

APPEL À CONTRIBUTIONS

Numéro spécial *Revue française de sociologie*

Big Data, Sociétés et Sciences Sociales

Coordination scientifique :

Gilles BASTIN (Laboratoire PACTE, Sciences Po Grenoble),
Paola TUBARO (Laboratoire de Recherche en Informatique, CNRS, Université de Paris-Saclay)

Peu de sujets ont suscité dans les années récentes autant d'intérêt dans le débat public et dans les sciences sociales que celui des *big data*. La montée en puissance de nouvelles sources de données massives produites principalement par les technologies numériques alimente, depuis le début des années 2010, une réflexion et des craintes sur le fonctionnement des sociétés contemporaines et la production du savoir sur ces sociétés. Une grande part de la littérature consacrée aux *big data* oscille encore entre deux approches. La première vise à caractériser de manière instrumentale ces données (par exemple par opposition aux données plus classiques utilisées par les chercheurs) et l'usage qui en est fait (Kitchin, 2013). La seconde met en avant de manière critique les risques induits par les *big data* : non seulement l'obsolescence de la méthode scientifique d'analyse des données appelée à être remplacée par des algorithmes sans lien fort avec les théories sociales (Anderson, 2006), mais aussi, et surtout, l'apparition d'une nouvelle forme de société « dirigée par les données » (Pentland, 2012), inaugurant des bouleversements de nos façons de vivre, de travailler et de penser (Mayer-Schönberger et Cukier, 2013) ou une nouvelle « gouvernamentalité » (Rouvroy et Berns, 2013).

On dispose cependant encore de peu de travaux interrogeant les effets sociaux et les implications scientifiques des *big data* à partir d'une expérience de première main dans le domaine. Cet appel à contributions vise à mobiliser la communauté des sociologues qui s'intéressent aux *big data* et qui les utilisent autour de deux grandes questions qui ne nous paraissent pas devoir être séparées : comment les *big data* transforment-elles la société ? Comment ces données affectent-elles la pratique des sciences sociales ? Cette double approche consistant, dans une perspective de sociologie des processus de *datafication*, à réfléchir à la quantification du social comme Alain Desrosières l'avait fait au sujet des statistiques, considérant en même temps « leurs apports de connaissance et les circuits sociaux de leur mise en forme et de leurs usages » (Desrosières, 2005). Notre objectif est par là de contribuer à la réflexion engagée depuis quelques années dans les sciences sociales sur les effets sociaux, économiques, juridiques et politiques, d'une part, méthodologiques et épistémologiques, d'autre part, du « déluge de données » (Hey et Trefethen, 2003) qui a accompagné trois phénomènes sociaux congruents. Le premier est le développement des pratiques de documentation et de digitalisation de pans entiers de l'expérience sociale des individus sur le web. Le second concerne la numérisation des produits culturels comme les livres, la musique ou les médias. Le troisième enfin est la généralisation de la collecte de traces d'usages et de pratiques sociales par des capteurs connectés au web dans des domaines très variés comme la santé, les déplacements ou la vie domestique.

Un premier type d'articles attendus concerne la généalogie du phénomène *big data*, aussi bien comme une forme d'innovation scientifique, dans des champs disciplinaires plus ou moins proches des sciences sociales, que comme un sujet du débat public (Beer, 2016). Les conditions historiques de diffusion du concept et des méthodes des *big data* pourront être examinées à l'aide d'études comparatives qui seraient utiles pour comprendre dans quelle mesure le contexte institutionnel pèse sur leur généralisation et la façon dont elles ont pénétré les espaces publics et académiques dans différents pays. La question du rapport (ou de l'écart) entre définition « publique » et définition « scientifique » des *big data* (Michael et Lupton, 2016) pourrait aussi être abordée dans des articles analysant ce phénomène sur un plan généalogique.

Les contributions pourront aussi aborder les nombreux problèmes publics qui ont émergé dans le sillage des *big data*. La façon dont la révolution des données modifie la perception de la vie privée et brouille la frontière entre sphère publique et sphère privée est un premier exemple de ces problèmes publics (Tubaro, Casilli et Sarabi, 2014). D'autres phénomènes peuvent être cités comme la question de la surveillance des populations rendue possible par la récupération massive d'informations sur les individus à partir des traces numériques de leurs activités. La question du « *digital labor* », souvent invisible et non rémunéré, qui permet d'alimenter les bases de données (Terranova, 2000 ; Cardon et Casilli, 2015 ; Scholz, 2012) comme celle du développement des algorithmes destinés à les traiter pourra aussi donner lieu à des propositions d'articles. Le regard peut aussi se porter sur la propriété privée des données et sur l'injection de ressources informationnelles massives dans la chaîne de valeur économique, qui sont souvent l'apanage de grandes entreprises et qui déplacent les rapports de pouvoir en faveur de celles-ci, par rapport aux États et aux autorités publiques. Les contributions proposées pourront s'intéresser aux formes de discrimination « algorithmique » résultant des segmentations de marché très fines que les *big data* facilitent, ou aux inégalités d'éducation et d'accès aux données. Enfin des articles pourront aborder la manière dont les *big data* peuvent *a contrario* faciliter l'implication citoyenne et la participation démocratique, donnant une voix à des couches de la population traditionnellement moins bien représentées, notamment par la mise à disposition de données dans le cadre de l'*open data* (mouvements sociaux, datajournalisme, etc.).

Dans une perspective réflexive de sociologie de la profession de sociologue, d'autres propositions pourront aborder les effets de l'usage de ce type de données dont l'origine se situe généralement dans la pratique des individus et des institutions, en dehors de toute démarche de recherche. Si l'on adopte le point de vue réflexif de Burrows et Savage dans leur fameux article sur la « *coming crisis* » de la sociologie empirique (Savage et Burrows, 2007), celle-ci serait simplement en train de perdre sa « juridiction » sur tout un pan de la connaissance de la société. L'entretien et l'enquête par questionnaire qui lui ont longtemps assuré cette juridiction seraient en effet dépassés par de nouveaux modes de représentation de la société sans lien évident avec les connaissances sociologiques acquises et fondées sur la *commodification* des données personnelles.

La *Revue française de sociologie* souhaite aussi susciter des réflexions d'ordre méthodologique sur les *big data*. Les chercheurs en sciences sociales se sont engagés récemment dans l'analyse et, de plus en plus, l'utilisation de ces nouveaux matériaux (Cardon, 2012). Il nous semble que ce numéro pourrait leur donner l'occasion de rendre visibles les écueils qu'ils rencontrent en chemin et la façon dont ces données mettent en tension les « assemblages » techniques, méthodologiques et déontologiques qui forment leur appareil de preuve (Ruppert, Law et Savage, 2013). Si, dans le discours courant, les *big data* sont souvent présentés en termes de rupture, ces contributions pourront notamment s'interroger à partir d'exemples concrets sur les éléments de continuité que l'on peut trouver dans leur utilisation par rapport à celle des *small* et des *medium data* dont s'est nourrie la sociologie tout au long de son histoire : l'enquête de terrain, l'interview, la statistique publique, etc. Un défi qui pourrait être discuté est celui des compétences

désormais exigées du sociologue et par là, de l'interdisciplinarité nouvelle qui se crée dans les lieux de production de ces données. L'usage des *big data* suppose en effet que le sociologue se forme aux méthodes de pointe de l'informatique (*machine learning*, visualisation des données) ou collabore avec des spécialistes de ces méthodes. Les effets de ces changements doivent être analysés dans le détail à l'image des opérations concrètes de ce que l'on peut appeler la *datafication*, par exemple le travail des « petites mains » des données (Dagiral et Peerbaye, 2012) ou des conflits d'interprétation qui naissent de la production des données d'enquête lorsque le sociologue doit les négocier avec leurs producteurs individuels, les plateformes d'agrégation de contenus et la CNIL (Bastin et Francony, 2016). La question de l'échantillonnage des données pourra aussi être au centre de propositions d'articles dans la lignée des débats soulevés par exemple par le *Great British Class Survey* (Savage et al., 2013 ; Mills, 2014).

Enfin, sur un plan épistémologique, des articles sur les aspects théoriques de l'épistémologie des sciences sociales à l'heure des *big data* pourront être publiés dans ce numéro. Les *big data* bouleversent-ils l'espace de l'enquête et le raisonnement sociologique (Passeron, 1991) comme l'avancent certains en craignant le glissement de l'explication causale à la simple mise en évidence de corrélations que provoqueraient les *big data* ? Comment comprendre l'émergence de notions comme celle de « trace » qui remet au goût du jour le paradigme de l'indice (Ginzburg, 1980) dans les études sur le numérique (Merzeau, 2009) ou encore celui de l'enquête fondée sur le modèle de la police scientifique — « *forensic social science* » (Goldberg, 2015) ?

Si ces quelques exemples n'épuisent pas l'ensemble des interrogations possibles sur le sujet de ce numéro spécial, il est opportun de préciser que des travaux discutant d'aspects méthodologiques spécifiques ou des défis techniques posés par les *big data* (par exemple, les problèmes de documentation ou de curation) seront les bienvenus pourvu qu'ils apportent des éléments permettant de les relier à la théorie ou la pratique sociologique. Réciproquement, sur les questions les plus théoriques soulevées par la thématique du numéro, un ancrage dans une étude de cas documentée précisément sera fortement apprécié.

Gilles Bastin et Paola Tubaro

REFERENCES

- Anderson, C. (2006). « The End of Theory: The Data Deluge Makes the Scientific Method Obsolete », *Wired*.
- Bastin, G. et Francony, J.-M. (2016). « L'inscription, le masque et la donnée. Datafication du web et conflits d'interprétation autour des données dans un laboratoire invisible des sciences sociales », *Revue d'Anthropologie des connaissances*, 10(4).
- Beer, D. (2016). « How should we do the history of Big Data ? », *Big Data & Society*, 3(1).
- Cardon, D. (2012). « Regarder les données », *Multitudes*, 2012/2, n° 49.
- Cardon, D. et Casilli, A.A. (2015). *Qu'est-ce que le digital labor ?*, Éditions de l'INA.
- Dagiral, É. et Peerbaye, A. (2012). « Les mains dans les bases de données », *Revue d'anthropologie des connaissances*, 6(1).
- Desrosières, A. (2005). « Décrire l'État ou explorer la société : les deux sources de la statistique publique », *Genèses*, n° 58.
- Ginzburg, C. (1980). « Signes, traces, pistes », *Le débat*, n° 6.
- Goldberg, A. (2015). « In defense of forensic social science », *Big Data & Society*, 2(2).
- Gray, J., Chambers, L. et Bounegru, L. (2012). *The data journalism handbook*. O'Reilly Media, Inc.

- Hey, A. J. et Trefethen, A. E. (2003). « The data deluge: An e-science perspective » In, Berman, F., Fox, G. C. and Hey, A. J. G. (eds.) *Grid Computing - Making the Global Infrastructure a Reality*, Wiley and Sons.
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*, Sage.
- Mayer-Schönberger, V. et Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*, Houghton Mifflin Harcourt.
- Merzeau, L. (2009). « Du signe à la trace : l'information sur mesure », *Hermès, La Revue*, 2009/1, n° 53.
- Michael, M. et Lupton, D. (2016). « Toward a manifesto for the 'public understanding of big data' », *Public Understanding of Science*, 25(1).
- Mills, C. (2014). « The great British class fiasco: A comment on Savage et al. », *Sociology*, 48(3).
- Passeron, J.-C. (1991). *Le raisonnement sociologique : l'espace non-poppérien du raisonnement naturel*, Paris, Nathan.
- Pentland, A. (2012). « Reinventing society in the wake of big data ». Edge. Available online at: https://www.edge.org/conversation/alex_sandy_pentland-reinventing-society-in-the-wake-of-big-data
- Rouvroy, A. et Berns, T. (2013). « Gouvernamentalité algorithmique et perspectives d'émancipation », *Réseaux*, 2013/1, n° 177.
- Ruppert, E., Law, J. et Savage, M. (2013). « Reassembling social science methods: The challenge of digital », *Theory, culture & society*, 30(4).
- Savage, M. et Burrows, R. (2007). « The coming crisis of empirical sociology », *Sociology*, 41(5).
- Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., Le Roux, B., Friedman, S. et Miles, A. (2013). « A new model of social class? Findings from the BBC's Great British Class Survey experiment ». *Sociology*, 47(2).
- Scholz, T. (ed.) (2012). *Digital labor: The Internet as playground and factory*, Routledge.
- Terranova, T. (2000). « Free labor: Producing culture for the digital economy ». *Social text*, 18(2).
- Thrift, N. (2005). *Knowing capitalism*, Sage.
- Tubaro, P., Casilli, A.A. et Sarabi, Y. (2014). *Against the Hypothesis of the End of Privacy*, Springer.

Les propositions de contribution (min. 1 000 mots – max. 1 500 mots, bibliographie non incluse), en français ou en anglais, devront être adressées à Christelle Germain (christelle.germain@cncs.fr), secrétaire de rédaction, au plus tard le 28 février 2017.

Elles feront l'objet d'un examen conjoint par les signataires de cet appel et un autre membre de la Rédaction. La notification d'acceptation sera rendue au plus tard le 30 mars 2017.

Les auteurs dont la proposition aura été retenue devront soumettre leur texte, dont la longueur ne dépassera pas 70 000 signes (espaces, bibliographie et figures compris), au plus tard le 15 septembre 2017. Chaque article sera évalué indépendamment par les coordinateurs scientifiques du dossier et, de manière anonyme, par le comité de lecture de la *Revue française de sociologie*.