

Technologies d'exploitation du *big data* dans les organisations et transformations organisationnelles : une étude de cas au sein du Service de santé des armées françaises

Operating Technologies of Big Data in Organisations and Organisational Change : A Case Study of the Service de santé des armées française

Marc Tanti, Ph. D.

Volume 63, numéro 4, octobre–décembre 2017

Les données et les sciences de l'information

URI : <https://id.erudit.org/iderudit/1042310ar>

DOI : <https://doi.org/10.7202/1042310ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Association pour l'avancement des sciences et des techniques de la documentation (ASTED)

ISSN

0315-2340 (imprimé)

2291-8949 (numérique)

[Découvrir la revue](#)

Citer cet article

Tanti, M. (2017). Technologies d'exploitation du *big data* dans les organisations et transformations organisationnelles : une étude de cas au sein du Service de santé des armées françaises. *Documentation et bibliothèques*, 63(4), 46–58. <https://doi.org/10.7202/1042310ar>

Résumé de l'article

L'exploitation des données sanitaires du *big data* via des technologies de collecte, de visualisation et de communication a permis au Service de santé des armées françaises de créer une valeur stratégique. Par exemple, après le tsunami de 2004 en Asie, cette exploitation a permis d'anticiper les épidémies pour les forces envoyées en soutien humanitaire. Plus récemment, dans le contexte de l'épidémie d'Ebola en Afrique, des vies humaines ont pu être préservées. L'appropriation de tels outils a également entraîné des bouleversements dans l'organisation et la culture organisationnelle, notamment un décloisonnement spatio-temporel des activités et un gain de coordination.

TECHNOLOGIES D'EXPLOITATION DU *BIG DATA* DANS LES ORGANISATIONS ET TRANSFORMATIONS ORGANISATIONNELLES : UNE ÉTUDE DE CAS AU SEIN DU SERVICE DE SANTÉ DES ARMÉES FRANÇAISES

Marc Tanti, Ph. D.

Adjoint au chef de service de veille sanitaire du Centre d'épidémiologie et de santé publique des armées (CESPA)
Chef de projets de recherches en sciences de l'information CESPA - UMR 912-SESSTIM-iNSERM-IRD-AMU
Chercheur associé à l'IRSIC-EA-4262
mtanti@gmx.fr

RÉSUMÉ | ABSTRACT

L'exploitation des données sanitaires du *big data* via des technologies de collecte, de visualisation et de communication a permis au Service de santé des armées françaises de créer une valeur stratégique. Par exemple, après le tsunami de 2004 en Asie, cette exploitation a permis d'anticiper les épidémies pour les forces envoyées en soutien humanitaire. Plus récemment, dans le contexte de l'épidémie d'Ebola en Afrique, des vies humaines ont pu être préservées. L'appropriation de tels outils a également entraîné des bouleversements dans l'organisation et la culture organisationnelle, notamment un décloisonnement spatio-temporel des activités et un gain de coordination.

Operating Technologies of Big Data in Organisations and Organisational Change: A Case Study of the Service de santé des armées française

The exploitation of health-related data of big data using collection, visualisation and communication technologies has enabled the *Service de santé des armées française* to produce strategic value. For example, following the 2004 tsunami in Asia, this exploitation enabled the humanitarian resources to better anticipate epidemics. More recently, during the Ebola epidemic in Africa, human lives were saved. The mastery of such tools also brought about important changes within the organisation and its culture, namely the spatiotemporal decompartmentalisation of activities and improved coordination.

Problème de recherche

Ces dernières années, le nombre de données circulant sur Internet a littéralement explosé. Ce phénomène de données massives, ou *big data*, amène les organisations à se positionner sur l'exploitation de tels volumes d'informations et notamment à mettre en place des processus organisationnels pour les aider à maîtriser cette pléthore d'informations à des fins d'intelligence pour la décision et l'action (Cohen 2013).

Ainsi, face à l'ampleur de ce phénomène, lié notamment à la prolifération des données des réseaux sociaux, téléphones intelligents, applications mobiles, etc., leur croissance exponentielle et les problèmes de repérage, d'accès, de gestion et de traitement de ces données, un certain nombre d'organisations, entre autres des entreprises du domaine privé, se sont engagées dans des travaux d'exploitation de ces données massives (Karoui 2014).

Par exemple, dans le domaine de la finance, notamment des banques et des assurances, les projets d'exploitations des données du *big data* permettent d'aider aux prises de décisions concernant les données à conserver dans le cadre réglementaire (Karoui 2014). Dans le secteur du *trading*, des projets d'exploitation des données massives ont été conçus pour identifier les mouvements de fonds suspects. C'est le cas par exemple pour l'entreprise BourseDirect (Karoui 2014). L'exploitation de ces données peut permettre également de créer de nouveaux services pour les clients. Dans ce cadre, Le Crédit Mutuel Arkéa a été primé en 2013 pour l'exploitation de ses données d'historiques d'opérations clients pour la création de nouveaux services aux usagers (Karoui 2014).

Le domaine médical est également un domaine où les données sont devenues pléthoriques sur Internet (Raghupathi 2014). Le secteur sanitaire est particulièrement touché, surtout en raison de l'émergence croissante des données

épidémiques et des données issues du séquençage des génomes. Dans ce secteur, les données accumulées sont chiffrées, analysées statistiquement (Alles 2014), mais elles sont aussi non chiffrées, non quantifiables, non structurées, de type socioculturel, par exemple.

Les organisations de santé publique portent une attention particulière à ces différentes formes de données, entre autres pour orienter les politiques décisionnelles. Ces organisations recherchent aujourd'hui des méthodes et des stratégies afin de les « faire parler » (Hamel 2013), notamment pour apporter une valeur ajoutée en matière de performance organisationnelle, de prise de décision, de stratégies (Brasseur 2013 ; Vayre 2014) ou d'amélioration de l'action publique (Hoppe 2015).

Le présent article a pour objectif de présenter le dispositif de veille sanitaire développé par l'organisation militaire française, considéré comme un dispositif infocommunicationnel. Selon Couzinet (2011), un dispositif infocommunicationnel est un lieu où humains, objets matériels et liens s'organisent pour mettre en œuvre des interactions. Un tel dispositif, dédié à la « mise en commun d'informations et au partage des savoirs », se compose « d'acteurs, de techniques et d'objets matériels en interaction permanente et dans un contexte défini, tous liés entre eux » (Couzinet 2011).

Dans le dispositif étudié, des mégadonnées de l'Internet sanitaire sont exploitées dans le domaine de la santé publique militaire (Boutin 2004). Dans ce cadre, notre article présente et interroge les technologies mises en œuvre pour collecter, analyser et communiquer ces données massives.

À partir de cette recherche-action, les enjeux organisationnels soulevés par l'exploitation de tels volumes d'informations sont également contestés, notamment les transformations organisationnelles induites (Parks 2014 ; Béranger 2016). L'article a également pour objectif de réfléchir aux apports en matière de construction des connaissances et de création de valeurs.

Cadre théorique

Big data

Définition

Selon Babinet (2015), le *big data* (ou mégadonnées ou données massives) désigne l'ensemble des données numériques produites par l'utilisation des nouvelles technologies à des fins personnelles ou professionnelles. Cela recoupe les données d'entreprise (courriels, documents, bases de données, historiques de processeurs métiers...) aussi bien que des données issues de capteurs, des contenus publiés sur le

Web (images, vidéos, sons, textes), des transactions de commerce électronique, des échanges sur les réseaux sociaux, des données transmises par les objets connectés (étiquettes électroniques, compteurs intelligents, téléphones intelligents...), des données géolocalisées.

Le *big data* correspond à une évolution de la *business intelligence* qui repose sur des entrepôts de données limités en taille (quelques téraoctets) et gérant difficilement des données non structurées et des analyses en temps réel. L'avènement du *big data* ouvre une nouvelle ère technologique qui offre des architectures et des infrastructures évoluées qui permettent en particulier des analyses sophistiquées, prenant en compte ces nouvelles données intégrées à l'écosystème de l'entreprise (Chen 2014).

Le *big data* peut se définir par trois composantes (Delort 2015) :

- Son volume : on le définit généralement à partir de cinq téraoctets de données à traiter, ce que les logiciels ou les bases de données classiques ne peuvent faire ;
- Sa variété : les données acquises sont brutes ou structurées, au format texte ou image, avec des propriétaires et des droits d'utilisation aussi différents que leurs sources ;
- Sa vélocité : il faut être capable d'intégrer en temps réel les dernières données disponibles et les lier aux autres jeux de données sans recommencer une analyse complète à chaque cycle.

Un quatrième « V » pour véricité peut être rajouté à cette définition pour évoquer la nécessité de vérifier la crédibilité des sources et la qualité des contenus afin de permettre une exploitation des données (Marr 2015).

Ces différentes composantes du *big data* nécessitent donc de nouvelles formes d'outils et de traitement de l'information (Pouyllau 2013).

Technologies

Selon Cointot, le *big data* recouvre l'ensemble des technologies, métiers, approches conceptuelles permettant d'exploiter l'ensemble des données générées par les hommes de façon consciente ou non et par tous les objets connectés ou non (Cointot 2014).

Les technologies qui permettent l'exploitation de ces données au volume important présentent quelques points particuliers, notamment pour permettre la collecte et la « fouille » dans ces données de grand volume. Ainsi, pour arriver à retrouver une épingle dans une meule de foin multidimensionnelle, sans savoir à quoi ressemble l'épingle ni si la meule de foin en contient une, des outils statistiques, de

L'avènement du big data ouvre une nouvelle ère technologique qui offre des architectures et des infrastructures évoluées qui permettent en particulier des analyses sophistiquées, prenant en compte ces nouvelles données intégrées à l'écosystème de l'entreprise.

classification et de visualisation des données sont nécessaires (Weinstein 2013).

Dans ce cadre, un certain nombre de solutions de visualisation de données du *big data* et plus largement de *data-scapes* du Web (paysages de données) sont aujourd'hui utilisables, notamment par des non-experts. Les topologies du Web peuvent ainsi être générées par des logiciels de calculs de graphes à partir de moissonnage de données et de liens utilisables dans l'exploitation des données massives (Boullier 2016). C'est par exemple le cas de Gephi¹, dont nous reparlerons plus tard dans cet article.

Des capacités matérielles conséquentes pour le stockage comme pour les ressources processeurs sont également indispensables au traitement de ces données massives. Dans ce contexte, Le «*cloud*», le nuage, symbolise la multitude d'unités informatiques et accessibles depuis l'Internet pour le traitement et le stockage des données du *big data* (Delort 2015).

Enfin, dans les technologies, les Environnements Numériques de Travail (ENT) ont aussi leur importance. Ainsi, lorsque les usagers se connectent, quel que soit le lieu, ils peuvent retrouver immédiatement leur poste de travail, dossiers, documents, courriels... De la même façon, on peut préciser que l'infonuagique est aussi centrée sur la tâche à accomplir. L'utilisateur ne se préoccupe ainsi plus des types d'applications à utiliser, mais bien de la tâche à réaliser (Delort 2015).

En conclusion, la technologie occupe une place centrale dans le développement et la transformation des organisations. Les relations entre les deux (technologie/organisation) ont d'ailleurs fait l'objet de nombreux travaux (Markus 1988; Orlikowski 1992; Orlikowski 2000; Kefi 2004). Cette approche qualifiée d'ailleurs de «*déterminisme technologique*» repose sur de nombreuses disciplines telles que la sociologie, l'économie et la gestion. L'hétérogénéité de ces approches permet, selon MacKenzie (1999), d'extraire plusieurs caractéristiques. Notamment, la technologie est considérée comme un artefact doté d'un ensemble de caractéristiques matérielles dont le développement est structuré par des lois scientifiques extérieures à la sphère organisationnelle, mais dont les effets sont appréhendés en termes d'impacts sur les usages et l'organisation (MacKenzie 1999).

Transformations organisationnelles

Le paradigme du *big data* n'est donc pas que technologique. Il a ainsi imposé à l'entreprise des transformations organisationnelles pour canaliser ce flux incessant de données et pour en tirer l'information nécessaire à la prise de décision (Cointot 2014; Delort 2015).

Selon Ferguson, le vrai problème est la conduite du changement dans la mesure où le vrai défi pour les organisations est d'arriver à faire les changements nécessaires au niveau des processus et des ressources humaines pour pouvoir mettre en place les initiatives analytiques que requiert un projet d'exploitation des données du *big data* (Ferguson 2013).

Selon une étude du CIGREF : «*Le big data est une composante de la transformation de l'entreprise, il est transversal et touche tous les métiers. En ce sens, le big data n'est pas un projet SI : c'est une manière nouvelle de penser et d'appréhender l'information. Il s'agit donc davantage d'une (r)évolution culturelle et technologique que d'un nouveau sujet SI.*» (CIGREF 2013)

Le phénomène du *big data* pour les entreprises recouvre ainsi deux réalités : d'une part cette explosion des données de façon continue, d'autre part la capacité technologique de traiter et d'analyser cette grande masse de données pour en tirer un profit (CIGREF 2013). Grâce au *big data*, les entreprises peuvent désormais gérer et traiter des données massives pour en extraire de la valeur, décider et agir en temps réel (CIGREF 2013). L'enjeu central pour elles est bien «*d'améliorer l'efficacité des prises de décision par l'exploitation d'informations protéiformes*» (CIGREF 2013). Elles deviennent ainsi plus réactives et plus concurrentielles. De fait, la capacité à traiter de grandes masses de données, à tisser des liens et corrélations entre des informations hétérogènes et à faire communiquer entre elles des données structurées ou non, ouvre la voie à des traitements probabilistes qui permettent l'amélioration des opérations managériales et la conception de produits et services innovants pour les clients (CIGREF 2013). Aussi, aujourd'hui, le *big data* est corrélé à l'avantage concurrentiel des entreprises et représente pour elles un atout considérable (CIGREF 2013).

Cette appropriation d'outils, de technologies devenues incontournables pour conserver un avantage concurrentiel impose ainsi à l'entreprise de nouvelles exigences en matière de compétences, d'autonomie et de capacité d'adaptation des travailleurs dont l'efficacité est censée croître (Benraiss 2005). Cette appropriation d'outils agit sur l'organisation du travail, impose de nouvelles méthodes, de nouvelles pratiques et affecte les attitudes et les comportements (Benraiss 2005).

Ces bouleversements liés à ces nouveaux usages touchent la structure même de l'entreprise qui se modifie, «*s'horizontalise*» (Lejeune 2010). La hiérarchie devient ainsi plus informelle. La communication se fait plus librement, tant sur le plan horizontal que vertical. L'information circule plus rapidement et plus efficacement. Le dialogue social se modifie. L'organisation économique subit également des modifications. Le changement ne se situe donc plus uniquement sur le plan professionnel ou organisationnel. Il se manifeste dans de nouveaux rapports sociaux et

1. <gephi.org/>.

économiques qui découlent de l'appropriation sociale des technologies (Lejeune 2010). La culture du travail et la relation au travail ne sont pas non plus épargnées et subissent des bouleversements profonds (Lejeune 2010).

Silva a réalisé une étude sur les nouvelles pratiques sociales liées à ces technologies dans les grandes entreprises (Silva 2009). Il a notamment analysé les changements de valeurs à l'origine de transformations sociales (Silva 2009). Selon cet auteur :

[...] les TIC favorisent les fragmentations des temporalités, l'éclatement des espaces et les différenciations de personnes [...]. Les rythmes et lieux de travail ne sont plus partagés par l'ensemble de l'entreprise. Il se développe, de fait, des territorialisations, c'est-à-dire des espaces temps propres à telle équipe ou à tel projet. La cohésion de l'entreprise peut être contestée [...]. Les TIC ont catalysé différents facteurs d'éclatement des groupes de salariés. En effet, elles favorisent la déconstruction des chaînes de responsabilités, dans la mesure où les messageries, par exemple, peuvent permettre de contourner certains processus hiérarchiques ou techniques, pour favoriser des liens informels qui s'étoffent [...]. L'informatisation va accompagner et augmenter cet effet de fragmentation du travail et des relations sociales dans l'entreprise. (Silva 2009)

Les TIC modifient de multiples fonctions de l'entreprise : informations décisionnelles, informations de communication, de création et de données sociales (Silva 2009 ; Silva 2010). C'est ainsi qu'émerge une nouvelle forme d'organisation du travail, centrée notamment sur les activités collaboratives qui constituent une rupture culturelle importante, notamment lorsqu'elles impliquent des personnes situées dans des lieux, des situations hiérarchiques et des temporalités différentes (Silva 2009 ; Silva 2010). Cette forme d'organisation est non plus seulement fondée sur l'individu, mais sur la capacité qu'il a de travailler avec d'autres, base de la valeur et de l'innovation pour les entreprises (Silva 2009 ; Silva 2010).

La multiplication des flux de données impacte donc directement la performance individuelle des personnes.

Construction de connaissances

Dans leur ouvrage *big data*, Mayer-Schoenberger & Cukier introduisent leur livre de cette manière :

Le phénomène des mégadonnées désigne tout ce qui peut être fait à une large échelle et non à une échelle plus réduite, afin d'extraire de nouvelles connaissances ou de créer de nouvelles formes de valeur, bouleversant ainsi les marchés,

les organismes, les relations entre citoyens et gouvernements, et bien plus. (Mayer-Schoenberger et al. 2014)

Le *big data* n'est pas en soi une discipline dédiée à la prédiction, mais plutôt à l'analyse brute de grande quantité de données (Cointot 2014). Dans bien des cas, ces analyses sont en elles-mêmes suffisantes pour permettre d'en extraire des informations de qualité dont il pourra être fait immédiatement usage (Cointot 2014). Par exemple, en médecine, des recherches de marqueurs de pathogènes dans les analyses médicales d'un grand nombre de patients peuvent permettre de savoir quelle population exposée à un environnement particulier peut être affectée d'une pathologie particulière. Cette construction de savoir peut ainsi constituer une valeur permettant la mise en place d'actions spécifiques auprès de cette population (Cointot 2014).

L'objectif est ainsi de créer, à partir de ces données quelque chose d'utile, de pertinent permettant d'accroître la connaissance (Hyeans 2016).

L'objectif, à partir de l'exploitation de ces données massives, est de créer quelque chose d'utile, de pertinent et permettant d'accroître la connaissance pour la prise de décision et l'action.

Pour Le Coadic (2004) : « La connaissance est le résultat de l'acte de connaître, acte par lequel l'esprit humain saisit un objet. Connaître, c'est être capable de former l'idée de quelque chose, c'est avoir présent à l'esprit. »

Selon cet auteur, « la connaissance peut aller jusqu'à la compréhension exacte et complète des objets [...] dans le but de permettre une adaptation à l'environnement (naturel et humain) » (Le Coadic 2004).

La connaissance possède quelque chose de supérieur. En effet, selon Foray (2009), elle donne à son détenteur une capacité d'action intellectuelle et physique.

Création de valeurs

Outre les quatre composantes décrites précédemment – volume, variété, vélocité, véracité (Delort 2015), une cinquième composante peut également définir le *big data* : la création de valeurs (Marr 2015 ; Babinet 2015). En effet, l'un des aspects notoires des *data*, c'est que l'on ne connaît pas nécessairement a priori le trésor qui se cache en son sein. L'objectif, à partir de l'exploitation de ces données massives, est de créer quelque chose d'utile, de pertinent et permettant d'accroître la connaissance pour la prise de décision et l'action (Babinet 2015).

La création de valeurs est un concept suscitant aujourd'hui un intérêt croissant dans différents domaines des sciences de gestion : management stratégique, finance d'entreprise, comptabilité, contrôle de gestion, organisation, marketing.

Bourguignon distingue trois acceptions de la valeur : la valeur au sens de mesure (en particulier dans les disciplines scientifiques que sont les mathématiques et la physique), la valeur au sens économique et la valeur au sens philosophique (Bourguignon 1998). Le terme de valeur est synonyme de celui de richesse. Le thème de la valeur fait ainsi l'objet de regards multiples ou de paradigmes, c'est-à-dire de visions communes aux membres d'un groupe donné (Kuhn 1983). La problématique de la valeur renvoie donc à la question des destinataires de la valeur créée : pour qui créer de la valeur ? Dans le cadre de la finance d'entreprise, la valeur est souvent une valeur financière pour l'actionnaire.

Créer de la valeur en économie, c'est ainsi la faire varier dans le sens de la hausse. A contrario, détruire de la valeur, c'est la faire baisser au cours du temps. La création de la valeur économique est au cœur de l'activité des organisations et au centre de leur vocation, de leur raison d'être et de leur stratégie (Savall 2008).

La notion de création de valeurs est elle-même polymorphe, car elle génère de nombreux indicateurs qui prennent essentiellement appui sur des bases comptables, boursières ou économiques. De ce fait, dans l'entreprise, la création de la valeur s'apprécie au niveau de l'actionnaire (Elidrissi 2010). Une entreprise crée de la valeur si le résultat dégagé par l'exploitation est supérieur au coût des capitaux investis. Elle s'apprécie aussi au travers les gains réalisés au-delà du coût du capital investi.

Un investissement en systèmes d'information (SI) peut générer trois types de gains pour l'entreprise. Elle peut générer des gains mesurables en unités monétaires. Elle peut générer des gains estimables en unités non monétaires (gain de volume, temps, espace, performance, stratégique, etc.) (Elidrissi 2010). Elle peut aussi générer des gains peu tangibles (autonomie de l'utilisateur de l'information, satisfaction des clients, meilleure collaboration en interne, accumulation de l'expertise technologique, etc.) (Elidrissi 2010).

Dans le cas des organisations publiques, comme dans notre étude, la création de la valeur ne peut a priori s'estimer qu'en unités non monétaires et qu'en gains peu tangibles, notamment qu'en termes de gain stratégique et décisionnel, car elles n'ont pas vocation à faire de profits.

La performance est aussi très liée au concept de création de valeurs afin de permettre la prise de décisions managériales, stratégiques idoines et adéquates d'une part, et de rendre l'organisation concurrentielle. Certains auteurs l'assimilent à l'efficacité, à la capacité ou à la compétitivité, d'autres à l'efficience, au rendement, à la productivité, et d'autres, enfin lui associent des notions telles que la santé, la réussite, le succès et l'excellence (Little 1973 ; Bourguignon 1995 ; Tchankam 2000). Par exemple, Little (1973) l'assimile comme la mise en œuvre et sa réalisation d'une action. Tchankam (2000) définit l'entreprise performante

comme celle qui fait mieux que ses concurrents sur le court, le moyen et le long terme, dans l'idéal d'un ensemble de paramètres définissant la performance, au minimum sur ceux jugés être les plus significatifs. Bourguignon (1995) pense que la définition de la performance est corrélée à l'appréciation du marché : part de marché, chiffre d'affaires, nombre de créances douteuses ou de plaintes de la clientèle dues à la qualité des produits.

La contribution originale des pionniers à cette dimension de la performance réside dans sa relation au concept d'innovation, processus managérial consistant à innover, c'est-à-dire à chercher à améliorer constamment l'existant, un concept souvent au cœur des préoccupations stratégiques des entreprises (Drucker 1957).

Méthode

L'analyse du dispositif s'est d'abord effectuée par une observation non participante au sein du service chargé de l'exploitation des données du *big data* sanitaire dans les armées (Tanti 2010). Dans le cadre de cette approche globale qualitative, l'observation non participante a permis, par le fait que l'observateur n'est pas membre du dispositif et par le fait qu'il reste en retrait durant les observations, de garder une part d'objectivité dans les analyses, constituant ainsi un outil de valeur dans l'analyse de l'objet de recherche.

Cette observation non participante s'est déroulée durant trois mois en 2014. Elle a ainsi permis une meilleure compréhension des acteurs du dispositif, de leurs rôles dans l'organisation, de leurs modes de travail, de leurs modes de communications, ainsi que des objets informationnels et techniques produits et mis en jeu dans l'organisation, notamment des produits documentaires et des systèmes d'information développés.

L'étude s'est poursuivie par des entretiens de ces acteurs rendus possibles par l'analyse des données récoltées antérieurement. Cette enquête, également réalisée en 2014 auprès des acteurs du dispositif (avec un total de six participants interrogés), a consisté en des entretiens semi-directifs, avec grilles. Ces entretiens ont ainsi permis une connaissance plus précise des interlocuteurs, l'expression de leurs opinions, des expériences vécues, une meilleure appréhension des outils de traitements de données, de leurs modes de partage et de production de la connaissance à l'échelle individuelle et collective, des valeurs créées et des modalités de fonctionnement des équipes, des technologies et de leurs usages, de l'organisation et des collaborations.

Enfin, une analyse du dispositif sous le prisme des concepts mobilisés dans le cadre théorique de cet article a été effectuée, notamment l'analyse des aspects construction de connaissances et création de valeur du dispositif.

Résultats

Technologies d'exploitation des données massives

Sur le Web, il existe encore plus de données que d'étoiles dans l'univers. Face à cette croissance exponentielle et afin d'exploiter ces données pour en retirer l'information sanitaire utile à la prise de décision, le dispositif utilise des outils de collecte automatisée et semi-automatisée et des outils d'aide à l'analyse et de classification/visualisation de données.

Outils de collecte automatisée ou semi-automatisée

Le dispositif collecte ainsi de manière automatique les données bibliographiques du *big data* médical à partir du service MyNCBI du logiciel PubMed (Bussi eres 2003). Ce service permet l'interrogation de la base de donn ees bibliographique en sciences de la sant e Medline (*MEDical Literature Analysis and Retrieval System on LINE*) produite par la National Library of Medicine (NLM) aux  .U. Il permet la collecte automatis ee de donn ees massives de cette base   partir d'un profil et d' equations de recherches d efinies. Les donn ees sont extraites selon ce profil et sont adress ees automatiquement par courriel   l'organisation soit imm ediatement d es leur parution, soit selon une p eriodicit e d efinie en fonction des th ematiques (Eveillard 2012). Par exemple, l'organisation suit avec attention, via cet outil, les derni eres donn ees de recherches sur les virus Chikungunya, Ebola et Zika.

Le syst eme collecte  galement de mani ere automatique et en temps r eel les donn ees  pid emiques   partir du programme international de veille  pid emiologique sur les maladies  mergentes PromMED-mail (Cowen 2006). Ce service permet notamment la r eception automatique d'alertes  pid emiques dans la bo ite aux lettres  lectronique. Dans ce cadre, lors des derni eres  pid emies de grippe A (H1N1), cet outil a permis un suivi quotidien des donn ees relatives   l' volution du nombre de cas   l' chelle mondiale.

  partir d'Europresse.com² et de Factiva³ et avec diff erentes  quations de recherche, le dispositif collecte automatiquement les donn ees sanitaires et  pid emiques issues de plus 80 m edias fran ais nationaux et internationaux. Par exemple, il extrait les donn ees des m edias fran ais LeMonde.fr ou Lefigaro.fr et des m edias internationaux comme elwatan.com, theguardian.com ou nytimes.com.   titre d'exemple, le dispositif a pu effectuer un suivi des rumeurs  pid emiques sur ces m edias lors de les derni eres  pid emies d'Ebola (2014-2015) et de Zika (2016). Une extraction des donn ees

informelles et de la rumeur est  galement effectu ee   partir du Web social, notamment de Twitter⁴ et des forums de discussions comme Doctissimo.fr⁵.

Gr ace au moteur de recherche Pickanews⁶ et via le moteur de recherche d'Europresse, l'organisation effectue un suivi quantitatif en rapport avec ses mots-cl es d'int er ets (Chikungunya, Ebola, Zika...) sur les diff erents m edias et sur Twitter.

Le dispositif utilise  galement des outils d'alertes⁷ pour surveiller les donn ees sanitaires en rapport avec les th ematiques d'int er ets sur le Web.

Outils d'aide   l'analyse, de visualisation et de classification des donn ees

Le dispositif fait usage d'un certain nombre d'outils d'aide   l'analyse permettant une visualisation des contenus et la classification de donn ees.

Ainsi, le moteur de recherche Pickanews, au-del a de permettre un suivi quantitatif de l' volution des donn ees d'int er ets pour la veille sur les m edias et Twitter, permet aussi de rassembler ces donn ees, de les classer et de les analyser par type de presse, m edias ou date. Ce moteur permet de surveiller les mots-cl es d efinis et d'acc eder aux articles identifi es.

Le logiciel bluenod⁸ permet une analyse des influenceurs et de leurs r eseaux sur Twitter. Il permet en particulier de g erer directement les listes de ce r eseau social, d'exporter des donn ees et de cibler les communaut es li ees aux sujets. Enfin, cet outil construit des repr esentations cartographiques de l'information et classe les donn ees. Il a permis par exemple de d eterminer quels  taient les protagonistes intervenant dans le champ de l' pid emie d'Ebola 2014-2015 et l' volution de leurs discours. Le logiciel MapDTweetmap⁹ est  galement utilis e dans le dispositif pour g eolocaliser sur Twitter les  changes de *tweets* et d eterminer les meilleurs *hashtags*.

L'outil Gephi¹⁰, dont nous avons parl e plus haut, permet la visualisation et la cartographie de donn ees sur Twitter, et l'analyse de r eseaux, notamment de l'activit e en r eaction   des  v enements. Par exemple, dans le cadre de l' pid emie d'Ebola 2014-2015, il a permis un suivi des r eseaux constitu es en r eaction aux pics  pid emiques et l'importation de cas.

2. <www.europresse.com/fr>.

3. <www.dowjones.com/products/factiva>.

4. <twitter.com>.

5. <www.doctissimo.fr>.

6. <www.pickanews.com>.

7. Google Alertes: <www.google.fr/alerts>.

8. <bluenod.com>.

9. <www.mapd.com/demos/tweetmap/>.

10. <gephi.org/>.

Enfin, l'outil AISummarizer¹¹ est un outil d'analyse linguistique utilisé dans le dispositif pour faire des résumés multilingues concis de pages Web et de documents. Dans ce cadre, il trouve son intérêt dans l'aide à l'analyse de données épidémiologiques de l'OMS ou de Santé publique France.

Outils de diffusion et de communication des données

Les données collectées, classées et synthétisées sont finalement mises en forme et diffusées en temps réel ou quasi réel sur une plateforme intranet sécurisée, dans un espace de travail collaboratif avec un Environnement Numérique de Travail (ENT) dédié au dispositif (Boutin 2004; Tanti 2010).

Il faut noter les capacités conséquentes de stockage sur des serveurs sécurisés pour le traitement de ces données massives, ainsi que des ressources processeurs élevées.

Les données captées sont diffusées dans cette plateforme principalement sous forme hypertextuelle, permettant de naviguer d'une unité d'information à une autre. La navigation peut aussi se faire via un moteur de recherche (Tanti 2010). Il existe différents profils d'utilisateurs, principalement des décideurs du Service de santé des armées (SSA)

et de l'État-major des armées (EMA), des experts militaires en santé et des médecins d'unités. Trois systèmes d'information (SI) hébergés sur cette plateforme ont ainsi été développés pour répondre aux besoins informationnels de ces différents usagers.

Le premier SI a été dénommé BEDOUIN, pour Banque Épidémiologique de Données sur l'OUtre-mer et la zone INter-tropicale (Figure 1). Il diffuse les données filtrées et analysées sur les risques sanitaires pour les forces hors de France issues du *big data*. Le contenu est architecturé par pays.

Le deuxième SI s'appelle REDUVES pour base de données en REcherches et Développements de l'Unité de VEille Sanitaire (Figure 2). Il diffuse les données scientifiques du *big data* sur les agents du risque biologique (en matière de nouveaux traitements, diagnostics...). Dans ce cadre, les contenus captés et filtrés sont architecturés par risques et par agents.

Le dernier SI s'appelle INTRACESPA, pour Intranet du CESP(A) (Figure 3). Il diffuse différentes données internes au SSA, principalement relatives aux investigations d'épidémies et aux programmes militaires de santé publique. Les données sont consultables par auteur de publication, thème, en mode multicritère, en recherche simple/avancée et par navigation arborescente.

11. <aisummarizer.com/summarizer/index.jsp?ui.lang=fr>.

FIGURE 1

Page d'accueil du SI BEDOUIN

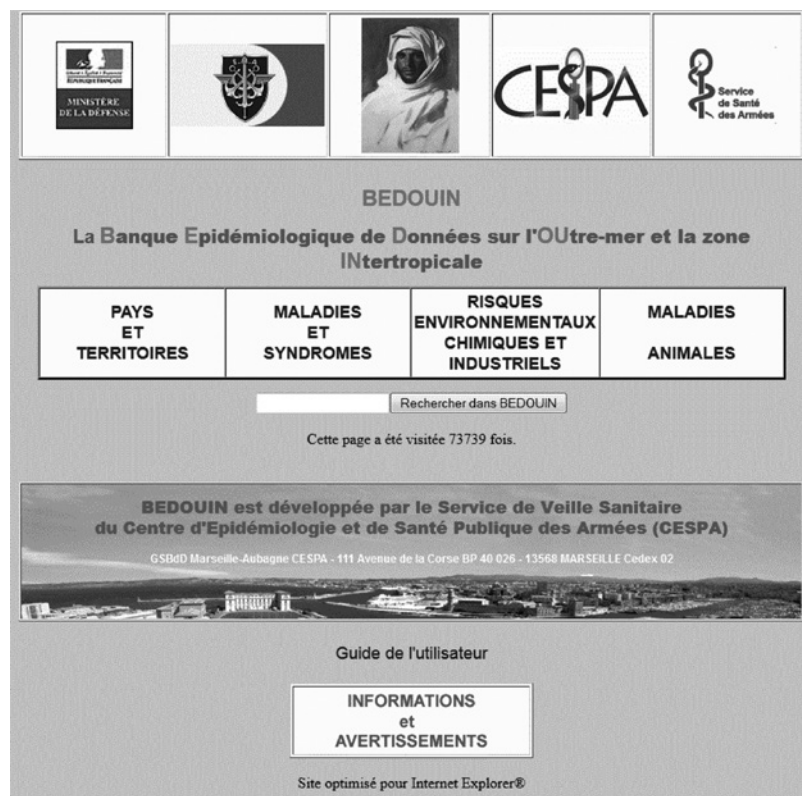


FIGURE 2

Page d'accueil du SI REDUVES

MINISTÈRE DE LA DÉFENSE

Service de Santé des Armées

CESPA

ACTUALITES EBOLA

ACTUALITES ZIKA

REDUVES
La base de données en
Recherches Et Développement de l'Unité de Veille Sanitaire

BULLETINS DE VEILLE

FICHES DE SYNTHESE AGENT MALADIE

FICHES DE SYNTHESE R ET D

DOCUMENTS SOURCE

OUVRAGES ET REVUES

Rechercher dans REDUVES

Cette page a été visitée 2727 fois.

FIGURE 3

Page d'accueil du SI intraCespa

INTR@-DESP : le portail des DESP nord et sud

Annuaire
IMTSSA
DESP Sud
DESP Nord
DCSSA
EVDG
CRSSA
EMA
SSA
Maia

Organigrammes
IMTSSA
EVDG

Divers & Pratiques
Outils survépid
Le Vidal
Bilan Social 2007
Lexibase
Dico Fra/Ang.

**CONSULTATION DES DOCUMENTS
EPIDEMIOLOGIQUES ET SANTE PUBLIQUE**

Consultation de tous les documents

Consultation par auteur Consultation par thème

Consultation par type de documents

Sélections multicritères (en cours d'élaboration)

paludisme

Rechercher dans INTR@DESP

Exploration des documents
par navigation arborescente

Dans cette plateforme intranet, outre les SI diffusant et stockant les données, les usagers peuvent communiquer entre eux, partager les informations, les commenter, se les réapproprier via des outils de travail collaboratif et un Environnement Numérique de Travail (ENT). La communication dans cet ENT se fait de manière asynchrone par messagerie électronique Outlook, qui tend à devenir un moyen de communication instantanée. À cette messagerie, un annuaire des experts, un agenda électronique et un gestionnaire de communautés de « projets » sont associés pour permettre une réappropriation et un partage collectif des données et la création d'intelligence, de nouvelles connaissances et de la valeur.

Transformations organisationnelles

Les entretiens avec les acteurs du dispositif ont révélé que les technologies de collecte, d'aide à l'analyse et de diffusion/communication des données utilisées dans l'exploitation du *big data* sanitaire ont induit des transformations dans l'organisation étudiée.

Transformations induites des outils de collecte et d'analyse

Les outils de collecte automatisée, de classification de données et d'aide à l'analyse ont permis à l'organisation un gain considérable de temps dans la gestion et la réalisation des tâches pour atteindre les objectifs définis. Les outils se sont ainsi mis au service des humains.

Un questionnement sur l'impact des usages des outils de collecte a été mené, notamment pour déterminer les nouveaux échanges et les nouvelles sociabilités induites. Il en ressort que les abonnements aux listes de diffusion professionnelles (PromMED-mail, MyNCBI) ont généré de nouvelles formes de sociabilités dans la mission de veille sanitaire de défense en créant de nouvelles communautés professionnelles virtuelles au service du partage et de la gestion de la même connaissance (Soulier 2004). De nouvelles activités cognitives ont été envisagées : la création, l'échange coopératif volontaire et la réutilisation des connaissances entre les professionnels du réseau (Lefebvre 2004). La communauté de pratiques qui est née a fait émerger une nouvelle unité de coordination qui a permis un traitement plus efficace des connaissances au profit de la mission de veille (Michaux 2004).

La mutualisation de ces outils et cette réappropriation collective des données ont ainsi permis de renforcer le « *coworking* » et le travail collaboratif. Le travail individuel et donc

collectif est devenu plus efficace. L'acteur de l'exploitation des données à son échelle a ainsi pu percevoir de manière plus fine son utilité à l'échelle collective.

Les outils, par leur mutualisation et leur réappropriation collective, ont également permis l'extraction de données « cachées » menant à une amélioration de la créativité, de l'innovation, de l'activité intellectuelle, de l'intérêt au travail et de ses conditions d'exercice, raison d'être du travail humain. Une transformation organisationnelle en a été induite, notamment en gain d'efficacité, de créativité, d'innovation, de productivité et de compétences individuelles et collectives.

Dans les transformations, il est important de noter que les outils de collecte et d'aide à l'analyse soutiennent les acteurs de l'exploitation et le dispositif face à un flux de mégadonnées de qualité variable, mais de quantité pléthorique. Ils permettent au dispositif, comme nous l'avons dit, un gain de temps dans la collecte, la sélection, la classification, la mise en forme, la visualisation des données. Mais ils ne remplacent pas l'analyse humaine et l'expertise qui demeurent indispensables ! Dans ce dispositif, il existe un réseau d'experts (médecins, scientifiques, vétérinaires...) qui exploite, interprète de manière intellectuelle la masse de données collectées. Malgré l'intérêt de l'ensemble des technologies, seul l'humain a la capacité de recouper les données collectées et de vérifier les erreurs de doublonage. Seul l'humain a la faculté de repérer dans les discours ce qui se fait au-delà de ce qui se dit (Lupton 1992) et d'apporter une expertise sur les échanges concernant les problématiques questionnées. Seul l'humain a le pouvoir d'effectuer une sélection et une validation des contenus, en fonction des événements traités, et par comparaison aux événements

antérieurs et connus. Seul l'humain peut analyser le caractère novateur d'un contenu, son intérêt, sa crédibilité, sa fiabilité, en accord avec une question déterminée au préalable par l'autorité de tutelle ou dans une approche inductive, en cherchant à contester les objets. Enfin, c'est le seul à pouvoir synthétiser et interpréter les données collectées en fonction de son expérience, selon les objectifs définis, à les mettre en forme et à les transformer en véritable information utile pour

l'utilisateur du dispositif (Leray 2008).

Transformations induites des outils de diffusion et de communication

Les outils de diffusion et de communication des données déployés sur les plateformes permettent des échanges de données en temps quasi réel entre les acteurs de

[Les outils de collecte et d'aide à l'analyse] permettent au dispositif [...] un gain de temps dans la collecte, la sélection, la classification, la mise en forme, la visualisation des données. Mais ils ne remplacent pas l'analyse humaine et l'expertise qui demeurent indispensables!

l'exploitation du *big data* et leurs usagers. Cela permet également une réappropriation collective des données entre les différents étages de l'organisation et un décloisonnement spatial, pour créer une intelligence collective. Les données permettent elles-mêmes de soutenir la communication dans l'organisation, car elles sont réappropriées collectivement pour créer de la valeur stratégique.

L'Environnement Numérique de Travail (ENT), de stockage de données sur l'intranet et la messagerie Outlook ont favorisé, dans l'organisation, en tant que technologie cognitive, l'émergence d'une culture numérique. De par leur efficacité technique, elles ont valorisé culturellement la pratique de la messagerie électronique et de l'ENT par l'organisation. Elles ont défini, en tant que dispositif matériel, de nouvelles formes d'expression fondées sur le réseau. À travers cette manière spécifique d'agir et de penser en réseau, elles ont instauré un nouveau rapport au savoir et à la connaissance. Elles ont valorisé une communication rapide moins formelle et ont promu les échanges immédiats, s'apparentant à de la communication synchrone.

La mise en place de la plateforme intranet qui héberge les différents SI a facilité l'accès aux flux d'informations, a accéléré l'acheminement des données, a créé de nouvelles formes de partage de la connaissance et a promu le *knowledge management*. Cette plateforme à guichet unique a modifié les relations sociales, la communication et les échanges entre les différents acteurs de l'organisation. L'autonomie des usagers en a été renforcée et les structures ont été décloisonnées. Des relations nouvelles en réseau entre les acteurs de l'exploitation et les utilisateurs ont été développées et redistribuées avec un effet très structurant sur l'organisation et la coordination des décisions. Le travail est devenu de plus en plus coopératif ou «collaboratif» entre tous les acteurs, entraînant notamment des effets «feedback». Cette ouverture a représenté un facteur important de création de valeur ajoutée à tous les échelons et un phénomène qui a transformé le cadre spatiotemporel de l'activité de veille. Les notions traditionnelles de localisation, de distance et de territoire ont été modifiées. Les différentes équipes et services répartis en des lieux différents du territoire peuvent ainsi consulter les données collectées, classées et synthétisées sans contrainte de distance. Le temps et les coûts des recherches documentaires pour l'ensemble des acteurs de l'organisation en ont été réduits.

En conclusion, ces transformations organisationnelles ont elles-mêmes constitué un enjeu managérial, dans la mesure où elles ont été portées par l'équipe de cadres militaires, dans un contexte global de conduite du changement.

Construction de connaissances

Ces mégadonnées «filtrées», hiérarchisées, «clustérisées», purgées de l'infobésité et mises à disposition des usagers,

sous la forme de cette plateforme intranet dédiée, a permis des corrélations inédites de données et la construction de nouvelles connaissances.

Un certain nombre d'exemples de construction de savoirs ont ainsi été donnés en entretiens.

Par exemple, en 2014-2015, lors de l'épidémie d'Ebola, l'exploitation des données de PromMED-mail a permis le suivi de la dynamique épidémique en temps quasi réel. Cet outil de collecte automatisée a ainsi permis de révéler plus rapidement l'émergence et l'expansion de la maladie que les méthodes de collectes traditionnelles sur le terrain. La diffusion des données filtrées, débarrassées du «bruit informationnel» et classées, sur la plateforme Web, a permis aux usagers décideurs de l'EMA un suivi épidémique en quasi instantané, l'anticipation d'éventuelles menaces pour les forces et la mise en œuvre de contre-mesures médicales rapides. De plus, dans ce cadre, l'exploitation des données de génomique et de protéomique de PubMed sur la maladie d'Ebola a permis pour les usagers du système d'ouvrir la voie vers la construction de nouveaux savoirs médicaux et d'explorer des innovations pharmaceutiques, notamment vaccinales et thérapeutiques. Il a permis par exemple de détecter les essais vaccinaux en cours au profit de la population militaire. Le croisement de ces différentes données a ainsi conduit à des corrélations inédites qui ont permis à l'étage décisionnel de développer des modélisations dans une optique d'aide à la décision. Des connaissances ont ainsi été construites et ont permis d'estimer les paramètres fondamentaux de la transmission de la maladie et d'évaluer a priori l'impact des stratégies de contrôle et d'interventions militaires, en y intégrant les aspects médico-économiques.

Création de valeurs

Dans cette étude, la valeur créée est principalement d'ordre stratégique. C'est-à-dire une valeur qui contribue à diminuer l'incertitude du décideur militaire, du médecin militaire de terrain, dans le choix de ses décisions notamment en situation d'urgence sur le théâtre ou avant déploiement ou encore en cas de crises sanitaires. Cette création de valeur permet d'anticiper sur les décisions militaires et prévenir les risques pour le soldat, notamment les risques sanitaires en opération extérieure. Elle permet aussi une meilleure orientation de l'action publique.

Dans notre étude, la valeur stratégique est une valeur plus vaste que la valeur économique ou financière, puisqu'elle intègre aussi des aspects médicaux, réglementaires et décisionnels. La création de la valeur demeure au centre des préoccupations des organisations militaires. Nous la définissons comme l'augmentation de la valeur stratégique pour le décideur militaire en santé. C'est celle qui va procurer des avantages pour anticiper sur les décisions militaires, pour prévenir les risques pour le soldat, notamment les

risques sanitaires, les risques épidémiques qui ont par définition un impact sur les opérations militaires.

Un certain nombre d'exemples de création de valeur ont été donnés en entretiens.

Par exemple en 2005, des forces maritimes françaises sont envoyées en Asie du Sud-Est en soutien humanitaire après le tsunami. Un certain nombre d'épidémies touche la région, notamment une épidémie d'une maladie peu connue à l'époque, l'encéphalite japonaise, qui fait énormément de décès dans le territoire. Il y a beaucoup d'incertitudes sur cette maladie, notamment son risque de transmission aux militaires déployés. L'ensemble des données collectées, notamment à partir de ProMED-mail et de MyNCBI, ont permis de révéler des corrélations inédites. Des données à forte valeur ajoutée mettant en évidence une létalité de 5 à 40 %, pouvant atteindre les 60 %, ont ainsi pu être détectées. Des informations ont également souligné que dans les formes graves, l'infection pouvait entraîner des séquelles neurologiques sévères à des fréquences non négligeables. Il a également été décelé, dans cette masse d'informations, l'existence d'un vaccin efficace. L'ensemble de ces données visualisées par l'étage décisionnel dans la plateforme dédiée ont été réappropriées collectivement et partagées. Elles ont conduit à une décision stratégique importante : l'achat de doses vaccinales au laboratoire le fabricant et l'administration à titre préventif aux forces militaires dans la zone pour prévenir tout risque pour la santé.

Nous pouvons également reprendre l'exemple précédent de l'épidémie d'Ebola 2014-2015 pour illustrer notre propos.

Dès l'émergence des premiers cas, l'exploitation des données massives collectées, analysées, filtrées et diffusées dans le dispositif a permis aux décideurs militaires et aux experts du Service de santé des armées (SSA) de suivre en temps quasi réel la dynamique épidémique de la maladie. Ces données ont été partagées dans la plateforme par les usagers et ont été réappropriées collectivement. Elles ont permis aux utilisateurs d'avoir une connaissance permanente et fine de la menace liée à cette maladie, en particulier de mortalité élevée, d'absence de traitements et de vaccins.

Cette connaissance a été particulièrement prégnante stratégiquement, lorsque pour faire face à cette épidémie d'ampleur sans précédent, un centre de traitement dédié exclusivement au personnel des soignants d'Ebola (CTS) est créé en Guinée en 2015. En effet, dans ce contexte, des experts militaires français ont été envoyés dans ce centre pour apporter leur soutien. Les données du dispositif ont

été partagées par ces experts avant le départ et pendant toute la durée du déploiement. Par des corrélations inédites et le croisement de données venant notamment du terrain, de nouveaux savoirs ont été construits. Ces savoirs ont particulièrement été réutilisés dans le cadre de communautés et de groupes de travail « Ebola ». Ainsi, les experts ont pu détecter des innovations thérapeutiques et vaccinales qui ont fait l'objet d'un traitement et d'un partage particulier étant donné les impacts en termes de mortalité et l'absence de moyens de prévention. L'identification précoce de ces innovations a ainsi permis de créer de la valeur en anticipant et prévenant les risques pour les personnels militaires déployés dans le centre. Par le partage de ces savoirs, la performance dans la prise de décisions en a été améliorée et a ainsi permis de préserver l'état de santé des militaires projetés dans les zones à risque.

Conclusions

L'exploitation des données sanitaires du *big data* via des technologies de collecte, de visualisation et de communication a permis au SSA la création de valeurs, en particulier stratégiques. Par exemple, cette exploitation a permis, après le tsunami en Asie du Sud-Est en 2004, d'anticiper sur les épidémies pouvant toucher les forces envoyées en soutien humanitaire. Plus récemment encore, dans le cadre de l'épidémie d'Ebola 2014 en Afrique, des vies humaines ont pu être préservées. Même si le dispositif d'exploitation des données sanitaires du *big data* décrit dans cet article a été développé dans une organisation publique qui ne fait pas de profits financiers, il serait intéressant d'évaluer, de manière quantitative et économique, la performance sur les plans coût/efficacité, productivité et rentabilité liée à cette exploitation. Cela nous amènerait également à réfléchir à

L'identification précoce de ces innovations a ainsi permis de créer de la valeur en anticipant et prévenant les risques pour les personnels militaires déployés dans le centre.

ce qu'est la performance dans les organisations publiques, notamment celles du secteur de la santé publique. Enfin, il a été mis en évidence que l'appropriation des technologies a entraîné des bouleversements de la culture organisationnelle du service en matière de découplage spatio-temporel des activités, de partage des connaissances et de coordination du travail. Cependant, les technologies n'ont pas remplacé le traitement humain. On peut même s'interroger si l'outil remplacera un jour l'homme dans l'analyse et l'interprétation des données. Enfin, le *big data* est aujourd'hui une innovation en matière de modèle économique et social. Mais, s'agit-il tout simplement d'une évolution de la performance des outils existants ou bien d'un simple effet de mode ? Des questions qui restent ouvertes et qui méritent d'être débattues.

SOURCES CONSULTÉES

- Alles, Michael & Miklos A Vasarhelyi. 2014. Thick data : Adding context to big data to enhance auditability. *International Journal of Auditing Technology* 2 (2) : 95-108.
- Babinet, Gilles. 2015. *Big data, penser l'homme et le monde autrement*. Paris : Le Passeur Éditeur.
- Benraiss, Laila *et al.* 2005. TIC et performance des salariés : quel rôle pour la responsabilité sociale de l'entreprise ? *Revue internationale sur le travail et la société* 3(2) : 915-933.
- Béranger, Jérôme. 2016. La valeur éthique des big data en santé. *Les Cahiers du numérique* 12 : 109-132.
- Boullier, Dominique. 2016. *Sociologie du numérique*. Paris : Edition Armand Colin.
- Bourguignon, Annick. 1995. La performance, essais de définition. *Revue Française de Comptabilité* 26 : 61-66.
- Bourguignon, Annick. 1998. *Management Accounting and Value Creation : Value Yes but what Value ? Working Paper*. Paris : ESSEC.
- Boutin, Jean-Paul *et al.* 2004. Pour une veille sanitaire de défense. *Medecine et Armées* 32 (4) : 66-372.
- Brasseur, Christophe. 2013. *Enjeux et usages du big data : technologies, méthodes et mise en œuvre*. Éditions Hermès.
- Bussièrès, Jean-François *et al.* 2003. Utilisation pratique et efficace de PubMed. *Québec pharmacie* 50 (1) : 20-22.
- Chen, Min *et al.* 2014. Big data : A survey. *Mobile Networks and Applications* 19 (2) : 171-209.
- Cigref. 2013. *Big data : la vision des entreprises : opportunités et enjeux*. Rapport d'octobre 2013.
- Cohen, Corine. 2013. *Business intelligence : The effectiveness of strategic intelligence and its impact on the performance of organizations*. Hoboken, NJ : Wiley-ISTE.
- Cointot, Jean-Charles & Yves Eychenne. 2014. *La Révolution Big data : les données au cœur de la transformation de l'entreprise*. Paris : Edition Dunod.
- Couzinet, Viviane. 2011. Les dispositifs : question documentaire. In *Approche de l'information-documentation : concepts fondateurs* sous la direction de Cécile Gardiès, Caroline Courbières, Viviane Couzinet *et al.* Toulouse : Cepaduès, 120-121.
- Cowen, Peter *et al.* 2006. Évaluation of ProMED-mail as an electronic early warning system for emerging animal diseases : 1996 to 2004. *Journal of the American Veterinary Medical Association* 229 (7) : 1090-1099.
- Delort, Pierre. 2015. *Le big data*. Paris : Presses Universitaires de France.
- Drucker, Peter. 1957. *La pratique de la direction des entreprises*. Paris : Les Éditions d'Organisation.
- Elidrissi, Djamila & Ali Elidrissi. 2010. Contribution des systèmes d'information à la performance des organisations : le cas des banques. *La Revue des Sciences de Gestion* 241 : 55 à 61.
- Eveillard, Philippe. 2012. PubMed en 10 leçons. *La Revue du Praticien* 62 : 1428.
- Ferguson, Renée. 2013. Big data and Big Change Management : A Path Forward. *MIT Sloan Management Review*.
- Foray, Dominique. 2009. *L'économie de la connaissance*. Paris : La Découverte.
- Hamel, Marie-Pierre & David Marguerit. 2013. Quelles possibilités offertes par l'analyse des big data pour améliorer les téléservices publics ? *Revue française d'administration publique* 146 (2) : 437-448.
- Hoppe, Magnus. 2015. Intelligence as a discipline, not just as a practice. *Journal of Intelligence Studies in Business* 5 (3) : 47-56.
- Hyeans, Andy. 2016. *Sport Data Revolution : l'analyse des données au service de la performance sportive*. Paris : Edition Dunod.
- Karoui, Myriam, Grégoire Devauchelle & Aurélie Duzert. 2014. Big data : mise en perspective et enjeux pour les entreprises. *Ingénierie des systèmes d'information* 19 (3) : 73-92.
- Kefi, Hajer & Michel Kalika. 2004. *Évaluation des systèmes d'information : une perspective organisationnelle*. Paris : Economica.
- Kuhn, Thomas. 1983. *La structure des révolutions scientifiques*. Paris : Flammarion.
- Le Coadic, Yves-François. 2004. *La science de l'information*. Paris : Presses Universitaires de France.
- Lefebvre, Philippe *et al.* 2004. Les théories des Communautés de Pratique à l'épreuve : conditions d'émergence et organisation des communautés. *Systèmes d'Information et management* 9 (1) : 25-48.
- Lejeune, Yannick. 2010. *TIC 2025 les grandes mutations. Comment Internet et les technologies de l'information et de la communication vont dessiner les prochaines années*. Paris : FYP Éditions.
- Leray, Christian & Lise Chartier. 2008. *L'Analyse de contenu : de la théorie à la pratique, la méthode Morin-Chartier*. Québec : Presses de l'Université du Québec.
- Little, William *et al.* 1973. *The Shorter Oxford English Dictionary*. Oxford, England : Oxford University Press.
- Lupton, Deborah. 1992. Discourse analysis : A new methodology for understanding the ideologies of health and illness. *Australian Journal of Public Health* 16 (2) : 145-150.
- MacKenzie, Donald & Jude Wajcman. 1999. *The social Shaping of Technology*. Second edition. Buckingham, UK : Open University Press.
- Markus M-Lynne & Daniel Robey. 1988. Information technology and organizational change : Causal structure in theory and research. *Management Science* 583-598.
- Marr, Bernard. 2015. *Big data : Using SMART Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance*. USA : John Wiley & Sons Ed.
- Mayer-Schönberger, Viktor, Kenneth Cukier *et al.* 2014. *Big data : la révolution des données est en marche*. Paris : Robert Laffont.
- Michaux, Valéry & Frantz Rowe. 2004. Complémentarité entre système d'information informatisé, communauté de pratiques et vigilance dans la haute fiabilité : le cas d'une compagnie d'assistance. *Systèmes d'Information et management* 9 (1) : 49-79.
- Orlikowski, Wanda. 1992. The duality of technology : Rethinking the concept of technology and organizations. *Organization Science* 3 : 397-427.
- Orlikowski, Wanda. 2000. Using technology and constituting structures : A practice lens for studying technology in organizations. *Organization Science* 4 : 404-428.
- Parks, Malcolm R. 2014. Big data in communication research : Its contents and discontents. *Journal of Communication* 64 : 355-360.
- Pouyllau, Stéphane. 2013. Web de données, big data, open data, quels rôles pour les documentalistes. *Documentaliste - Sciences de l'Information* 50 : 32-33.
- Raghupathi, Wullianallur & Viju Raghupathi. 2014. Big data analytics in healthcare : Promise and potential. *Health Information Science and Systems* 2 : 3.

- Savall, Henri *et al.* 2008. *Mastering Hidden Costs and Socio-Economic Performance*. Charlotte, USA: Information Age Publishing.
- Silva, François & Stéphane Hugon. 2009. *Émergence des nouveaux usages liés à l'introduction des nouveaux outils dans les grandes entreprises*. Rapport CIGREF – ORSE.
- Silva, François & Anis Ben Ali. 2010. Émergence du travail collaboratif: nouvelles formes d'organisation du travail. *Management & Avenir* 36 (6): 340-365.
- Soulier, Eddy. 2004. Les communautés de pratique au cœur de l'organisation réelle des entreprises. *Systèmes d'information et management* 9 (1): 3-24.
- Tanti, Marc *et al.* 2010. Computerized support of decision-making in medical intelligence. In *Information systems and economic intelligence* sous la direction de Malek Ghenima, Kamel Smaili & Sahbi Sidhom S. Tunis: IHE Edition, 57-65.
- Tchankam, Jean-Paul. 2000. *Performance comparée des entreprises publiques et privées: une étude empirique dans un PVD*. Cahier de recherche, LAREMA.
- Vayre, Jean-Sébastien. 2014. Manipuler les données. Documenter le marché. Les implications organisationnelles du mouvement big data. *Les Cahiers du Numérique* 10 (1): 95-125.
- Weinstein, Marvin *et al.* 2013. *Analyzing big data with Dynamic Quantum Clustering*. SLAC National Accelerator Laboratory.