

Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication

Présentation du dossier

Production of data, "Production of Society". Big Data and Algorithms in the field of Information and Communication Sciences

Producción de datos, "Producción de la sociedad". Big Data y Algoritmos en el campo de las Ciencias de la Información y de la Comunicación

Article inédit, mis en ligne le 15 novembre 2018.

Vincent Bullich

Vincent Bullich est maître de conférences en sciences de l'information et de la communication et membre du Groupe de Recherche sur les Enjeux de la Communication (GRESEC – Université Grenoble Alpes). Ses travaux concernent principalement l'analyse socio-économique des industries culturelles et communicationnelles ainsi que l'économie politique de la propriété intellectuelle.

Viviane Clavier

Viviane Clavier est maître de conférences HDR au Groupe de Recherche sur les Enjeux de la Communication (Gresec) à l'Université Grenoble Alpes. Ses recherches s'inscrivent dans le champ de l'information spécialisée et professionnelle et portent plus particulièrement sur l'organisation des connaissances dans les dispositifs info-communicationnels.

Le dossier 2018 des *Enjeux de l'information et de la communication* est consacré au couple « Big Data et algorithmes », termes récemment devenus prégnants dans les discours sur « la révolution numérique ». Portés tant par des journalistes, des essayistes, des acteurs politiques ou économiques que des représentants de la société civile, ces discours, tantôt apologétiques, tantôt apocalyptiques, opposent souvent promesses d'un renouveau économique aux risques majeurs pour la démocratie et la sécurité nationale, les libertés individuelles, voire l'avenir de l'espèce humaine que feraient

courir la captation et l'exploitation algorithmiques de volumes massifs de données. Concomitamment à cette mobilisation « triviale », le sujet s'est, sur le plan scientifique, échappé des cénacles des spécialistes en informatique pour s'imposer comme thématique des sciences humaines et sociales.

Au-delà des considérations méthodologiques relatives à une nouvelle heuristique fondée sur l'obtention et le traitement d'une quantité inédite de données (Anderson, 2008 ; Aiden et Michel, 2014) et que nous laisserons ici de côté, ce qui motive l'intérêt des chercheurs en SHS se rapporte aux mutations sociales imputables à ces nouveaux dispositifs « *Big Data* et traitement algorithmique ». En effet, ceux-ci seraient en passe d'investir de multiples sphères professionnelles et domestiques et configureraient un nombre croissant d'activités tant publiques que privées. Illustrant le phénomène, le néologisme « data-ification » s'impose progressivement dans les entreprises et les administrations afin de désigner « le nouvel âge » des systèmes d'information au sein de ces organisations (Delort, 2015). Cet essor de la production et du traitement des données est particulièrement visible au niveau des services de marketing ou des ressources humaines notamment (*Ibid.*), mais également dans le cadre des « stratégies de production de l'information publique » (Bardou-Boisnier et Paillart, 2012) et du « gouvernement par les nombres », selon l'expression d'A. Desrosières (2008), conduits par l'Etat et ses administrations. En cela, le couple « *Big Data* et algorithmes » apparaît symptomatique de l'accentuation rapide et significative de « l'informationnalisation » de la société, c'est-à-dire de la circulation croissante et accélérée des flux d'information, ainsi que leur contribution accrue à l'ensemble des dimensions de la vie sociale (Miège, 2004). La dynamique est déjà ancienne et fut initialement abordée par le prisme exclusif du rapport de l'information à la production : elle apparaît, par exemple, à la fois comme la cause et la conséquence de l'émergence de la « société de l'information » comme projet industriel et projet de société à partir des années 1970-1980 (Mattelart, 2001). Cependant, depuis 2001, on assiste à une explosion sans précédent du volume de données médiatisées, explosion qui serait donc à même de provoquer des mutations majeures, en cours ou à venir, à laquelle l'ensemble des acteurs sociaux doivent se préparer, afin d'en tirer le meilleur parti ou d'organiser la résistance à ce qui pourrait être un « panoptique » au niveau mondial (Kitchin, 2014).

Malgré l'ampleur récente des discours profanes et experts oscillant entre promesses exubérantes et anathèmes radicaux et malgré l'intérêt nouveau des chercheurs en SHS pour la thématique des *Big data* et algorithmes, la production même de ces données massives et des instruments de traitement algorithmique reste encore peu appréhendée comme objet de recherche par les sciences humaines et sociales. C'est donc sur cet aspect que le dossier 2018 des *Enjeux de l'information et de la communication* a souhaité se centrer. Plus précisément, celui-ci a visé à une meilleure connaissance du procès de production des *Big Data* et algorithmes en tant que fondement de dispositifs décisionnels de « production de la société ».

La référence à l'ouvrage d'A. Touraine (1973) n'est pas fortuite : bien que la démarche ne se veuille pas (nécessairement) sociologique, il s'agit bel et bien d'envisager ces dispositifs comme des « instruments de production de la société par elle-même », « instruments » foncièrement « historicisés », c'est-à-dire à la fois reflets d'une époque et moteurs de l'évolution sociétale à cette période (*Ibid.*). Nous faisons donc, en filigrane, l'hypothèse que la compréhension des effets sociaux (*lato sensu*) de ces dispositifs passe immanquablement par la compréhension de la manière dont ces données sont construites, traitées et utilisées. Dans cette perspective, l'analyse du procès concret de production des dispositifs constitue une étape indispensable à la connaissance du type de société qu'ils seraient à même de produire. Or, cette analyse passe par l'étude des instruments et techniques mobilisés, par celle des activités des intervenants qui les conçoivent et les mettent en œuvre, par la prise en compte des discours qui orientent et accompagnent leur ancrage social, ainsi que par l'éclairage des conditions institutionnelles dans lesquelles il se réalise. Il s'agit ainsi de

refléter l'épaisseur des multiples médiations *qui s'exercent sur* ces dispositifs, et *qu'exercent ces* dispositifs dans la configuration du rapport des individus aux mondes sociaux. C'est précisément à cet ensemble de domaines que se rattachent les travaux qui composent ce dossier. Ceux-ci ont été répartis en quatre axes, en fonction des objets de recherche et positionnements scientifiques de leur(s) auteur(s), axes dont la complémentarité est à même de favoriser la compréhension des enjeux liés à la production sociale telle que configurée par la production de données.

Les Big Data au prisme des stratégies industrielles et marchandes

Ce premier axe se rapporte à l'étude des stratégies des acteurs économiques, aux logiques de production industrielle des données et de leur traitement, dès lors envisagés comme ressources, ainsi qu'à l'étude de la construction des marchés sur lesquels ces mêmes ressources sont valorisées. Au cœur de la démarche, se trouve l'identification des positionnements de ces acteurs et l'analyse des rapports qu'ils entretiennent. Il s'agit, en outre, d'aborder les données massives et algorithmes comme des facteurs de production (à l'origine de la désormais fameuse « *Data-Driven Economy* ») au sein de filières dont le fonctionnement reste à éclairer et des ressources valorisables *per se* sur des marchés. Enfin, cet axe s'intéresse aux discours portés par ces acteurs et contribuent fortement à inscrire ce couple *Big Data* et algorithmes dans des « économies de la promesse » au sein desquelles il est présenté comme « ressource essentielle » et « solution ultime » aux défis économiques et sociaux de demain (Bullich, 2016).

Considérant l'opacité qui règne dans « l'économie numérique » et la culture du secret industriel qu'y manifestent les acteurs industriels (*a fortiori* les acteurs les plus importants), il faut généralement prendre des chemins détournés afin d'observer la conduite de ces mêmes acteurs industriels. C'est précisément la démarche qu'emprunte Caroline Creton. Elle s'intéresse à la façon dont Facebook est à même de valoriser auprès des musiciens les données dont l'entreprise dispose sur les goûts culturels des utilisateurs de son dispositif. La stratégie de l'entreprise californienne est ainsi mise en lumière par un travail sur les pratiques et usages de musiciens « à notoriété locale », leurs « tactiques », suivant la terminologie de de Certeau, pour répondre aux logiques commerciales de la plateforme et la compréhension qu'en ont ces usagers. Ceux-ci se trouvent face au dilemme que l'auteur exprime dans le titre de l'article : faut-il payer ou non pour toucher son public ? La pratique rappelle le « *pay per play* » pratiqué aux Etats-Unis dès le début du vingtième siècle par certaines salles de concerts et producteurs de spectacles et qui consiste à faire payer aux musiciens le droit de jouer sur scène en arguant de la force promotionnelle du lieu. Il en va de même en 2018 mais ce sont les données personnelles dont dispose Facebook qui font ici figure d'argument commercial décisif. Caroline Creton montre qu'au-delà de l'enjeu économique, il s'agit pour ces musiciens d'acquérir des compétences de l'ordre du marketing numérique afin d'utiliser au mieux les outils de la plateforme. En cela, « l'offre industrielle de Facebook participe aux mutations qui font du musicien un artiste-entrepreneur » (Creton, 2018).

Le deuxième article étudie les stratégies liées aux données massives d'un autre acteur majeur de la Silicon Valley : Netflix. Gabrielle Silva Mota Drumond, Alexandre Coutant et Florence Millerand proposent ainsi d'étudier les ressorts du dispositif de recommandation qui a fait le succès – tant médiatique qu'économique – de l'entreprise. La thèse est que ce dispositif serait à même de « produire l'utilisateur » ainsi que l'indique le titre de l'article. Afin d'étayer leur démonstration, les auteurs ont mis en place une méthodologie fondée principalement sur l'étude de la mise en visibilité des contenus de Netflix ainsi que sur les discours de l'entreprise sur son propre système de recommandation. Ce qui apparaît en effet remarquable est que celle-ci, au travers de son « *Techblog* », communique de façon régulière à ses différents publics sur les choix opérés par l'entreprise quant à ce dispositif censé « optimiser » l'expérience utilisateur. Or, cette optimisation

est ramenée, ainsi que nous l'indique les auteurs, à une consommation intensive, « excessive », de contenus audiovisuels (Drumond, Coutant, Millerand, 2018).

Le troisième et dernier article se rapporte quant à lui au champ politique et s'intéresse à la prégnance toute contemporaine des *big data* dans les stratégies électorales. Anaïs Théviot présente ainsi les stratégies des « prestataires en *big data* électoral » et retrace l'émergence puis l'affirmation d'une « croyance en l'efficacité » des données massives, croyance désormais partagée avec les décideurs politiques et militants. Ce que l'auteur met en évidence est que le moteur de cette technicisation des campagnes électorales réside dans la « promesse » portée par ces entrepreneurs en données quant aux possibilités de prédiction, d'orientation des stratégies électorales qu'une connaissance des électeurs serait en mesure de produire (Théviot, 2018). Ces discours de promotion ont eu une portée perlocutoire considérable puisque le recours aux données massives est maintenant perçu par les responsables des équipes électorales comme indispensable dans la panoplie des outils de campagne. En cela, l'article d'Anaïs Théviot illustre parfaitement les effets sociaux (en l'occurrence politiques même) produits par les stratégies des acteurs économiques autour de l'obtention et du traitement des *Big Data*.

La production des *data* : dispositifs et pratiques

Le deuxième axe concerne l'analyse du procès de production lui-même. Il s'agit ainsi d'ouvrir la « boîte noire », la focale se faisant plus resserrée. Les données sont souvent considérées comme « brutes », comme si elles étaient uniquement le point de départ de traitements algorithmiques, alors qu'elles sont « toujours déjà » le résultat de traitements élaborés, qu'elles portent sur des gisements informationnels ou documentaires volumineux préalablement construits (archives institutionnelles ou patrimoniales, bibliothèques numériques par exemple). Il s'agit dès lors de s'interroger à la fois sur la provenance des données mobilisées, sur leur transformation en informations, sur le « travail » des agents impliqués dans cette production ainsi que sur les modalités organisationnelles qui président à leur co-opération. C'est donc le *process* lui-même qui est ici au cœur de l'investigation : il s'agit d'appréhender à la fois empiriquement et théoriquement les différentes phases qui le composent, de l'obtention des *data* comme *inputs* à la production spécifique d'*outputs*, résultat, généralement automatisé, du traitement algorithmique. En outre, la transformation qui y est opérée doit être abordée sous l'angle de ce qui se perd, ou au contraire ce qui est ajouté, en termes informationnels, par le *process* même. Dans le cadre de ce deuxième axe, il s'agit donc de se centrer à la fois sur des dispositifs comme ensemble d'instruments et de techniques ainsi que sur des pratiques, fondées sur des savoirs et savoir-faire. Il s'agit d'interroger aussi les formes de (re-)contextualisation des *data*, les modalités de leur interprétation, la place des langages et celle des représentations visuelles.

En 2006, Bruno Maresca soulignait que le recours systématique aux enquêtes de fréquentation des bibliothèques de lecture publique en France faisait « apparaître ce domaine de la culture comme l'un des plus préoccupés de la quantification des publics ». En s'intéressant aux méthodes de recueil de données destinées à documenter les usages de la bibliothèque numérique Gallica, Philippe Chevallier (2018) s'inscrit à la fois dans une continuité, celle des études de publics, et une rupture, celle de l'incursion de méthodes complexes déléguées à des spécialistes de la fouille de données. Dans cette première contribution, Philippe Chevallier évoque comment les données de masse recueillies sous forme de logs, peuvent renseigner finement sur la temporalité des connexions (fréquence, durée) des usagers, sur leurs actions-types, sur leurs parcours de navigation ou encore sur les types de documents consultés. Il indique également que ces méthodes sont complémentaires d'autres enquêtes et observations qualitatives en vigueur à la BnF et met en garde

sur la nécessité d'une approche concertée de la donnée entre professionnels de bibliothèques et *data scientists* à des fins de gouvernance.

Le deuxième article proposé par Jean-Marc Francony (2018) s'intéresse aux flux de données publicisés dans les API de Twitter et porte un regard expert sur la fabrication des données et les logiques de publication. Tirant profit de sa notoriété, Twitter, qui figure parmi les dispositifs les plus utilisés, a récemment repensé son modèle économique en enrichissant l'accès à ses interfaces de données. Parallèlement, de nombreuses restrictions qualitatives et quantitatives ont été opérées sur les données, cette réduction informationnelle fondée sur des contraintes techniques conduisant in fine à réduire les possibilités d'accès aux réseaux d'acteurs et aux flux informationnels. Ce faisant, Twitter préserve et garde la maîtrise de son offre de service.

Enfin, l'axe se termine par la contribution de Valentyna Dymytra (2018) qui focalise son attention sur les médiations accompagnant l'ouverture des données numériques urbaines. A partir d'une sélection d'applications mobiles destinées à assister les usagers dans leurs déplacements en ville, l'auteur distingue plusieurs types de médiations, notamment informatiques et info-communicationnelles. Valentyna Dymytra montre de manière très convaincante qu'il est indispensable de considérer les conditions de production des applications, la nature des acteurs, les moyens matériels, technologiques et symboliques mobilisés pour révéler et éclairer les médiations techniques à l'origine des choix des développeurs. Ainsi, cette étude qui s'appuie sur les résultats d'une enquête menée en France dans le cadre d'une ANR (OpenSensingCity), illustre parfaitement notre propos. Elle montre que les nouveaux services urbains qui reposent sur la circulation et la valorisation des données sont toujours le résultat « d'une double médiation technique et sociale et reflètent des orientations des acteurs dominants qui en maîtrisent la production et la distribution ». (*Ibid.*)

L'intégration des *Big Data* et algorithmes au sein des secteurs professionnels et des champs sociaux

Le troisième axe rassemble des analyses sectorielles. Il s'agit de s'intéresser ici aux modalités d'intégration des dispositifs au sein de secteurs professionnels et de champs sociaux : assiste-t-on à une reconfiguration des activités qui se déploient en leur sein ou, plus modestement, à des mutations partielles de celles-ci ? Comment de grands ensembles d'activités comme la culture, le marketing, la santé, le journalisme ou l'éducation sont-ils affectés de façon spécifique et/ou de façon commune, comparable ? Afin d'apporter des éléments de réponses à ces questions, il s'agit de porter ici une attention plus soutenue aux professionnels qui participent à cette production. Les récentes fonctions aux dénominations anglophones qui se multiplient au sein des organisations telles que « *Data Scientist* », « *Data Analyst* », « *Chief Data Officer* » ou « *Data Protection Officer* », manifestent l'intérêt croissant des entreprises et des administrations pour l'obtention, l'agrégation et l'analyse de données massives préalablement à la conception et la mise en œuvre de leurs stratégies. Ces fonctions ainsi que leur intégration organisationnelle ont ainsi fait l'objet d'études spécifiques.

Premièrement, Jean-Sébastien Vayre s'est intéressé dans son article aux « technologies d'apprentissage artificiel appliquées à la gestion de la relation client ». Il a décortiqué minutieusement les dispositifs mis en place par différentes entreprises et un des points forts de l'étude est de révéler l'importance du « travail de cadrage de leurs activités inférentielles » (Vayre, 2018). Ce travail est le fruit de « *data scientists* » employés par les entreprises considérées. Au moyen d'une étude portant sur leur activité complétée par une analyse de l'architecture technico-logique de ces « machines apprenantes », l'auteur affirme que ces dispositifs sont à même, au

travers de ce qu'il nomme la « socialisation algorithmique » (c'est-à-dire des processus automatisés d'acquisition de compétences à la fois cognitives et communicationnelles), d'élaborer leurs propres critères de pertinence révélant leur capacité à produire, plus que reproduire, de l'information.

Le deuxième champ d'application abordé dans le cadre de cet axe est le domaine militaire. Olivier Koch propose ainsi un article exposant les modalités d'inclusion de machines prédictives, fondée sur l'obtention de données massives et leur traitement exhaustif *via* des algorithmes informatiques, au sein de l'armée des Etats-Unis d'Amérique. Depuis le début de la décennie 2000, l'Etat-Major américain s'est doté d'outils destinés à repérer et prévenir les foyers insurrectionnels sur les principaux terrains d'intervention où son armée s'est déployée (Afghanistan et Irak). Toutefois, au-delà du caractère directement opérationnel en zone de conflits et situation de « guerre irrégulière », ce que montre l'auteur est que l'intérêt de ce type de dispositif réside dans les possibilités (ou les « promesses » tout du moins) « d'optimisation de la prise de décision politique » (Koch, 2018). En effet, malgré « l'inefficacité des programmes mis en œuvre au regard des finalités que leur assignent leurs artisans-concepteurs », ces dispositifs continuent d'être financés et exploités précisément en raison, une nouvelle fois, d'une croyance « au plus haut niveau de l'Etat fédéral » dans l'efficacité de leur prédictions, non pas sur le plan de la polémologie mais sur celui de la gouvernementalité (*Ibid.*).

Le dernier article de ce troisième axe se rapporte aux « résistances épistémologiques, ruptures économiques et adaptations professionnelles » liées à l'intégration des *Big Data* au sein des médias d'informations. L'insertion des dispositifs liés à l'acquisition et au traitement des *Big Data* dans les champs professionnels ne se fait pas sans heurts et c'est précisément ce qu'illustre l'article d'Alexandre Joux et Marc Bassoni. Spécialistes du journalisme, les deux auteurs exposent les principales mutations du métier liées à la « data-ification » des rédactions. Ils procèdent pour ce faire à une « cartographie des études » récentes sur le sujet et mobilisent également des paroles d'acteurs qui témoignent de l'attitude pour le moins ambivalente de certains professionnels face au « journalisme robotisé » (Joux, Bassoni, 2018). Au final, ce que les auteurs critiquent foncièrement est la « vision ultra-techniciste » qui présente le traitement des données massives comme un nouveau mode d'investigation journalistique qui se distinguerait par son objectivité totale, et en cela, une « prophétie de la disparition du métier » pour les plus pessimistes (*Ibid.*). Les faits militent pour une perspective (évidemment) plus nuancée, bien que les mutations à venir s'annoncent profondes.

Les processus de régulation : gouvernance, lois et réglementations, normalisation technique et organisationnelle

Le quatrième et dernier axe porte sur les questions de régulation du procès de production « *Big Data* / algorithmes de traitement ». Par régulation, nous entendons aussi bien les règles s'imposant « verticalement » aux acteurs sociaux que celles produites par ceux-ci, généralement dans une visée d'optimisation de la coordination des activités. Ainsi, le regard se pose-t-il tant sur les lois et règlements produits par l'Etat et ses administrations que sur les normes *ad hoc* configurant les dispositifs et les pratiques professionnelles. Il s'agit notamment de rendre compte de la plasticité des règles produites ainsi que des asymétries en termes d'agentivité entre producteurs de *data* et producteurs d'informations à partir de ces *data*, entre les acteurs maîtrisant cette production d'information et leurs concurrents ne la maîtrisant pas, asymétries que ces règles sont susceptibles de renforcer ou de réduire. Il s'agit, de surcroît, de porter le regard sur les processus de production des normes et les jeux argumentatifs, sur les instances de délibérations ainsi que sur les discours qui y sont énoncés, considérant que ces discours sont porteurs de visions

du monde à partir desquelles ces normes se construisent et qui les justifient, les légitiment aux yeux des acteurs sociaux.

Enfin, ce dernier axe concerne également les travaux s'intéressant à l'inscription de la thématique des *Big Data* et algorithmes dans les politiques publiques. Au-delà des questions portant sur l'encadrement légal *stricto sensu*, il s'agit d'apprécier comment l'Etat investit ce domaine, dans le cadre des mesures concernant l'*Open Data* évidemment, mais aussi au travers de ses actions favorisant le développement des activités liées à la captation et au traitement des données massives : dans le domaine de la formation, dans celui de la standardisation technique, etc.

L'article d'Anne Lehmans s'inscrit précisément dans cette optique. L'auteur s'intéresse aux « politiques et pratiques de médiation, de valorisation et d'éducation autour des données ouvertes ». A la suite d'une enquête collective « sur plusieurs entités productrices et utilisatrices de données ouvertes dans l'académie de Bordeaux » et dont l'objectif est une meilleure connaissance de « la façon dont les enseignants recourent effectivement à ces données ouvertes dans un objectif pédagogique », l'auteur montre que les mesures traduisant les politiques autour de l'*Open Data* sous-estiment généralement le poids des contraintes d'appropriation (Lehmans, 2018). Par conséquent, si la mise à disposition de données publiques est susceptible de contribuer à un élargissement de l'espace public, en favorisant notamment un dialogue renouvelé entre « les collectivités publiques, les entreprises et les citoyens », l'auteur insiste sur le fait que le rapport aux données nécessite des compétences, une « culture », dont ne dispose pas systématiquement chaque citoyen ni même chaque entreprise ou administration. Un effort de formation voire « d'acculturation » est donc nécessaire (*Ibid.*).

On retrouve les interrogations sur la normalisation dans l'article de Guillaume Sire. A partir d'une étude du Web sémantique, l'auteur expose comment « la discrétisation des informations présentes sur le Web, c'est-à-dire leur transformation en *Big Data* » contribue à industrialiser des « politiques du sens » (Sire, 2018). Pour ce faire, il compare trois syntaxes et procédures de normalisation afférentes et montre comment les pratiques des producteurs de contenus se sophistiquent dans la perspective d'un référencement optimisé. Pis, il identifie des « stratégies prédatrices » qui passent par l'imposition d'éléments syntaxiques que mènent les géants du Web, au premier desquels se trouve Google/Alphabet. Derrière des procédures de normalisation et de régulation pour plus d'interopérabilité, se produit une « tectonique [...] dont dépendent les modalités concrètes de production et de mise en circulation à grande échelle des signes, c'est-à-dire des informations transformées en données qualifiables, puis qualifiées ». Or, ainsi que le pose l'article, cette « tectonique » manifeste avant toute chose des mouvements industriels, des affrontements entre « visions [...], projets, intérêts » divergents (*Ibid.*).

Le dernier article de ce quatrième axe renvoie à la question du droit à la protection des données personnelles au sein de la recherche. Dans leur contribution, Julien Rossi et Jean-Edouard Bigot indiquent que les chercheurs en SHS seraient actuellement mis au défi d'intégrer les équipements numériques pour mener leurs recherches. Ce faisant, les chercheurs seraient progressivement confrontés à divers types de données numériques (traces numériques, réponses à des questionnaires, entretiens, etc.) qui, en raison de leur caractère personnel, tomberaient sous le coup de la loi, notamment du Règlement général de protection des données entré en vigueur le 25 mai 2018. A la faveur d'entretiens semi-directifs conduits auprès de chercheurs en SHS engagés dans une ANR, les auteurs ont enquêté sur les formes d'autorégulation et sur les questionnements éthiques soulevés par ce cadrage juridique des données personnelles (Rossi, Bigot, 2018).

Conclusion

Les articles composant ce dossier 2018 des *Enjeux de l'information et de la communication* envisagent donc la question des données massives aux prismes des approches et des méthodologies multiples mobilisées en sciences de l'information et de la communication. Si la question de l'épistémologie n'est pas abordée de front, il ressort toutefois de ces 12 contributions un consensus fort quant à l'exigence scientifique de considérer ces données comme, avant toute chose, des construits sociaux. En cela, elles manifestent autant les perspectives, les positionnements, voire les stratégies de leurs (co-)concepteurs que les caractéristiques supposées « immanentes » des objets analysés ; ignorer leurs conditions de production revient donc à se priver d'un élément majeur de compréhension de leur fonctionnement et de leurs effets sociaux. Ces contributions éclairent également les modalités de production du social induites en retour par les traitements des données. Ce faisant, les différents articles exposent, chacun à leur façon, comment les dispositifs contribuent à produire ou reproduire des relations entre les individus ou alors les relations des individus aux institutions qui les gouvernent.

Au travers de ce parcours en quatre étapes, nous espérons présenter au lecteur un large panorama des modes de production « algorithmique » de la société, en alternant les points de vue, les focales et en multipliant les objets visés.

Avant de clore cette présentation, nous souhaitons remercier très chaleureusement l'ensemble des auteurs ayant contribué au dossier : pour la qualité de leurs travaux ainsi que pour leur grande réactivité lors des échanges que nous avons eus avec le comité éditorial.

Références bibliographiques

Aiden E. et Michel, J.-B., *Uncharted: Big Data as a Lens on Human Culture*, New York : Riverhead Books, 2014.

Anderson C., « The End of Theory: The Data Deluge Makes the Scientific Method Obsolete », *Wired*, 2008 [en ligne] : <http://www.wired.com/2008/06/pb-theory/> (consulté le 10/04/2017).

Bardou-Boisnier S. et Pailliant I. (coord.), dossier : *Information publique : stratégies de production, dispositifs de diffusion et usages sociaux*, *Les Enjeux de l'Information et de la Communication*, n° 13/2, 2012, [en ligne] : http://w3.u-grenoble3.fr/les_enjeux/pageshtml/art2012.html#dossier (consulté le 12/04/2017).

Bullich, V., « Big Data : stratégies industrielles et économie de la promesse. », dans Pilati A. (dir.), *La comunicazione multipla. Media, piattaforme digitali, Over the Top, Big Data*, Rome : Magna Carta Edizioni, 2016, p. 41-71.

Certeau, M. (de), *L'invention du quotidien. Tome 1. Arts de faire*, Paris : Gallimard, 2002 (1^{ère} édition : 1980)

Chevallier, P., « Les données au service de la connaissance des usages en ligne : l'exemple de l'analyse des logs de Gallica », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication*, *Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Creton, C., « To pay or not to pay : les musiciens à notoriété locale face à la publicité ciblée sur Facebook », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication*, *Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Delort, P., *Le Big Data*, Paris : Presses Universitaires de France, 2015.

Desrosières A., *Gouverner par les nombres*, Paris : Presses des Mines ParisTech, 2008

Drumond, G., Coutant, A., Millerand, F, « La production de l'utilisateur par les algorithmes de Netflix », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Dymytrova V., « Les médiations de l'open data au prisme des applications liées à la mobilité », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Francony J.-M., « L'éditorialisation des données aux bornes des API : Enjeux et perspectives pour une analyse empirique », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Joux, A., Bassoni, M., « Le journalisme saisi par les Big Data ? Résistances épistémologiques, ruptures économiques et adaptations professionnelles », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Kitchin, R. *The Data Revolution*, Londres : Sage, 2014.

Koch, O., « Les données de la guerre. Big Data et algorithmes à usage militaire », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Lehmans, A., « Les réinventions de la démocratie à l'aune de l'ouverture des données : du discours de la participation aux contraintes de la gouvernance », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Maresca B. « Les Enquêtes de fréquentation des bibliothèques publiques ». *Bulletin des bibliothèques de France (BBF)*, 2006, n 6, p. 14-19. [en ligne] <http://bbf.enssib.fr/consulter/bbf-2006-06-0014-003> (consulté le 25 septembre 2018)

Mattelart A., *Histoire de la société de l'information*, Paris : La Découverte, 2001.

Mayer-Schönberger V et Cukier K., *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Boston : Eamon Dolan/Houghton Mifflin Harcourt, 2014.

Miège B., *L'information-communication, objet de connaissance*, Paris : INA Editions, 2004.

Rossi J. et Bigot J.-E., « Traces numériques et recherche scientifique au prisme du droit des données personnelles », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Sire, G., « Web sémantique : les politiques du sens et la rhétorique des données », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Théviot, A., « « Une économie de la promesse » : mythes et croyances pour vendre du Big data électoral », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.

Touraine A., *Production de la société*, Paris : Le Seuil, 1973.

Vayre, J.S., « Les machines apprenantes et la (re)production de la société : les enjeux communicationnels de la socialisation algorithmique », dossier : *Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication, Les Enjeux de l'Information et de la Communication*, n° 19/2, 2018.