurité et de la Justice, *Vers une police 3.0, enjeux et perspectives à l'horizon 2025,* juin 2016. En ligne https://www.inhesj.fr/sites/default/files/fichiers_site/les_publications/les_travaux_des_auditeurs/gds3.pdf

[79] Babak Akhgar, Gregory B. Saathoff, Hamid R Arabnia, Richard Hill, Andrew Staniforth & Petra Saskia Bayerl, *Application of Big data for National Security*, Oxford (UK), 2015, p. 35.

[80] Viktor Mayer-Schönberger & Kenneth Cukier, *op., cit.,* p. 194.

[81] Ronald Meeus Longread, *Comment iPolice peut sécuriser la nation*, 24/06/2016. En ligne http://datanews.levif.be/ict/actualite/longread-comment-ipolice-peut-securiser-la-nation/article-longread-516395.html?utm_campaign=Echobox&utm_medium=social&utm_source=Facebook

[82] Thomas Cavaillé-Fol, *De nouveaux algorithmes pour scruter le cyberespace*, octobre 2016. In Sciences et Vie. & Direction de l'information légale et administrative, *Loi du 24 juillet 2015 relative au renseignement*, 1/12/2015. En ligne le 8/10/2016 http://www.vie-publique.fr/actualite/panorama/texte-discussion/projet-loi-relatif-au-renseignement.html

5.3 Au niveau légal et éthique

Certaines dérives possibles doivent être soulignés dès lors qu'elles entâchent la performance des services de renseignement, notamment en terme de légalité et d'éthique.

Entre sécurité publique et vie privée

L'efficacité des services de police et de renseignement est parfois mise à mal quant au nom du respect de la vie privée, les données sont davantage protégées. Tout est affaire de mesure et de point de vue. Il semble toutefois illusoire de penser qu'on peut anonymiser totalement des données. Car même quand les données ne paraissent pas explicitement porter sur des informations personnelles, elles permettent, au travers de procédures employées avec les *big data*, de remonter jusqu'à la personne concernée, dont on peut déduire des détails intimes.[83] Selon Paul Ohm, professeur de droit, il n'existe pas de solution miracle en matière de désanonymisation. A partir d'un certain volume de données, une anonymisation parfaite est impossible, quels que soient les efforts déployés.[84]

Dans un autre registre, la protection des communications, si elle sert la vie privée, peut déservir la sécurité publique. Ainsi le chiffrement des données de *WhatsApp* a été pointé du doigt: "Nous avons intégré le chiffrement de bout en bout dans les dernières versions de notre application. (…) vos messages, photos, vidéos, messages vocaux, documents et appels sont protégés pour ne pas tomber entre de mauvaises mains. Le chiffrement de bout en bout de *WhatsApp* garantit que seuls vous et la personne avec qui vous communiquez pouvez lire ce qui est envoyé; il n'y a donc pas d'intermédiaires, pas même *WhatsApp.* (…) Afin d'assurer une protection supplémentaire, chaque message que vous envoyez a son propre cadenas unique et sa clé unique.[85] Si *WhatsApp* contente par là ses utilisateurs, elle fait beaucoup moins d'heureux parmi les services de police et de renseignement dont la tâche est rendue plus compliquée. La difficulté est de situer le juste milieu dans un débat opposant la sécurité publique à la vie privée. Dans le même registre, on pourrait évoquer le cas d'un iPhone qu'il fallait craquer pour raisons d'enquête terroriste, ce qu'Apple refusait pour ces mêmes raisons de protection de la vie privée. Impasse ! Jusqu'à ce que l'iPhone soit piraté à la demande du FBI, délivrant ainsi les informations tant attendues par les policiers.[86]

---

[83] Viktor Mayer-Schönberger & Kenneth Cukier, *op., cit.* p.186.
[84] *Id.,* p. 190.
[85] Whatsapp, *Le chiffrement de bout en bout*, 5/06/2016. En ligne https://www.whatsapp.com/faq/fr/general/28030015
[86] Le Monde, *Le FBI assigné en justice pour révéler comment il a pu débloquer l'iPhone de San Bernardino*, 16/09/2016. En ligne http://www.lemonde.fr/conflit-apple-fbi/article/2016/09/16/le-fbi-assigne-en-justice-pour-reveler-comment-il-a-pu-debloquer-l-iphone-de-san-bernardino_4999085_4870067.html

Détournement de la finalité, à l'insu de tous ... ou presque

Test-Achats, association belge de protection des consommateurs, a soulevé un problème éthique en matière de données de santé: En Belgique, l'utilisation des données de santé est en principe interdite, sauf si la personne concernée donne son consentement explicite. Le consommateur doit exactement savoir de quelle manière ses données seront utilisées, et également, indépendamment de toute pression, marquer son accord à ce sujet. Les données de santé bénéficient donc en Belgique d'une protection renforcée. Toutefois, les données, par exemple, sur la qualité du sommeil ou sur l'IMC peuvent-elles être qualifiées de données de « santé » ? Ou parle-t-on plutôt de données sur le « bien-être » d'une personne ? Le statut des données recueillies via ces applications et d'autres appareils de mesure reste pour le moins incertain. De surcroît, les utilisateurs de ces applications et autres gadgets de mesure sont rarement informés de ce qui advient des données collectées. En outre, leur consentement explicite pour l'utilisation de ces données de santé n'est, dans la majorité des cas, jamais demandé. Test-Achats demande donc à la Commission de la vie privée d'examiner la problématique.[87]

Un autre problème se pose lorsque l'on prend une photo. Le plus important réside dans ce qui est visible, c'est une question de point de vue. Mais lorsque que l'on prend une photo avec un smartphone ou avec un appareil photo numérique, l'image est généralement accompagnée de caractéristiques appelées métadonnées. Celles-ci sont écrites dans le format «EXIF» (Exchangeable Image File Format). On retrouve au rang de ces métadonnées la date, l'heure, les caractéristiques techniques de l'appareil, les réglages de la prise de vue, etc. Mais on retrouve surtout dans ces métadonnées la géolocalisation de l'image. Au centimètre près. Nombre d'utilisateurs l'ignorent et dès lors ne modifient pas ce paramètre par défaut. Il suffit ensuite, et le petit logiciel est facilement accessible en ligne, d'extraire les métadonnées des photos visées pour en obtenir les coordonnées GPS. Ainsi, si, par exemple, si l'on met en ligne la photo d'un objet que vous souhaitez vendre, le risque est grand de communiquer votre adresse, à votre insu … à n'importe qui. Cet « état de fait » est exploité par les services de Gendarmerie nationale française qui, à l'aide du petit utilitaire GendEXIF conçu pour répondre à leurs besoins opérationnels, extraient les coordonnées GPS des photos trouvées dans les smartphones de suspects. Ils les transfèrent ensuite sur un fond cartographique tel que Google Maps dont la fonction Street View permet une vue terrain à 360°. La méthode s'est

[87] Test-Achats, *Que deviennent les données d'« automesure »? Test-Achats tire la sonnette d'alarme*, 14/10/2015. En ligne https://www.test-achats.be/action/espace-presse/communiques-de-presse/2015/privacy

révélée très efficace dans le cadre d'enquêtes pour terrorisme. Mais la géolocalisation du smartphone peut être désactivée, ce qui empêchera les coordonnées GPS d'accompagner les photos. Des logiciels existent également pour supprimer après coup les métadonnées accompagnant les photos. Et certains réseaux sociaux effacent les métadonnées de photos uploadées.[88]

Big data sous le Prism de la NSA

Le système va plus loin quand il organiserait la transmission de données via GAFA[89] à la NSA. C'est ce qu'Edward Snowden a dénoncé: PRISM prévoit le ciblage de personnes vivant en dehors des Etats-Unis. Ce programme fut créé en décembre 2007, dans le cadre du Protect America Act de 2007 et du FISA[90] Amendments Act. Selon The Register, un magazine britannique d'actualité technologique, ces amendements « autorisent explicitement les agences de renseignements à surveiller, pour une durée maximale d'une semaine, les appels téléphoniques, les courriels et d'autres communications de citoyens américains sans mandat d'un tribunal » quand l'une des parties n'est pas sur le sol des Etats-Unis.[91] Un document, remis par Snowden, explique que PRISM est « la source première de renseignements bruts utilisés pour rédiger les rapports analytiques de la NSA.[92] PRISM était autorisé par la Foreign Intelligence Surveillance Court. Le 29 juin 2013, le Washington Post publie quatre nouvelles diapositives de présentation du programme PRISM qui montrent qu'il permet de surveiller en temps réel les courriels, les communications par «chat», la participation à des forums de discussion, la diffusion de photos et de vidéos et les appels téléphoniques de «cibles».[93] Ainsi, au travers de PRISM, la NSA disposerait d'un accès direct aux données hébergées par les géants américains des nouvelles technologies, parmi lesquels Google, Facebook, YouTube, Microsoft, Yahoo!, Skype, AOL et Apple.[94] De plus, Microsoft et le FBI ont développé une solution permettant l'interception des "chats" cryptés d'Outlook.com avant que ce service ne

[88] Sandrine Mathen, *Sachez-le, vos photos vous trahissent !*, 27/06/2016. En ligne http://www.reputation365.eu/analyses-decryptages/vos-photos-vous-trahissent-exif/
[89] Google Apple Facebook Amazon
[90] Foreign Intelligence Surveillance Act
[91] Neil McAllister, *Senate Votes to Continue FISA Domestic Spying Through 2017*, 29/12/2012. In The Register. En ligne http://www.theregister.co.uk/2012/12/29/senate_fisa_extension_vote/
[92] The Washington Post, *NSA Slides Explain the PRISM Data-Collection Program*, 6/06/2013, http://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents/
[93] The Washington Post, *NSA Slides Explain the PRISM Data-Collection Program*, 6/06/2013, http://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents/
[94] James Ball & Dominic Rushe, *NSA Prism program taps in to user data of Apple, Google and others, 6/06/2013.* En ligne http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data

soit lancé publiquement, une solution pour permettre l'accès via PRISM au service de stockage en ligne SkyDrive.[95] Les géants dénient.

6. Conclusions et perspectives

La croissance exponentielle du volume des données disponibles par, entre autres, la multiplication des objets connectés peut très vite noyer une administration comme un service de renseignement qui y trouve un intérêt, si une méthode de traitement et d'analyse efficaces n'est pas appliquée. Nous avons vu que certaines agences s'y attèlent par le recrutement d'experts en gestion de l'information. L'infobésité peut atteindre le niveau de performance des services dès lors qu'il leur serait impossible d'extraire des données massives les valeurs immédiates et cachées qu'elles révèlent.

Les collecteurs de données et les analystes doivent composer avec cette situation. Appréhender autrement les informations, en protégeant le cœur de métier 'identificateur' du renseignement, mais en acceptant un changement structurel profond. Cette nouvelle donne amène en effet avec elle de nouvelles approches de l'information et du renseignement qui provoqueront nécessairement des bouleversements dans toute la structure organisationnelle. Les services de renseignement doivent saisir la balle au bond sous peine de perdre en efficacité. L'histoire montre que le renseignement s'est toujours adapté aux situations nouvelles que cela soit pour la disponibilité de l'information et des vecteurs nouveaux qui la véhiculent. Anticiper, prévoir, émettre des hypothèses et planifier, cœur du métier du renseignement sont des actions qui peuvent sortir renforcées par ce que les données massives efficacement analysées peuvent promettre. Ce sont aussi des objectifs qu'une organisation inéluctablement en transition doit se fixer.

Quand on parle *big data*, on parle donc valeur et potentiel. Valeur immédiate, définie, et valeur à plus long terme, éventuellement non encore définie. Avec un potentiel insoupçonnable. Nul ne sait ce que l'on pourra en tirer demain. La valeur réside aussi dans le chef de ceux qui pourront extraire de ces *big data* l'information, le renseignement utile. Celui qui mettra en avant une corrélation, celui qui permettra une prédiction. Dans ce nouveau paradigme amorcé, l'analyste garde toute son importance. L'humain y a toute sa place. Aidé

[95] The Guardian, *How Microsoft handed the NSA access to encrypted messages,* 11/07/2013. En ligne https://www.theguardian.com/world/2013/jul/11/microsoft-nsa-collaboration-user-data

d'algorithmes, certes, mais doté d'une intuition, d'une inventivité, d'une créativité que la machine n'égale pas et qui lui permet de mettre du sens là où les tableaux croisés n'en voient pas. L'humain esquisse des hypothèses, des relations causales mais n'attend pas leur confirmation pour avancer. Il suppose et garde la main.
L'intuition, la créativité, l'innovation sont des processus à prendre en compte face à des quantités massives de données dont il faut extraire de la valeur. Les algorithmes peuvent mettre sur la voie de corrélations mais l'analyste doit se montrer intuitif dans son interprétation, audacieux dans ses hypothèses. La plus-value humaine est toujours d'actualité. Le *big data* ne l'a pas détrôné.

Finalement, pour le renseignement, « le jeu en vaut-il la chandelle ? » Sans aucun doute oui, un service de renseignement ne peut se permettre de 'louper' une source d'informations qui vient s'adjoindre aux autres sources traditionnelles, pour autant que : (1) l'analyse de l'adéquation entre les nouvelles ressources engagées (ce que coûte l'adaptation structurelle d'un service en terme de ressources humaines, financières, techniques …) et les *outputs* soit positive ; on parle alors d'*efficience productive*; (2) que l'adéquation entre les objectifs que le service de renseignement s'est fixé en matière de gestion des *big data* et les *outcomes* réels soit positivement évaluée. On parle d'*efficacité ;* (3) que les *outputs* et l'impact sur les groupes-cibles (les clients des services) en terme de degré d'utilisation soient en adéquation.
Mais pour un service de renseignement de taille moyenne, un débat politique mériterait d'être entamé, d'autant que ces services seraient tentés, par manque de ressources humaines, techniques et financières d'externaliser (outsourcing) les capacités de collectes et de filtrage des *big data.* Dans ce cas, des problèmes d'indiscrétion, par rapport aux opérations menées par le renseignement, des problèmes de cadre légal, de règles de déontologie et de comportements éthiques devraient être pris en compte[96].

[96] Lucas Wagner, *Une coopération entre le secteur public et le secteur privé (sous la forme d'une externalisation) pour effectuer du renseignement de sources ouvertes (OSINT) est-elle envisageable ? Analyse des limites et opportunités.* Mémoire Master en science politique, université de Liège, année académique 2015 - 2016

BIBLIOGRAPHIE

Livres

Babak Akhgar, & al., *Application of Big data for National Security*, Oxford (UK), 2015,

Franck Bulinge, *Maîtriser l'information stratégique,* Editions Deboeck, 2014

Pierre Delort, *Le Big data*, Paris, 2016

Olivier Forcade & Sébastien Laurent, *Secrets d'Etat – pouvoirs et renseignement dans le monde contemporain,* Editions Armand Colin, 2005

Mark Lowenthal, *Intelligence. From secrets to policy,* CQ press, 2009.

Viktor Mayer-Schönberger & Kenneth Cukier, *Big data*, *A Revolution That Will Transform How We Live, Work, and Think,* Paris, 2014

David Omand, *Securing the State,* Editions C. Hurst & Co. Ltd, 2011

Harry Ransom, *Central Intelligence and national security*, Cambridge, Harvard University Press, 1958

Le guide du routard de l'intelligence économique, 2012, Paris.

Le livre blanc, défense et sécurité nationale, Direction de l'information légale et administrative, Paris, 2013

Articles scientifiques

Franck Bulinge, « Le cycle du renseignement, analyse critique d'un modèle empirique », in *Market Management,* 2006/3, vol 6

Thomas Cavaillé-Fol, « De nouveaux algorithmes pour scruter le cyberespace », in, *Sciences et Vie*, octobre 2016.

Erhard Friedberg, *« L'analyse sociologique des organisations ».* In *Pour,* N°28, 1972, (fruit des travaux de Michel Crozier, et du Centre de Sociologie des Organisations)

Arthur Hulnick, "What's wrong in the intelligence cycle", in *Intelligence and national security,* Vol 1, N°6, Decembre 2006

Peter Knoepfel, Frédéric Varone, « Mesurer la performance publique: méfions-nous des terribles simplificateurs », in *Politiques et management public*, Vol 17, n°2, 1999

William Lahneman, *"The need for a new intelligence paradigm." In International Journal of intelligence and Counterintelligence*, Vol 23, N°2, Summer 2010

Patrick Leroy, « La communauté du renseignement belge, essai de définition », in, *Revue Militaire Belge,* n°12, Institut Royal Supérieur de Défense, 2016

Yvon Pesqueux, « La notion de performance globale », *2003,* in Thierry Le Nedic, *La performance dans le secteur public, outils, acteurs et stratégie*, Ecole des Mines de Paris, 2009

Guy Rapaille, Dirk Peeters & Patrick Leroy, « Ethique et analyse : la relation avec le décideur », in, *Renseignement et éthique : le moindre mal nécessaire,* Groupe Européen de Recherche en Ethique et Renseignement, 2014

Anselm Strauss, « La dynamique des professions », in *La trame de la négociation - Sociologie et interactionnisme,* Editions Lharmattan, 1992,

Ouvrages académiques

Lucas Wagner, *Une coopération entre le secteur public et le secteur privé (sous la forme d'une externalisation) pour effectuer du renseignement de sources ouvertes (OSINT) est-elle envisageable ? Analyse des limites et opportunités.* Mémoire Master en science politique, université de Liège, année académique 2015 – 2016

Guillaume Gustav De Valk, *Dutch Intelligence – Towards a qualitative framework for Analysis,* Rijksuniversiteit Groningen, 2005.

Sabine de Bethune, Présidente du Sénat, Préface de l'ouvrage *Regards sur le contrôle : vingt ans de contrôle démocratique sur les services de renseignement*, Wouter Van Laethem & Johan Vanderborght, Ed., Editions Interscientia, 2013

Sources électroniques

Sandrine Mathen, *Sachez-le, vos photos vous trahissent !*, 27/06/2016.
En ligne http://www.reputation365.eu/analyses-decryptages/vos-photos-vous-trahissent-exif/

Neil McAllister, *Senate Votes to Continue FISA Domestic Spying Through 2017*, 29/12/2012. In The Register.
En ligne http://www.theregister.co.uk/2012/12/29/senate_fisa_extension_vote/

The Washington Post, *NSA Slides Explain the PRISM Data-Collection Program*, 6/06/2013, http://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents/

James Ball & Dominic Rushe, *NSA Prism program taps in to user data of Apple, Google and others, 6/06/2013.*
En ligne http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data

The Guardian, *How Microsoft handed the NSA access to encrypted messages,* 11/07/2013. En ligne https://www.theguardian.com/world/2013/jul/11/microsoft-nsa-collaboration-user-data

Whatsapp, *Le chiffrement de bout en bout*, 5/06/2016.
En ligne https://www.whatsapp.com/faq/fr/general/28030015

Le Monde, *Le FBI assigné en justice pour révéler comment il a pu débloquer l'iPhone de San Bernardino*, 16/09/2016.
En ligne http://www.lemonde.fr/conflit-apple-fbi/article/2016/09/16/le-fbi-assigne-en-justice-pour-reveler-comment-il-a-pu-debloquer-l-iphone-de-san-bernardino_4999085_4870067.html

Test-Achats, *Que deviennent les données d'« automesure »? Test-Achats tire la sonnette d'alarme*, 14/10/2015.
En ligne https://www.test-achats.be/action/espace-presse/communiques-de-presse/2015/privacy

Ronald Meeus Longread, *Comment iPolice peut sécuriser la nation*, 24/06/2016.
En ligne http://datanews.levif.be/ict/actualite/longread-comment-ipolice-peut-securiser-la-nation/article-longread-516395.html?utm_campaign=Echobox&utm_medium=social&utm_source=Facebook

James Clapper, *Worldwide Threat Assessment of the US Intelligence Community*, *Statement for the Record*, februari 9, 2016. In the future, intelligence services might use the IoT for identification, surveillance, monitoring, location tracking, and targeting for recruitment, or to gain access to networks or user credentials. En ligne sur le site du Comité du Sénat US sur les Services armés https://www.armed-services.senate.gov/imo/media/doc/Clapper_02-09-16.pdf

Le Monde, *Le directeur du renseignement américain reconnaît s'intéresser aux objets connectés*, 10/02/2016.
En ligne sur le site web de Le Monde. http://www.lemonde.fr/pixels/article/2016/02/10/le-directeur-du-renseignement-americain-reconnait-s-interesser-aux-objets-connectes_4862587_4408996.html#9bjRf4mTzPGspPWS.99

T2M, *Les objets connectés : attention, prêt, connecté !*.
En ligne le 9/11/2015 http://www.time2marketing.fr/marketing3.0/objets-connectes.html

Le Monde, *INTRUSION 2.0 – Avec Shodan, contrôlez des webcams et imprimez chez les autres*, 10/06/2014.
En ligne le 5/09/2016 http://bigbrowser.blog.lemonde.fr/2014/06/10/avec-shodan-controlez-des-webcams-et-imprimez-chez-les-autres/

Big data ou Small Data, 04/05/2012.
En ligne http://www.analysepredictive.fr/marketing-predictif/enjeux-marketing/big-data-ou-small-data

Dominique Boullier & Audrey Lohard, *Opinion mining et Sentiment analysis*, 30/04/2012.
En ligne http://books.openedition.org/oep/202

Jean-Luc F., *Devenez ingénieur Big data pour le renseignement militaire,* 21/09/2016.
En ligne le 15/10/2016 http://www.defense.gouv.fr/ema/interarmees/la-direction-du-renseignement-militaire/servir-a-la-drm/les-metiers-de-la-drm/ingenieur-big-data/devenez-ingenieur-big-data-pour-le-renseignement-militaire

Association des professionnels de l'information et de la documentation, *Information grise*.
En ligne le 26/09/2016, http://www.adbs.fr/information-grise-17424.htm?RH=ACCUEIL

Philippe Vion-Dury, *Loi de programmation militaire : le scandale qui fait sploutch*, 19/12/2013.
En ligne http://rue89.nouvelobs.com/2013/12/19/loi-programmation-militaire-scandale-fait-sploutch-248467 ;

Frédéric Bergé, *Loi de programmation militaire : Jacques Attali juge "ahurissant" l'article 20*, 23/12/2013.
En ligne http://www.01net.com/actualites/loi-de-programmation-militaire-jacques-attali-juge-ahurissant-larticle-20-610836.html ;

La rédaction, *La loi de programmation militaire 2014 à 2019 est déjà promulguée*, 19/12/2013.
En ligne http://www.zdnet.fr/actualites/la-loi-de-programmation-militaire-2014-a-2019-est-deja-promulguee-39796465.htm

[1] *The NSA's "Boundless Informant" Program* » [archive], sur http://www.spiegel.de/ [archive], Der Spiegel, 29/06/2013 & Laura Poitras, Marcel Rosenbach, Fidelius Schmid, Holger Stark & Jonathan Stock, *Cover Story: How the NSA Targets Germany and Europe*, 1/07/2013.
En ligne http://www.spiegel.de/international/world/secret-documents-nsa-targeted-germany-and-eu-buildings-a-908609.html

Sébastian Seibt, *Le renseignement américain puise dans les données de Facebook et Google*, 8/06/2013.

En ligne http://www.france24.com/fr/20130607-prism-espionnage-nsa-cybersurveillance-donnees-facebook-google-yahoo-microsoft-scandale-verizon

Le Monde, *Cybersurveillance : la NSA sait déjouer le chiffrement des communications*, 5/09/2013.
En ligne http://www.lemonde.fr/technologies/article/2013/09/05/cybersurveillance-la-nsa-a-contourne-les-garde-fous-qui-protegent-les-donnees_3472159_651865.html

Anouch Seydtaghia, *Inviolable, le chiffrement de WhatsApp irrite le FBI,* 6/04/2016.
En ligne https://www.letemps.ch/economie/2016/04/06/inviolable-chiffrement-whatsapp-irrite-fbi

Khalid Benabdeslem, Christophe Biernacki & Mustapha Lebbah, *Les trois défis du Big data - Éléments de réflexion*, juin 2015.
En ligne sur le site web de la Société française de statistique http://statistique-et-societe.fr/ojs/index.php/stat_soc/article/view/445/419

https://www.cia.gov/news-information/featured-story-archive/2012-featured-story-archive/big-data-at-the-cia.html, consulté le 18 novembre 2016

https://www.cia.gov/news-information/featured-story-archive/2012-featured-story-archive/big-data-at-the-cia.html, consulté le 18 novembre 2016

Jean-Laurent Cassely, *Big data: ce n'est pas la taille qui compte*, 13/05/2013.
En ligne le 7/09/2016 sur http://www.slate.fr/economie/72361/big-data-small-data

http://www.smartdatacollective.com/xanderscho/364797/how-cia-reinventing-case-big-data, consulté le 18 novembre 2016

François Assémat, *Traders et terroristes : les risques extrêmes en 2016,* 6/01/2016. In La Tribune.
En ligne http://www.latribune.fr/opinions/tribunes/traders-et-terroristes-les-risques-extremes-en-2016-540285.html

Vinny Eastwood, *How To Predict The Future With Linguistics And Technology, Cliff High.*
En ligne le 13/04/2015 https://youtu.be/VLnon-t9fVQ?t=1m42s