



HAL
open science

Les sciences sociales face aux traces du Big Data

Dominique Boullier

► **To cite this version:**

Dominique Boullier. Les sciences sociales face aux traces du Big Data : Société, opinion ou vibrations ?. Revue Française de Science Politique, 2015, 65 (5-6), pp.805 - 828. 10.3917/rfsp.655.0805 . hal-02365435

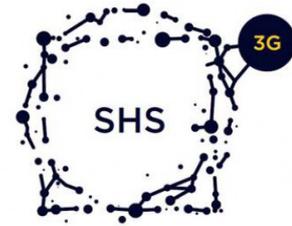
HAL Id: hal-02365435

<https://sciencespo.hal.science/hal-02365435>

Submitted on 15 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Programmes scientifiques

Les sciences sociales face aux traces du big data ? Société, opinion et répliques

Dominique Boullier

N°88 | avril 2015

Une troisième génération de sciences sociales doit voir le jour pour assumer la spécificité du monde de données et de traces créées par les réseaux numériques, sans se contenter de prolonger les acquis des sciences de la « société » et de l'« opinion ». Ces entités ont été construites dans une époque précise dont la généalogie est restituée pour être comparée avec le travail des agences qui exploitent les traces numériques et qui peuvent produire toute la réflexivité nécessaire en devenant prédictives. Il est proposé de penser les traces numériques en tant que « répliques » que les sciences sociales doivent suivre avec des méthodes adaptées car elles constituent désormais un nouveau continent du social.

Working Papers Series

Les sciences sociales face aux traces du big data ? Société, opinion et répliques

Dominique Boullier

Avril 2015

L'auteur

Dominique Boullier est professeur des universités en sociologie à Sciences Po. Docteur en sociologie de l'École des hautes études en sciences sociales (EHESS, 1987), diplômé de linguistique (Rennes 2, 1991), HDR en sciences de l'information et de la communication (Bordeaux 3, 1995). Souvent défini comme un « entrepreneur de recherche » en raison de son expérience passée de chef d'entreprise, de créateur de labos, de leader de projets multipartenaires (ANR ou européen) et de partenaire d'entreprises dans ses projets de recherche, il est directeur exécutif du projet Idefi Forccast (formation par la cartographie des controverses à l'analyse des sciences et des techniques) (2012-2019) et responsable de la pédagogie numérique à Sciences Po (avec Pascale Leclercq).

Parmi ses publications : *Événements et sécurité. Les professionnels des climats urbains*, Paris, Les Presses des Mines, 2012 (avec Stéphane Chevrier et Stéphane Juguet) ; *Opinion mining et sentiment analysis. Méthodes et outils*, Open Press Éditions (collection Sciences Po médialab) (avec Audrey Lohard) ; *La ville événement*, Collection «La ville en débat», PUF, 2010 ; *La Télévision telle qu'on la parle. Trois études ethnométhodologiques*, Paris, L'Harmattan, 2004 ; *L'Outre-lecture. Manipuler, (s')appropriier, interpréter le Web*, Paris, Bibliothèque Publique d'Information/Centre Pompidou, 2003 (avec Franck Ghitalla et alii).

Le texte

Ce texte sert de sous-basement à la réflexion collective inscrite dans le séminaire SHS 3G 2015, animé par Dominique Boullier et Françoise Thibault.

Pour en savoir plus : shs3g.hypotheses.org

Remerciements

Ce texte a bénéficié des commentaires critiques de E. Didier, G. Bowker, B. Latour, N. Marres et F. Thibault. Qu'ils en soient remerciés. Il a fait l'objet d'une première présentation sommaire au colloque Big Data du Collège de France en juin 2014 (chaire de P.-M. Menger).

Citer ce document

Dominique Boullier, *Les sciences sociales face aux traces du big data ? Société, opinion et répliques*, FMSH-WP-2015-88, avril 2015.

© Fondation Maison des sciences de l'homme - 2015

Informations et soumission des textes :

wpfmsh@msh-paris.fr

Fondation Maison des sciences de l'homme
190-196 avenue de France
75013 Paris - France

<http://www.fmsch.fr>

<http://halshs.archives-ouvertes.fr/FMSH-WP>

<http://wpfmsh.hypotheses.org>

Les Working Papers et les Position Papers de la Fondation Maison des sciences de l'homme ont pour objectif la diffusion ouverte des travaux en train de se faire dans le cadre des diverses activités scientifiques de la Fondation : Le Collège d'études mondiales, Bourses Fernand Braudel-IFER, Programmes scientifiques, hébergement à la Maison Suger, Séminaires et Centres associés, Directeurs d'études associés...

Les opinions exprimées dans cet article n'engagent que leur auteur et ne reflètent pas nécessairement les positions institutionnelles de la Fondation MSH.

The Working Papers and Position Papers of the FMSH are produced in the course of the scientific activities of the FMSH: the chairs of the Institute for Global Studies, Fernand Braudel-IFER grants, the Foundation's scientific programmes, or the scholars hosted at the Maison Suger or as associate research directors. Working Papers may also be produced in partnership with affiliated institutions.

The views expressed in this paper are the author's own and do not necessarily reflect institutional positions from the Foundation MSH.

Résumé

Une troisième génération de sciences sociales doit voir le jour pour assumer la spécificité du monde de données et de traces créées par les réseaux numériques, sans se contenter de prolonger les acquis des sciences de la « société » et de l'« opinion ». Ces entités ont été construites dans une époque précise dont la généalogie est restituée pour être comparée avec le travail des agences qui exploitent les traces numériques et qui peuvent produire toute la réflexivité nécessaire en devenant prédictives. Il est proposé de penser les traces numériques en tant que « répliques » que les sciences sociales doivent suivre avec des méthodes adaptées car elles constituent désormais un nouveau continent du social.

Mots-clefs

sciences sociales, numérique, big data, épistémologie, sociologie, statistique, opinion

Social sciences and traces of big data. Society, opinion, and vibrations

Abstract

The third generation of Social Sciences currently emerging must assume the specific nature of the world of data created by digital networks, without falling back on the Sciences of «Society» or «opinion.» All of these entities were constructed in specific political, institutional situations and with specific techniques. «Social» Big Data is taken into account by agencies specializing in the processing of this data to produce potentially predictive correlations primarily for the benefit of brands. The risk is that the Social Sciences would be lastingly disqualified from producing reflexivity that has hitherto been their raison d'être. Beyond «society» and «opinion» for which the text lays out a genealogy, appear the «traces» that must be theorized as «vibrations» by the Social Sciences in order to reap the benefits of the uncertain status of entities' widespread traceability.

Keywords

social sciences, digital, big data, epistemology, sociology, statistics, opinion

Sommaire

L'âge du numérique	5
Ni personnes ni identités, les traces sont la matière première	5
Les traces sont produites par des plates-formes	6
L'emprise des marques sur les traces	7
La construction « de l'opinion »	8
Les médiations qui font exister l'opinion publique sont constituées	8
Des marchés et des publics nationaux : les échelles des médias	9
L'opinion publique existe, je l'ai mesurée	9
La fabrication de « la société »	11
Le temps des calculs et des machines à calculer	12
Le pouvoir d'agir des dispositifs techniques de calcul	13
Ce que les sciences sociales peuvent faire du numérique, ce que le numérique fait aux sciences sociales	13
Les traditions de l'opinion et de la société équipées par le numérique	14
La fin de la théorie sociale ?	14
Les propriétés des sciences sociales de troisième génération	16
<i>Volume et exhaustivité</i>	16
<i>Variété et représentativité</i>	16
<i>Vélocité et traçabilité</i>	16
Des traces aux répliques	18
Conclusion	20
Tableau des trois générations de sciences sociales	20
Références	22

Les sciences humaines et sociales dans leur ensemble et la sociologie en particulier ont depuis plusieurs années adapté leurs méthodes et leurs techniques à la puissance de calcul et aux nouvelles données fournies par le numérique. Dans certains cas, les traditions plus anciennes s'équipent de façon plus efficace, dans d'autres cas, de nouvelles méthodes sont expérimentées (enquêtes en ligne, croisement des bases de données, etc.), et des sous-disciplines tentent de faire leur place comme les web studies, au sein d'ensembles plus vastes comme les humanités numériques. Ces phénomènes s'accompagnent d'une timide reconnaissance institutionnelle du besoin de très grandes infrastructures de recherche pour les sciences humaines et sociales, ce qui n'est pas sans susciter quelques émois au sein des communautés scientifiques et des décideurs. La rapidité des transformations a laissé à la traîne une réflexion d'ampleur sur ces phénomènes. Souvent,

le numérique est traité soit comme une technique de plus, soit comme un domaine à part, et il ces approches ne suffisent pas à rendre compte de ce qui est en jeu dans la mutation numérique en cours. En effet, parce que la modalité de réflexivité sur les sociétés fournie par les sciences sociales repose sur un ensemble de dispositifs parfois très longs à constituer mais qui, font convention (Eymard-Duvernay et al., 2004), il nous semble urgent d'explorer de nouvelles pistes de recherche. Nous voudrions proposer ici une grille de lecture de la mutation numérique en tant que nouvelle réflexivité offerte aux sociétés : de nouvelles sources de données sont désormais disponibles, au-delà des recensements et des registres (Desrosières, 1993) ou des sondages et des questionnaires¹. De ce fait, de nouvelles entités sont rendues accessibles par le numérique, que ni « la société », et ses propriétés socio-démographiques, ni « l'opinion » ne peuvent englober. Une nouvelle génération de sciences sociales frappe à la porte (la 3G qui aura cependant peu de chances d'exister si l'on ne mesure pas que des acteurs, des plates-formes et des producteurs/ capteurs de traces tendent à occuper tout le terrain. Pour le dire rapidement, le marketing et les « computer sciences » s'approprient et génèrent des outils de suivi de la vie sociale, sous forme de suivi de marques, de réputations, de communautés, de réseaux sociaux, etc.

1. Blondiaux, 1998. Plusieurs remarques prennent appui sur ce travail fondamental pour notre réflexion.

qui peuvent se passer des interprétations et des modèles des sciences humaines et sociales car ils compensent par une puissance de calcul et une traçabilité inédite, celles du Big Data. Leur souci principal reste l'action, la réaction, et non l'analyse ou la compréhension telles que les traditions de la sociologie et des autres sciences humaines et sociales les avaient définies. Traces et non plus données, réactivité et non plus réflexivité, le monde numérique se trouve façonné par des principes qui laissent de moins en moins de place aux SHS. Il n'y a en effet aucune raison pour que l'autorité des SHS ne soit pas remise en cause comme le furent toutes les autorités depuis l'avènement du numérique en réseaux. Mais comme le dit Geof Bowker, ce glissement n'est pas sans perturber tous nos repères : « For those of us brought up learning that correlation is not causation, there's a certain reluctance to examine the possibility that correlation is basically good enough. It is surely the case that we are moving from the knowledge/power nexus portrayed by Foucault to a data/action nexus that does not need to move through theory: All it needs is data together with preferred outcomes. » (Bowker, 2014).

L'âge du numérique

Ni personnes ni identités, les traces sont la matière première

Depuis de nombreuses années, mais de façon étendue avec les réseaux sociaux, les « computer sciences » calculent et modélisent le social comme si les traces récoltées permettaient d'accéder aux « vrais » individus mieux que tous les sondages, toutes les enquêtes et tous les recensements. Prenons deux exemples, l'un académique et l'autre commercial :

- « The Web does not just connect machines, it connects people. » (Knight Foundation, 14 September 2008). Voilà ce que déclarait Sir Tim Berners-Lee, fondateur du web en 1991 avec René Caillau, voulant insister ainsi sur le passage à une dimension du réseau non plus technologique (internet) ni documentaire (le WWW), mais sociale.
- Facebook de son côté a réussi le tour de force de rendre « normal » du point des acteurs eux-mêmes de déclarer son identité véritable, c'est-à-dire celle fournie par l'état-civil, son nom et son prénom, contre la tradition d'anonymat sur le web. La plate-forme prétend

ainsi devenir le monde de référence, voire l'état-civil de substitution, en compétition avec Google sur ce plan.

Or, rien ne permet de garantir quelque lien que ce soit entre les identités de Facebook ou « les gens » de Berners-Lee et des personnes identifiables par l'état-civil. Ce qui est connecté ne sont que des comptes et les données récupérées ne sont que des traces d'activité d'une entité qui peut prendre éventuellement les formats de l'état-civil. Dans le cas des scores qui permettent de classer les sites sur un moteur de recherche, la topologie des sites qui en résulte ne traite jamais de leurs contenus en tant que tels, mais des liens entrants et des liens sortants qui produisent un rang d'autorité ou de hub, au sens de la topologie des réseaux (Keinberg and al, 1998) et non d'un statut social.

Précisons ici d'emblée ce que nous entendons par traces dans le but de les distinguer des données. Les traces peuvent aller de signaux (« bruts », générés par des objets) à des verbatims non structurés, elles peuvent être des traces (liens, clics, likes, cookies)² exploitées en bases de données par les opérateurs ou les plates-formes mais aussi captées indépendamment de cela à travers les API et qui ne relèvent pas alors de bases de données relationnelles (Bowker and Star, 1999). Les traces ne sont pas nécessairement préformatées pour un calcul précis ni dépendantes de l'agrégation que l'on peut appliquer ensuite. Il est aisé de soutenir que, malgré tout, « derrière » les sites ou « derrière » les clics, il y a bien des humains, mais cela n'enlève rien au fait que les algorithmes, eux, ne s'intéressent pas à cette propriété et que, de plus, aucune garantie ne peut être apportée sur ce plan. Les traces entendues en ce sens restreint, sont produites par les plates-formes et les systèmes techniques numériques, mais ne sont pas les « signes » ou les indices d'autre chose qu'elles-mêmes tant que les relations avec d'autres attributs ne sont pas créées. Cela les distingue radicalement des données que l'on peut récupérer en masse sur des fichiers clients ou encore à partir d'actes administratifs. Certes, les méthodes de calcul du Big Data peuvent y être appliquées dans les deux cas, mais les traces sont a priori indépendantes des autres attributs, en particulier socio-démographiques qui sont rarement mobilisées dans les corrélations recherchées entre traces. Les

2. D. Cardon en a proposé une typologie comprenant les liens, les clics, les likes et les traces (Cardon, 2013)

relations avec des paramètres plus classiques dans les sciences des données se limitent au temps (un timestamp) et au lieu (les tags de géolocalisation), qui permettent de produire des timelines et des cartes qui deviennent des modes de présentation simplifiés des traces.

Les traces sont produites par des plates-formes

Pour Amazon ou Apple (puisque le web n'est plus distribué mais bien accaparé par ces quatre plates-formes GAFa qui centralisent la majorité du trafic), ce ne sont pas non plus des personnes qui sont mises en relation mais avant tout des goûts (livres ou musique à l'origine), exprimés par des traces d'achat, de choix, qui peuvent être traitées en masse pour produire des patterns, des profils, indépendamment des informations personnelles. Certes, il convient de ne pas oublier que toutes ces plates-formes sans exception sont aussi très friandes de données de type état-civil, numéros de téléphones et autres ressources très intéressantes pour les annonceurs à qui elles les revendent. Les méthodes de marketing qui en découlent reposent largement sur de l'adressage de masse de publicités ou de mails à des adresses IP ou mails qui ont cliqué sur un article (retargeting) mais beaucoup plus rarement sur des mises en relation sophistiquées avec les autres attributs des supposées personnes attachées à ces adresses ou à ces clics (profiling).

Les traces de comportements numériques constituent ainsi une « matière première » particulièrement profitable, sans qu'il soit nécessaire de faire appel aux sciences sociales. Que doivent faire les sciences sociales face à cette situation ? Deux options s'offrent à elles : soit elles se cantonnent à leur monde fait de données administratives, d'enquêtes et de sondages, en relativisant l'intérêt de ce type de traces et en privilégiant les données ; soit elles décident de chevaucher le tigre et de prendre ces traces comme matière première à leur tour à condition de les réorienter, de faire du « repurposing » comme le propose Richard Rogers. Elles doivent alors accepter de dépendre des plates-formes qui produisent ces traces, sans pouvoir peser d'un quelconque poids sur leur formatage voire en dépendant totalement des conditions de fourniture de ces données, qui peuvent varier dans le temps et selon les plates-formes. Les puissants phénomènes de viralité propres à la plate-forme Facebook et à son mécanisme de likes, ne laissent

pourtant pas les chercheurs indifférents tant ils sont spectaculaires comme dans l'affaire du bijoutier de Nice (création de la page le 11 septembre 2013, 1 000 000 de likes le 14 septembre jusqu'à 1 635 000 likes le 7 Novembre). Cela suscite des analyses de tous types, des plus constructivistes critiques (« tous les likes sont achetés ») jusqu'aux plus réalistes (« c'est bien la preuve que l'opinion, voire même « les gens », a basculé dans le réflexe sécuritaire ». Les limites de la qualité des traces sont observables sur toutes les plates-formes mais ces limites peuvent être intrinsèques, lorsqu'elles ne répondent pas au critère de traçabilité que nous considérons comme décisif pour les exploiter, ou extrinsèques lorsqu'on critique leur absence de relation fiable avec le monde « réel ». C'est cette dernière posture que l'on trouve chez boyd et Crawford à propos de Twitter: « Some users have multiple accounts. Some accounts are used by multiple people. Some people never establish an account, and simply access Twitter via the web. Some accounts are 'bots' that produce automated content without involving a person. Furthermore, the notion of an 'active' account is problematic. While some users post content frequently through Twitter, others participate as 'listeners'. Twitter Inc. has revealed that 40 percent of active users sign in just to listen. » (boyd and Crawford, 2011) D'autres travaux (Driscoll and Walker, 2014) ont testé les données produites à partir de différentes méthodes d'accès offertes par Twitter par exemple et ont montré que l'API Search, l'API Streaming et le Gnip Power Track (service payant) fournissent des résultats très différents, la dernière méthode récoltant un bien plus grand nombre de tweets, mais non de façon uniforme selon les requêtes ! C'est dire que les traces collectées sont entièrement dépendantes des dispositifs de collecte, ce qui ne saurait étonner mais qu'on a tendance à oublier pour d'autres méthodes plus anciennes qui sont devenues conventionnelles.

L'emprise des marques sur les traces

D'où vient cette fascination pour les traces malgré leurs limites, par comparaison avec les données des registres et des enquêtes ? Les traces sont en fait une des ressources clés pour les marques pour suivre les effets de leurs propres actions sur leur public. La réputation, la notoriété, ne se traduisent plus seulement dans des mesures d'audience qui seraient une importation simpliste des mesures longuement construites pour les

mass média. Sur les réseaux, il faut mesurer à la fois une forme d'audience (le reach), des activités les plus élémentaires de ces publics incertains (likes, étoiles), mais aussi des activités plus élaborées, comme leurs commentaires, qui constituent ce qu'on appelle leur « taux d'engagement ». Les marques sont friandes de ces traces et ce sont elles qui alimentent le chiffre d'affaires de toutes ces plates-formes et par là de tout le web. Les outils d'opinion mining et de sentiment analysis (Boullier et Lohard, 2012) constituent ainsi la réponse à cette angoisse du marketeur après le lancement de produit. Cependant, l'extension de ce domaine de la marque atteint toutes les activités, qu'elles soient commerciales, culturelles, politiques, institutionnelles voire interindividuelles lorsque chacun doit mesurer son excellence à l'aide de rankings, comme le font les chercheurs (Bruno et Didier, 2013). Dès lors, ce sont les méthodes des marques qui prennent partout le dessus et imposent leur loi et leur rythme, jusque dans les services publics. Or, ce qui préoccupe avant tout ces marques ne sont pas des données structurées et construites pour tester des causalités par exemple, mais bien des traces, qui fonctionnent comme indices et alertes, même approximatifs, non pas au niveau individuel mais au niveau de tendances, de trends. De même, ce n'est pas la réflexivité qui est recherchée mais avant tout la réactivité, la capacité à déterminer sur quel levier agir en fonction des dimensions (features) de la marque qui sont affectées. Le monde politique lui-même est désormais pris dans cette spirale de la réactivité et son addiction aux tweets nous a conduits à considérer que nous étions entrés dans l'ère du High Frequency Politics (Boullier, 2013) à l'image du High Frequency Trading de la finance spéculative.

Nous avons ainsi dressé un tableau qui mérite d'être systématisé. Le numérique en réseaux génère

- des traces
- assemblées et formatées par des plates-formes
- pour des marques
- en vue d'une réactivité
- pour produire des rankings ou des patterns.

Cette situation n'est pas nouvelle. Deux autres moments clés de l'existence des sciences sociales et en particulier de la sociologie et des sciences politiques, doivent être mis en parallèle selon la

même méthode pour comprendre la portée des changements en cours.

La construction « de l'opinion »³

La situation contemporaine n'est sans doute pas si éloignée d'un moment clé dans l'histoire des sciences sociales qui nous aiderait à comprendre ce qui se passe. Si l'on donnait à l'époque actuelle des traces numériques le libellé de « 3G », pour troisième génération, il faudrait alors donner à l'émergence de l'opinion à la fin des années 30 le libellé de 2G. En 1936 en effet, Georges Gallup parvint à prédire l'élection de Roosevelt face à Landon avec une étude sur 50 000 personnes. Roper et Crossley avaient fait de même au même moment. Non seulement Gallup impressionna les médias et les décideurs mais il disqualifia radicalement les méthodes anciennes (straw polls), celles du Literary Digest fondée sur les réponses de 2 millions de personnes, en prédisant même leurs propres résultats erronés. Ce qu'il fondait ainsi dans ce geste spectaculaire, c'était la fiabilité du sondage et des méthodes d'enquête par échantillonnage, le sampling, qui certes perdait l'exhaustivité des enquêtes sur une population entière mais parvenait à des résultats corrects à condition de respecter des conditions de représentativité. Il échouera cependant en 1948 à prédire la victoire de Truman dont les électeurs se décidèrent dans les dix derniers jours. Les méthodes ainsi appliquées à la vie politique et à une épreuve grandeur nature aussi importante qu'une élection présidentielle avaient été testées auparavant sur les études de lectorat pour lesquelles Gallup avait rendu opérationnel l'échantillonnage stratifié. A vrai dire, ces méthodes avaient été déjà mises en place par le norvégien Kiaer en 1894. De même, ainsi que l'a montré Emmanuel Didier (Didier, 2009), les méthodes statistiques, dans le domaine agricole puis dans le domaine du chômage, au début des années 30 aux USA, vivaient une mutation profonde depuis les méthodes des correspondants vers l'échantillonnage aléatoire, à base d'approches probabilistes (Desrosières, 2001). Les méthodes des quotas, fondées sur le « choix judicieux » où l'on fait correspondre l'échantillon sélectionné avec certaines propriétés de la population telles que le recensement les a établies, étaient pourtant

3. Les travaux de Loïc Blondiaux (op. cit.) et de Joelle Zask (2000) développent cette histoire largement.

différentes de ces méthodes de sélection aléatoire stratifiée, et restaient même quelque peu méprisées par les statisticiens⁴ (cf. Stephan cité par E. Didier). Les données recueillies étaient aussi très différentes puisque les statisticiens des administrations agricoles ou du travail voulaient récupérer des « faits », mais étaient cependant obligés de les faire reposer sur des déclarations, et non sur des machines de mesure, même s'ils le tentèrent avec des « crops meters ». Mais l'opération de légitimation de l'échantillonnage en général réussit avant tout grâce aux performances de Gallup (1939), entièrement dédiés à d'autres mondes sociaux, ceux de « l'opinion publique », et non plus de « la société » qui restait la référence des statisticiens de l'Etat fédéral et de ses bureaux. C'est bien dans le contexte des médias de masse que leur importance fut reconnue. Avec Ogilvy en effet, Gallup étudia les audiences des films puis chez Young et Rubicam, avec Crossley, les audiences de la radio à partir d'interviews téléphoniques avant même de proposer ces sondages électoraux. Le nom de Gallup doit être de ce point de vue associé à celui de Lazarsfeld, qui, dans la même période, en 1936, lançait un « Radio Research Program », fondé sur ses travaux d'étude d'audience de la radio commencés en 1930⁵. Avec Merton, ils lancèrent les méthodes de focus groups dès 1941 et l'étude de Decatur en 1945 fournit les données pour l'analyse de « Personal Influence », publié en 1955 (Katz and Lazarsfeld, 1955), qui établit le cadre d'analyse du « two-step flow » dans lequel les médias de masse jouent un rôle, mais à travers les médiations des relations d'influence de divers types.

Les médiations qui font exister l'opinion publique sont constituées

Le lien entre les médias de masse et la vie politique est ainsi constitutif des nouvelles méthodes statistiques d'échantillonnage stratifié (certes fondées sur des quotas et non aléatoires). Ainsi que le note Alain Desrosières (op.cit.), la condition de prédictibilité d'une élection nationale dépendait en fait de la constitution d'un espace public médiatique commun à l'échelle des Etats-Unis, et seule la radio pouvait le faire de façon

4. Les quotas seront d'ailleurs abandonnés quasiment partout après la guerre sauf en France.

5. Les couples opérationnels/académiques ont été constitués plutôt de Gallup-Cantrill d'un côté et de Roper-Lazarsfeld de l'autre mais l'histoire en a retenu surtout Gallup et Lazarsfeld, voir Blondiaux sur ce sujet.

à rendre comparable l'état de connaissances des électeurs à propos des différents candidats. Une mutation médiatique considérable, les médias de masse (la radio à l'époque), a donc constitué les conditions d'émergence et de validation d'une technique d'enquête, qui ouvre ainsi toute une nouvelle époque, notamment pour les sciences politiques. Plus encore, c'est l'« opinion publique » elle-même qui prend une existence mesurable, par ces méthodes d'échantillonnage dont la puissance performative dépassera largement la phase expérimentale. Les instituts de sondage se répandront partout dans le monde, dont celui de Gallup (devenu nom commun pour sondage en Finlande par exemple) et celui de Stoetzel en France en 1938 (IFOP, institut français d'opinion publique, dès 1938, donc, même si l'activité réelle ne prendra que bien après la guerre), à tel point que les sciences sociales devront faire appel à eux pour constituer leurs échantillons jusqu'à récemment. Depuis quelques années, les laboratoires de recherche disposent enfin d'échantillons de populations disponibles pour leurs études sans avoir à passer par les instituts de sondage. Cette situation n'est pas sans rappeler celle que nous avons évoquée précédemment à propos du suivi des traces sur le web par les marques et les plates-formes et qui montre bien l'intrication étroite entre un système médiatique, des dispositifs de calcul innovants, et des intermédiaires générant une activité commerciale à partir de leurs savoir-faire.

Des marchés et des publics nationaux : les échelles des médias

Le maillon manquant dans toute notre description reste en effet le levier d'intéressement financier à de tels investissements pour connaître un public. Les agences de communication comme les instituts de sondage ne peuvent en effet vivre de leurs seules activités électorales quand bien même elles leur apportent une grande visibilité et notoriété. Leur cible est au départ constituée par les médias de masse, disions-nous, pour une raison essentielle : la mesure d'audience devient la clé de répartition des espaces publicitaires et cela dès l'origine avec la radio puis avec la télévision (en 1941, sont diffusées les premières publicités à la télévision américaine pour les montres Bulova pendant un match de baseball). Mais ces mesures permettent aussi de suivre les effets de ces campagnes publicitaires sur les esprits des consommateurs, donnant un essor sans précédent au marketing qui pilote des stratégies de communication

de plus en plus sophistiquées (Cochoy, 1999). Les marques sont ainsi présentes dès l'origine dans les méthodes d'enquête d'opinion par échantillonnage dès lors que ces enquêtes visent avant tout les audiences des médias de masse. Les études de marché des biens de consommation se développent à la même époque à partir des années 30 dans un même mouvement de standardisation nationale des produits, comme le rappelle Desrosières et notamment à travers le packaging qui crée la relation directe de la marque au consommateur final, au détriment des grossistes et des détaillants, comme le montre F. Cochoy (2002). La production d'un territoire national unifié, via les médias dont font partie les transports et la poste, constitue donc une nouvelle condition de félicité pour ces mesures par sondage. Cela nous permet de faire directement le parallèle avec la constitution récente d'un marché mondial, cette fois-ci, à travers la domination des plates-formes numériques. Google, Apple, Facebook et Amazon ont produit, avec l'aide des porte-conteneurs, le même effet d'échelle territoriale que la radio et le chemin de fer pour le territoire des marchés nationaux. Cette remarque s'inscrit ainsi dans le droit fil des travaux de Mc Luhan (1964) pour qui le changement d'échelle est constitutif en tant que tel d'un autre monde bien plus que les biens ou les contenus échangés.

L'opinion publique existe, je l'ai mesurée

Le travail réalisé par Gallup pour le côté opérationnel et Lazarsfeld pour le côté scientifique, n'est donc pas une simple opération marketing ou un lifting des sciences sociales : il fournit à des sociétés entières les méthodes pour s'auto-analyser, pour se représenter elles-mêmes comme opinions. Tarde (1989) avait certes mis en évidence l'importance de ces opinions, c'est seulement lorsque les métriques sont mises en place et produites de façon conventionnelles que l'opinion finit par exister. Et seules la commande des médias et leur capacité à produire de façon unifiée un public sur un territoire national permettaient de faire tenir ce montage méthodologique. Le « tout » dont parle les sondages, c'est en fait à l'origine le public constitué par les médias, qui permettent de faire émerger cette audience comme opinion publique, de la rendre visible et mesurable en permanence. Cette parenté entre mesures d'audience et méthodes de suivi de l'opinion publique, parenté technique et historique,

doit être considérée comme la clé du dispositif : les médias veulent avant tout mesurer des audiences, comme le fit Gallup pour la lecture, mais les techniques mises en place se transformèrent en outils prédictifs de votes, ce qui justifia le pari sur une opinion publique. Le tout « audience » voire « public » a ainsi muté en « opinion publique » et a pu se détacher de son auto-référence aux médias (qui se mesuraient eux-mêmes) au point d'être exploitable pour les marques pour mesurer l'influence de leurs campagnes. Les parties (Latour *et al.* 2012) que sont les expressions individuelles sont préformatées pour être enregistrables et calculables mais le lien entre parties et tout n'est réalisé que par les boîtes noires des instituts de sondage. Les précautions scientifiques de rigueur sont prises grâce aux intervalles de confiance (définis en 1934 par Neyman), qui permettent de garder une référence avec l'exhaustivité de la population étudiée. Bowley (1906) avait proposé ces principes dès 1906 en parlant de l'erreur probable, ce qui permet de relier clairement les sondages avec l'apparition des probabilités chez les statisticiens, comme le fait E. Didier. Mais rapidement, ces précautions disparaîtront des restitutions, comme on le voit dans les médias contemporains, et des intervalles de 1% dans des sondages sur un effectif de 1 000 personnes pour une population de 40 millions d'électeurs sont désormais considérés comme significatifs par les médias et les acteurs eux-mêmes alors que les principes de précaution sont pourtant fournis (ce qui entrainera une bonne partie des déceptions et dénonciations de l'incapacité des sondages à prédire les comportements réels). A cet instant, chacun sait que l'opinion existe, quelque soit le travail de compte-rendu des artefacts nécessaires pour la faire exister et quoi qu'en dise Bourdieu (1984)⁶. Elle a été naturalisée, « taken for granted » et les méthodes d'échantillonnage enfouies sous l'effet performatif puissant de ces indices agrégés et immédiats. L'approximation reste acceptable notamment parce que la répétition d'un même questionnaire dans le temps (par panel, par échantillon indépendant ou par échantillon tournant), dans des conditions identiques « toutes choses étant égales par ailleurs », permet

de lisser les biais, qui deviennent alors acceptables par convention.

Le travail de convention ainsi réussie porte sur les mêmes assemblages de médiations déjà évoqué pour les traces :

- des « surveys » et des « polls » (à partir d'expressions individuelles cadrées par des questions et ainsi rendues calculables)
- assemblés et formatés par des instituts de sondage
- sous garantie de représentativité d'échantillons (sampling)
- pour des médias
- en vue d'un monitoring
- pour produire de l'opinion publique (et des audiences)

Comme le dit Alain Desrosières (2008), l'essentiel n'est pas de savoir si ces données sont des reflets ou des miroirs de la société ou d'autre chose, mais de « faire quelque chose qui se tient ».

Notons qu'un élément nouveau intervient ainsi dans cette chaîne : celui de la contrainte méthodologique, exprimée en terme de représentativité des échantillons, car cet élément manque encore pour les traces numériques, ce qui explique en grande partie l'incertitude et la suspicion sur tous les résultats obtenus par comparaison avec les sondages, dont les biais sont bien connus mais contrôlés par convention depuis les années 40. La « consolidation », qu'Emmanuel Didier (*op. cit.*) décrit pour les statistiques et les sondages hors études d'opinion, reste à faire.

Ce retour un peu long sur la fabrication réussie de l'opinion était nécessaire pour comprendre à la fois les analogies entre cette époque et celle que nous vivons mais aussi pour mesurer la distance et le travail nécessaire pour produire des conventions de qualité équivalentes (Eymard-Duvernay, *op. cit.*) qui fassent exister les « traces » comme entités reconnues pour les sciences sociales. Il nous faut bien considérer l'opinion comme une réalité sociale qui vit sa vie et ne pose plus question grâce à la qualité des montages techniques et institutionnels qui ont stabilisé son mode d'apparition. Certes, les mondes des sciences sociales, dont les sciences politiques, et ceux du marketing diffèrent : pourtant, ils ont utilisé pendant des années les mêmes méthodes, voire les mêmes échantillons tout en étant capables de

6. Pp. 222-235. Précisant sa pensée à la fin de l'article, il indique : « L'opinion publique dans l'acception implicitement admise par ceux qui font des sondages d'opinion ou ceux qui en utilisent les résultats, je dis simplement que cette opinion-là n'existe pas ».

s'en distinguer. La question posée à ce nouveau monde des traces qui émerge sur le web est du même type : comment pouvons-nous inventer les sciences sociales qui leur correspondent tout en admettant les conditions de production et d'utilisation de ces traces ?

La fabrication de « la société »

Cette référence historique à l'opinion pourrait sembler trop voisine du monde numérique en réseaux en raison de l'implication des médias et des marques. De ce fait, le monde des traces produites sur le web pourrait finalement se limiter à une extension permanente du domaine des marques, sur d'autres métriques. Pourtant, il nous semble qu'un autre moment historique des sciences sociales nous permettrait de complexifier le panorama et de le percevoir dans la longue durée. Nous prétendons en effet ici que Durkheim a réussi une opération identique à celle de Gallup et de Lazarsfeld qui inventèrent « l'opinion publique », car il parvint à faire exister « la société ». Autant le caractère conventionnel de la notion d'opinion peut encore être admis, autant l'évidence de la société ne souffre pas discussion. D'autant que le terme ne date pas de Durkheim, même si son histoire n'est pas si longue. L'archéologie de la notion de société (Latour, 2005) pourrait encore être enrichie par l'appel aux travaux de Quételet (1846) produisant son « homme moyen » qui resta longtemps la clé de toute la statistique⁷. A la fin du XIX^{ème} siècle cependant et avec le coup de génie de Durkheim en grande partie, se produit un changement d'existence pour la notion de société. Les premiers travaux de Durkheim sur « la division du travail social » (1893) ne s'appuyaient pas sur une méthode statistique mais posaient les bases d'un modèle des types sociaux,

agregés en solidarités mécaniques et organiques. L'examen détaillé des régimes juridiques servait à la démonstration et de ce fait s'appuyait sur des ensembles constitués ou en cours de constitution que sont les systèmes juridiques, dans leur version traditionnelle ou plus moderne. Avec « Le Suicide » (1897), la méthode se met en place pour prolonger cette discussion des types qui va faire émerger l'anomie comme situation problématique. Mais l'appui sur les registres de données produites par les Etats, issus de ses diverses composantes (ministères, préfectures, administrations) devient clé dans la démonstration. Ce sont en effet ces agrégats qui sont expliqués ou explicatifs, grâce à une méthode de comparaison entre pays, entre régions, départements ou districts quand c'est possible et nécessaire. La méthode dépend entièrement des données disponibles et ne peut se payer le luxe de critiquer ou de mettre en doute les procédures de production de ces données, malgré les innombrables limites relevées dès la publication. Ainsi, le seul fait que le suicide soit une catégorie produite et enregistrée par les services administratifs devrait encourager une investigation de type ethnométhodologique sur les modes de catégorisation propre à chacune de ces administrations, sans considérer que la comparaison soit immédiatement possible entre pays. Mais ce constructivisme-là n'est pas de mise face au réalisme des approches statistiques et surtout il ignorerait l'effet de réalité produit par ces conventions. En organisant tout son dispositif de preuve autour de ces statistiques administratives nationales, Durkheim trouve un analogue quantitatif à son parti-pris conceptuel qui place « la société » dans un statut à part de toutes les manifestations et comportements individuels. Le tout de Durkheim devient une entité de second degré, « la société », alors que les recensements et autres registres de données des états ne font pourtant qu'un travail de récupération d'événements administratifs individuels (état-civil, procédures judiciaires, etc.), formatés dans des catégories identiques et agrégés pour faire apparaître des comportements de populations. Toute la force de conviction de Durkheim sera de faire exister ces populations statistiques comme équivalentes de sa « société ». L'appareil statistique rend visible cette société de la même façon que le sondage rendra visible l'opinion et dès lors, indépendamment de la validité statistique, le cadrage (framing) ainsi opéré gagne en puissance. Il faut en effet remarquer qu'une forme « d'alliance

7. Nous évitons ici de revenir sur l'archéologie des sciences sociales qui serait nécessaire et qui supposerait de faire leur place à Smith, Ricardo ou Walras, à Tocqueville et à Le Play, à Marx et à Auguste Comte, pour ne citer que quelques-uns de ces grands ancêtres. La capacité de Durkheim à produire la méthode pour faire tenir les faits sociaux et à les appuyer sur des statistiques constitue selon nous un saut, à la fois conceptuel et opérationnel, dans son couplage avec l'Etat. Sans doute, faudrait-il cependant s'attarder sur Le Play et sa vision du « tout » à travers ses monographies, et reprendre les travaux des anthropologues, au-delà des approches quantitatives ici mises en avant. Nous disposons aussi de ces éléments pour alimenter la discussion en montrant les trois générations de méthodes qualitatives mais ce serait entrer dans une présentation trop longue ici.

objective » se constitue entre les producteurs de données issus des administrations de l'état et les sciences sociales naissantes. Ensemble, ils vont produire l'entité « société » comme l'objet à suivre par l'Etat pour des raisons de gouvernement et à expliquer pour des raisons scientifiques. Le résultat tiendra dans une évidence partagée, « la société » existe, et les méthodes qui permettent de la faire exister n'ont pas lieu d'être interrogées puisqu'elles démontrent à la fois leur valeur scientifique et leur valeur opérationnelle, outil de preuve et outil de gouvernement comme le dit Desrosières (2014). Processus et alliances tout-à-fait identiques à celles que l'on rencontre entre les médias et les instituts de sondage qui s'entendent pour faire exister l'opinion et la rendre naturelle, taken for granted, après un long travail de montage de conventions.

Le temps des calculs et des machines à calculer

Dans le cas de Durkheim, il faut noter des voisinages historiques, qui ne valent pas causalité mais qui permettent de comprendre le gain de puissance de cette façon de faire exister la société. En effet, en 1890, Hollerith utilise sa machine (qu'il a inventée quelques années auparavant et pour laquelle il a déposé une demande de brevet en 1886) pour réaliser le recensement américain. Le Bureau of the Census n'avait pas réussi à finir de traiter le recensement précédent qui datait de 1880 lorsqu'il fallut déjà lancer le suivant. Un changement de technique était nécessaire et disponible. La machine de calcul mécanographique de Hollerith fit le travail et fut commercialisée pour les mêmes objectifs de recensement dans plusieurs pays dont la France. La compagnie de Hollerith sera transformée par Watson en IBM en 1926. On comprend mieux comment la puissance gagnée dans le dénombrement et dans la description des populations consolide le statut de l'Etat et lui offre des sources de renseignements supposés utiles à son gouvernement. La prétention à l'exhaustivité du comptage accomplit la promesse du concept de société : les dispositifs techniques de saisie du tout existent, ce sont les machines de Hollerith équipant les procédures de recensement. Il faut signaler à quel point ces investissements de forme (Thévenot, 1986) que sont les recensements finissent par faire événement dans une population donnée et par devenir incontestables comme une forme de rituel d'apparition du tout social à lui-même. Cependant, lorsque pour

étaler le traitement de ces données et obtenir des mises à jour plus fréquentes, comme cela était fait dans bien d'autres pays depuis longtemps, l'INSEE décida en 2004 de pratiquer par recensement partiel, en distribuant l'enquête dans tout le territoire de façon inégale selon les tailles de communes et opaque pour le grand public, les vertus de ce tout finirent par disparaître. Lorsque l'INSEE produit les données du tout, il apparaît clairement que ce sont des projections, des reconstitutions, et ce, dès le départ, sans pouvoir faire appel à un mythe de l'origine commune, où tous les individus sont captés ensemble et au même moment. Les conséquences administratives et financières de tels ajustements peuvent être très mal vécues par les maires (cf. Saint Malo en décembre 2013) puisque leur commune peut perdre des aides liées à un niveau de population dès lors qu'elle décline. La critique de la méthode est plus aisée puisqu'aucune prétention au tout originellement capté ne peut plus être faite par l'INSEE.

Alain Desrosières avait largement démontré ces processus en montrant comment le concept de société de Durkheim rendait compte de la genèse des Etats nations, qui reposent autant sur ces chiffres que sur des infrastructures. C'est ainsi que le territoire devenait une médiation clé, dont on trouve la trace dans l'émergence des politiques de protection sociale, dans le développement de marchés de biens nationaux grâce au chemin de fer, puis de campagnes électorales à travers des médias nationaux (la presse diffusée aussi par le chemin de fer, puis la radio, qui deviendra un vecteur d'émergence de l'opinion).

Les opérations réalisées par Durkheim ont eu des conséquences sur la production d'autres « tout » puisqu'il a insisté en permanence, et contre Tarde en particulier, pour séparer la sociologie de la psychologie. Celle-ci désormais travaille avec un autre tout, l'individu, et peut le mesurer à l'aide de dispositifs expérimentaux dès lors qu'elle devient psychologie cognitive. Mais l'opération de construction du tout « économie » avait été réalisée bien avant par les économistes et comme l'a montré la sociologie économique, l'économie-chose a bien été performée par l'économie-discipline, de façon à la séparer de tous les liens qui la faisait tenir aux autres processus. La valeur est devenue la clé de la mesure alors que la monnaie en est en fait le support (Orléan, 2011). L'anthropologie tenta de faire exister aussi le tout «

culture » mais il se déclina immédiatement en cultures au pluriel, en reprenant les distinctions entre tradition française et allemande, ce qui ne permit pas d'obtenir la force de frappe d'un concept pourtant fort utile à l'entreprise coloniale. Les sciences politiques ont, elles, adopté le tout « Etat » comme référence et la difficulté à faire admettre l'extension du politique en dehors des sphères institutionnelles indique bien que ce concept a gardé toute sa puissance définitoire de la discipline elle-même.

La performance de Durkheim aura ainsi été de faire tenir un assemblage de médiations fort puissant :

- des recensements
- assemblés et formatés par des administrations publiques
- sous garantie d'exhaustivité
- pour des Etats
- en vue d'un gouvernement
- pour produire de la « société » (à partir des populations)
- à l'aide de machines de calcul mécanographiques

Le pouvoir d'agir des dispositifs techniques de calcul

Nous introduisons ici la dimension technique des supports de calcul qui produisent les données, car ces capacités de calcul et leur augmentation jouent un rôle essentiel. Les machines IBM qui servent les grands calculs des états vont irriguer toutes les institutions pendant 80 ans, et pénétrer de plus en plus profondément dans l'équipement de tous les services administratifs centraux puis locaux.

Peut-on trouver pareille situation pour l'invention de l'opinion publique ? Au moment même où Gallup adapte le sampling pour les sondages d'opinion et en fait la démonstration lors de l'élection de 1936, Alan Turing (1936) écrit son fameux article qui constituera les fondations de toute l'informatique. Avec Von Neumann (1945) qui pensa quelques années plus tard l'architecture type de l'ordinateur, les conditions de développement de l'informatique et des calculs rapides émergent. Or, dans le cas de l'opinion publique, la perte de l'exhaustivité doit se compenser par un suivi plus fréquent et une réactivité plus

importante nécessaire pour les médias. Seules les capacités des ordinateurs, associés aux réseaux téléphoniques pour la transmission des données, permettent d'unifier et d'accélérer les calculs de ces échantillons représentatifs à une échelle nationale. La vitesse fut aussi décisive pour le suivi du chômage, comme le raconte E. Didier, puisqu'en 1937, le calcul fut réalisé en un mois seulement sur tout le territoire américain et montra la fiabilité des sondages face aux autodéclarations (3 M de chômeurs en plus).

Dans la même veine, on mesure dès lors la mutation actuelle en cours avec internet puis avec le web. La capacité de suivi des traces telle que Google l'a pensée et équipée en 1998 dépend entièrement d'une architecture technique du web inventée en 1990 par Berners-Lee et Cailiau. Dans ce cas, la dépendance technique est totale car il n'existe pas d'autres moyens de faire émerger ces liens entre sites, ces traces laissées par des clics et autres comportements des internautes. C'est aussi pour cela que le couplage entre marques/ réseaux techniques/traces est nettement plus fort que celui entre médias/ informatique/ opinion ou entre Etats/ calcul mécanographique/ société.

Ce que les sciences sociales peuvent faire du numérique, ce que le numérique fait aux sciences sociales

Replacer les mutations numériques dans cette longue histoire des sciences sociales permet de mieux comprendre les mouvements contemporains dans l'usage des traces. Trois postures peuvent se présenter :

- l'une qui tente de reprendre le cours des sciences sociales des générations précédentes pour appliquer leurs méthodes et leurs concepts de « société » et d'« opinion » aux traces du web,
- l'autre qui accepte ce nouveau monde des traces en s'immergeant dans ses exigences et ses principes en abandonnant les traditions et les impératifs scientifiques,
- la dernière qui s'affronte à la radicale nouveauté de cette configuration socio-technique et qui tente de comprendre quelle peut être la place des sciences sociales dans

la production de nouvelles conventions pour exploiter ces traces.

Les traditions de l'opinion et de la société équipées par le numérique

La première orientation consiste à reprendre les méthodes et concepts et à les appliquer aux traces recueillies sur le web, ou, de façon encore plus raisonnable, à exploiter les potentiels des réseaux numériques pour mettre en œuvre exactement les mêmes méthodes. Ainsi, les enquêtes par questionnaire et les sondages sont non seulement assistés par ordinateur, sur des plates-formes équipées pour cela, mais peuvent être entièrement réalisées sur internet et permettent parfois de récupérer des attributs socio-démographiques suffisants des répondants pour assurer une représentativité de l'échantillon. Ces sondages ou enquêtes en ligne sont parfaits pour rendre encore plus réactive l'opinion et la suivre dans ces états intermédiaires sur un rythme plus fréquent. Le numérique amplifie (Einsenstein, 1991) ainsi « la réalité de l'opinion ». De même, les modes de saisie des recensements peuvent être équipés de terminaux informatiques pour accélérer et standardiser encore le recueil des données, ce qui rend encore plus fiable « la société » ainsi rendue visible. Mais les web studies issues des sciences sociales mobilisent les mêmes cadres sur ces nouveaux supports médiatiques : études économiques de préférence à partir des requêtes Google, études de sociabilité, suivi longitudinal de « communautés », autour de thèmes, ou de sites particuliers, approches d'« opinion mining » et de « sentiment analysis » mobilisées pour augmenter le suivi de l'opinion publique ou le repérage de tendances.

L'approche la plus apte à se saisir des traces disponibles sur le web serait sans aucun doute l'analyse de réseaux. L'analyse sociale de réseaux (et non l'analyse de réseaux sociaux au sens du web 2.0) s'est souvent appuyée sur le suivi de la diffusion d'innovations (Coleman, 1990, entre médecins), d'informations (Granovetter, 1995, sur la recherche d'emploi) ou sur les traces des échanges permanents au sein d'organisations (Lazega, 2001). Le terme de « réseaux » permettrait de créer des passerelles avec l'analyse des graphes effectués sur les réseaux du web ou sur les plates-formes de réseaux sociaux. En réalité, il est plutôt trompeur. En effet, les réseaux des analystes de réseaux sont souvent des réseaux complets et l'on comprend bien à nouveau ici l'impératif d'exhaustivité. Dans

le cas de réseaux de diffusion, l'étendue des configurations étudiées est limitée, en partie pour des raisons de faisabilité mais il permet de prétendre à une généralisation par une forme de représentativité, la comparaison permettant de corriger les éventuels effets déformants d'un échantillonnage trop réduit. De plus, les concepts mobilisés dans ces approches doivent avant tout servir à approfondir des questions posées en termes de pouvoirs, d'inégalités, questions à caractère structural pour le tout qu'est la société. Il est ainsi plus aisé de comprendre pourquoi l'analyse de réseaux a si peu entrepris d'exploiter le web puisqu'elle exploitait ces concepts dans le cadre de sciences sociales de première génération.

Le même phénomène a pu se reproduire en partie avec la scientométrie dont l'analyse de réseaux constitue aussi une ressource mais qui a développé à côté de cela des approches spécifiques. La scientométrie (Courtial *et al.*, 1993) a l'avantage énorme de travailler sur un corpus de données normalisées par une communauté qui a effectué un travail considérable de discipline sur elle-même, à travers ses « disciplines » précisément, pour connecter toutes les références aux auteurs, aux articles, aux revues, aux institutions, aux mots-clés eux-mêmes. Cette tendance s'est encore accentuée avec la tendance contemporaine aux classements internationaux (rankings) devenus essentiels pour les carrières comme pour les financements des institutions et pour leur attractivité. Les méthodes mises en œuvre pour étudier les réseaux de citations comme les évolutions de thématiques par exemple bénéficient largement des capacités de calcul du numérique et des nouvelles méthodes de graphes. En revanche, elles ne peuvent en rien être transposées pour l'étude des traces car leurs impératifs d'exhaustivité, voire de représentativité sont aisément remplis dans un univers aussi normé que celui des publications scientifiques alors que le web n'offre aucun cadre pour se repérer, aucun tout de référence.

La fin de la théorie sociale ?

Chris Anderson du magazine *Wired* avait lancé un pavé dans la mare avec son court papier intitulé « The End of Theory » en 2008. Sa remarque se fondait sur une observation très juste des pratiques observées dans plusieurs domaines scientifiques (génomique ou physique) où l'usage massif des données à des fins de découverte de corrélations avait remplacé, selon lui, la nécessité

des modèles, des hypothèses et des épreuves construites pour les tester. Pour lui, le web constituait un autre monde qui ne décalquait rien mais mettait à disposition les outils pour capter les traces, les agréger, les calculer, les évaluer et les utiliser pour les modifier en retour, dans le même mouvement. Les likes de Facebook n'ont pas besoin de théorie : la plate-forme capte les traces des actions des internautes ou des machines qui cliquent, sous un format standardisé, elle les agrège et produit un score, qui est affiché et peut être utilisé par la plate-forme elle-même pour afficher des tendances qui permettent d'orienter les placements publicitaires des annonceurs qui eux aussi cherchent à obtenir des effets et à optimiser leurs choix de placements ou de communication. Dans un format simplifié, voilà la chaîne qui a été produite. La théorie sociale n'a quasiment aucune utilité dans un tel dispositif opérationnel où le mécanisme performatif fonctionne quasiment à l'identique de celui des mesures d'audience. Certains tentent alors de développer une critique montrant que les likes agrègent des comportements très différents, jusqu'à des likes achetés. Mais cela ne préoccupe guère les opérateurs des plates-formes ni les annonceurs. Leur action/réaction fonctionne sur le mode performatif, où les likes font émerger une réalité qui enclenchera des stratégies en vue d'agir sur les likes, dans un cycle auto-référentiel que l'on a pu attribuer aussi aux mesures d'audience. Cependant, dans le cas des mesures d'audience, tous les annonceurs et les programmeurs se sont entendus sur des critères stables, ont produit une convention partagée et son évidence a fini par s'imposer avec force tous les matins dans les directions de programmes des grands médias de masse. Les plates-formes de réseaux sociaux et les annonceurs n'ont pas encore stabilisé ce compromis, ce qui explique la prolifération des services qui prétendent devenir le standard, comme nous l'avons montré dans le cas de Klout, et qui veulent devenir le Nielsen ou le Médiamétrie de ces mesures (Boullier et Lohard, 2013). Mais pour tous ces instituts, pas de théorie nécessaire, si ce n'est un haut niveau de qualité statistique, car toute avancée théorique sur l'audience et ses processus supposerait une renégociation des accords, ce qui reste le seul critère de validité de toute théorie. Il est aisé de constater l'écart entre ces principes et les traditions des sciences humaines et sociales comme le fait G. Bowker et de montrer leur réductionnisme furieux : "If I am defined by my clicks and purchases and so

forth, I get represented largely as a person with no qualities other than "consumer with tastes." However, creating a system that locks me into my tastes reduces me significantly. Individuals are not stable categories—things and people are not identical with themselves over time. (This is argued in formal logic in the discipline of mereology and in psychiatry by, say, ethnopsychiatry.) The unexamined term the "individual" is what structures the database and significantly excludes temporality." (Bowker, *art. cit.*)

Bowker a raison de s'inquiéter du point de vue de la société, mais les sciences sociales de troisième génération ne s'intéressent pas tant à « la société » qu'à d'autres processus créés par d'autres dispositifs et qui pourtant nous font agir. Les marques, les réputations et les recommandations telles qu'elles sont exploitées par Amazon peuvent, certes, être réinjectées de force dans une matrice « société » pour leur faire dire ce qu'elles ne sont pas faites pour dire. Mais elles disent aussi quelque chose d'elles-mêmes, d'un autre monde, celui du pouvoir des recommandations et des contagions que les sciences sociales rechignent à entendre. Tout se passe comme si la sociologie de « la société » revivait une expérience analogue à celle provoquée par l'anthropologie, celle du nécessaire décentrement par rapport au monde moderne et à ses catégories. Certes, Durkheim s'en est d'abord servi pour analyser les religions (Durkheim, 1912) et mettre les sociétés traditionnelles et le totémisme au service de sa démonstration de la force de la société sur les individus. Mais Mauss (1950) fit un grand pas de côté en reconnaissant le pouvoir des choses et de l'esprit qui persistait en leur sein, le mana, ce dont Lévi-Strauss lui fit reproche. Un nouveau pas de côté fut fait par P. Descola (2005) qui oblige à situer les sciences en général dans une ontologie (naturaliste) parmi plusieurs disponibles dans le monde. Avec Viveiros de Castro (2009), le défi est encore poussé plus loin puisque l'anthropologie adopte un perspectivisme qui met en jeu ses propres fondations dès lors qu'elle accepte d'être affectée par le monde qu'elle observe, à savoir des collectifs qui constituent des assemblages bien éloignés des principes de la société de Durkheim.

Prenons le risque de nous comporter de la même façon que ces anthropologues avec le monde du big data, avec l'internet des objets, avec les corrélatifs qui font proliférer les liens alors que la

science voudrait les épurer et les réduire à des causes. Le monde qui naît est aussi différent du monde moderne et des sciences sociales que ne l'étaient les Achuar de Descola, si l'on accepte comme lui de ne pas projeter toutes nos catégories mais de penser de l'intérieur et de « devenir avec » (Haraway, 2003). Alors, il devient inutile de se plaindre de l'imperfection des données et de leur approximation puisqu'on a désormais affaire à des traces, à un processus de traçabilité généralisée. Cette traçabilité connecte des entités qui n'avaient pas d'existence auparavant, mais qui sont désormais dotées d'une adresse IP (grâce à IPv6, soit $3,4 \times 10^{38}$ adresses disponibles) et qui peuvent donc échanger tout comme les humains via leurs machines. L'état-civil qui constitue la base de référence des sciences sociales de 3ème génération n'est plus le recensement mais un index des adresses IP, totalement agnostique sur les entités qui « sont derrière », car toutes agissent de façon quasi équivalentes et font agir les autres. Ce décentrement peut sembler radical, mais il permet de faire cohabiter les approches des précédentes générations tout en observant le monde déjà là avec les outils et les catégories pertinentes.

Les propriétés des sciences sociales de troisième génération

Il reste en effet à observer, voire à imaginer ce que pourrait être la troisième génération de sciences sociales qui assumerait le caractère radicalement nouveau de ces traces hétérogènes, qui ne les rabattrait pas sur un statut de traces ou de symptômes d'un vrai social (la société ou l'opinion) et qui pour autant ne se laisserait pas happer par le système autoréférentiel de production/suivi des traces qui se dispense de théorie car il a d'autres visées. Nous présenterons d'abord les propriétés générales d'un tel programme, à la façon d'un cahier des charges, avant de revenir plus en détail sur les médiations jusqu'ici mobilisées et sur les choix qui s'imposent dans ce domaine pour faire tenir ces approches.

Les tendances du Big Data peuvent fournir des premières pistes qui méritent d'être confrontées à celles des sciences sociales précédentes. Ainsi les critères de qualité du Big Data sont souvent résumés aux 3V : Volume, Variété, Vélocité. La parenté avec les exigences des sciences sociales est assez frappante.

Volume et exhaustivité

Le volume correspond à l'exigence d'exhaustivité traduite sous un mode quelque peu limité, puisque personne ni rien ne permet de définir les frontières des univers de données rassemblées. Dès lors, il conviendra de fixer un équivalent de ce volume qui se rapproche des exigences traditionnelles de l'exhaustivité, sans pour autant pouvoir les respecter lorsqu'on traite du web. Nous devons clairement faire notre deuil de l'exhaustivité mais cela ne dispense pas de fixer les cadres conventionnels de toute démarche en sciences sociales traitant de traces numériques.

Variété et représentativité

Le second critère, la variété, est lui aussi une forme de transcription de l'exigence de représentativité qui a permis à toutes les sciences sociales de procéder par enquêtes, par sondages, à base d'échantillonnage. Là encore, le critère est une version lâche de la représentativité, qui suppose que l'on accepte un niveau suffisant de variété. Pour les sciences sociales de troisième génération qui acceptent de perdre la contrainte de représentativité telle qu'elle a été construite dans le cas des sondages, il reste à définir ce que serait cette variété. La constitution d'un ensemble de sources (sourcing) lors d'études du web par exemple devrait alors répondre à quelques critères propres aux méthodes numériques et au domaine étudié. Nos travaux sur l'opinion mining nous ont conduit à considérer qu'aucune description du social-société, du social-opinion ou du social-traces ne peut être produite en généralité sur les réseaux numériques. La prolifération des traces rend paradoxalement impossible toute prétention à une référence à un tout posé a priori ou constitué a posteriori. Les sciences sociales doivent accepter de ne traiter que des « issues » (Marres, 2007 ; Marres and Weltewrede, 2013), ou des points de focalisation d'attention, ou des « engagements situationnels finalisés » (Hannerz, 1983) dont le numérique peut garder les traces, des traces qui seront spécifiques à chaque issue ou à chaque engagement. Cela réduit considérablement la portée totalisante des prétentions du Big Data mais cela rend possible un certain équivalent de représentativité et d'exhaustivité.

Vélocité et traçabilité

Enfin, le dernier critère, la vélocité ne trouve guère d'équivalent dans les sciences sociales de première et deuxième génération. A vrai dire, ces processus

dynamiques ne constituaient pas leur point fort ni leur préoccupation. Il est en effet essentiel de trouver avant tout à représenter les positions à un instant t , pour montrer la force d'imposition de « la société » sur la diversité des comportements individuels ou pour montrer comment l'opinion publique se structure au-delà des expressions singulières obtenues dans les enquêtes. Certes, à travers un suivi longitudinal des mêmes populations ou la reprise des mêmes questionnaires, il est possible de restituer un équivalent d'une dynamique, sans jamais cependant pouvoir suivre à la trace les médiations qui permettraient de produire ces évolutions. La vitesse semble donc hors du champ des approches plus classiques.⁸

Cependant, une branche des sciences du web s'est, elle aussi, emparée de cette question de la vitesse à sa façon en exploitant les traces des mêmes qui se propagent sur le web. Il est très significatif que Kleinberg, celui-là même qui avait exporté les méthodes de la scientométrie vers l'étude de la topologie du web, méthodes qui furent reprises par Google, s'est intéressé depuis plusieurs années⁹ (2002) à la mise au point d'un « meme tracker » avec Leskovec (Leskovec and al., 2009). Leur étude la plus fameuse a porté sur la propagation des citations durant la campagne Obama, ce qui leur permit de réaliser une visualisation spectaculaire de la focalisation de l'attention en courbes à montées et descentes très rapides (streams and cascades) autour de certains incidents de la campagne¹⁰. Leur méthode agrège tous les types de traces que peuvent laisser ces citations, traitées comme des chaînes de caractères dont on peut trouver la trace dans tout le web, et en produit une métrique ancrée dans le temps, au jour le jour, voire minute par minute désormais avec Twitter (l'unité de mesure étant devenue le Tweet per Second). La prise en compte des mêmes nous paraît prometteuse sous réserve que l'on suive aussi les transformations-traductions de

ces mêmes (ou la dérive mémétique¹¹, Dawkins, 1976) dans des milieux différents, car « toute existence va différenciant », comme le disait Tarde (1895).

Dès lors, il devient possible de trouver un équivalent de la vitesse du Big Data : la traçabilité. Elle devient le critère essentiel de qualité des entités que l'on peut étudier. Quelques conditions de félicité doivent être réunies pour y parvenir :

- a/ Les traces en question doivent avoir une continuité suffisante pour qu'il soit encore possible de dire qu'il s'agit d'un même processus, sans avoir la contrainte de la première mémétique (identique) ni le laxisme d'une intertextualité généralisée, où tout pourrait être signe et reprise de tout.
- b/ Les traces en question doivent permettre des suivis d'associations hétérogènes, c'est-à-dire une puissance de connectivité suffisante. Pour cette raison, des traces dont le format est trop spécifique à une plateforme peu connue ne peuvent donner lieu à extension ni à suivi.
- c/ Le suivi des traces en question doit permettre de dater avec précision tous les événements, toutes les transformations et toutes les associations. Les timelines sont ici l'équivalent d'autres conventions, comme les points cardinaux pour la topographie ou les inégalités de richesse pour les sciences sociales de première génération.

Ces conditions rendent possible le basculement des sciences sociales vers le suivi d'éléments qui ne sont plus ni des individus ni des groupes, ni la société, ni l'opinion. Le numérique a dissous toute certitude sur leurs statuts, qui avaient été déjà remis en cause. La propagation de ces éléments, qu'il nous faudra qualifier au-delà des traces, devient l'objet d'étude de la troisième génération de sciences sociales car ce sont les propriétés de ces entités qui leur permettent de créer de petites différences et par là de circuler et d'affecter les individus et les groupes, la société et l'opinion. Mais aucune structure ou loi de propagation n'est à rechercher comme voudrait le faire la physique sociale (Pentland, 2014) car chaque milieu (plate-forme, problème, format, conjoncture) est spécifique.

8. Nous ne mentionnons pas ici les approches de la physique sociale, de l'éthologie ou de l'épidémiologie sociale qui ont produit des modèles sociaux sans référence aux traditions des sciences sociales.

9. J. Kleinberg part ici d'une détection de « topics » qui provoquent des « bursts » dans des flux d'emails, son objectif étant d'aider à les structurer et non de les suivre en tant que tels. Mais il fait le lien avec les problèmes du temps dans la narration tels que traités par Genette.

10. Notons la parenté des streams avec les « stream of thought » et « stream of consciousness » de W. James dans ses *Principes de Psychologie*, pour qui « ça pense ». James, 1892).

11. La piste des mêmes est intéressante sur le plan méthodologique, sous certaines conditions et sans obligation d'accepter l'idéologie qui accompagne le discours de Dawkins.

Mais ces conditions ne sont pas nécessairement considérées comme essentielles de la part des plates-formes (GAFA en résumé) et des marques qui font vivre tout cet écosystème par l'exploitation publicitaire de cette traçabilité. De même que la sociologie durkheimienne s'est associée de fait avec les institutions étatiques productrices de données pour produire « la société » en combinant de fait enquêtes (des statisticiens) et registres (des administrations) (Desrosières, *op. cit.*), les instituts de sondage de Gallup et de Lazarsfeld se sont associés aux médias grands consommateurs de données sur les publics pour produire leur « opinion publique », en rapprochant ainsi « mesures d'audience » (le public des médias) et « opinion publique » (le public de la sphère publique au sens politique (Lippmann, 1925) puis « tendances de consommation » pour le marketing. Les sciences sociales de troisième génération ne pourront guère faire autrement que de s'associer à ces plates-formes et à ces marques pour produire la science des traces qui seraient alors traitées comme des « vibrations », comme nous le proposons plus loin. Les traces produites étant dépendantes des plates-formes, on ne peut guère espérer les modifier à la source. Il est en revanche possible d'exploiter les traces produites par les plates-formes en les détournant de l'usage pour lequel elles avaient été conçues (*repurposing*, Rogers, 2013). La règle veut ici qu'on ne prenne en compte aucune explication d'un autre niveau ou d'un autre monde, mais que l'on puisse comparer des vitesses de propagation, des rythmes, des transformations éventuelles (par exemple par contamination sur d'autres domaines, etc.). La différence devra se faire sur la capacité à faire voir des processus qui n'avaient jamais encore été mis en évidence, en raison soit des limites techniques pré-numériques, soit des cibles adoptées par les générations précédentes de sciences sociales.

Cependant, l'ère des traces ne fait que commencer et les plates-formes ne sont pas et ne seront pas les seuls fournisseurs de traces en masse. L'internet des objets n'est plus un fantasme d'ingénieur, et la vie ordinaire commence à se peupler d'échanges sans contact, de puces RFID et d'autres géolocalisations automatiques qui dépendent non plus des personnes mais des objets eux-mêmes (cf. l'interobjectivité de Latour, 1994). Leurs traces durant leur parcours, leur état (activé ou non) permettent de piloter des processus logistiques, transactionnels, qui sont souvent confinés aux mondes professionnels concernés. Cependant leur extension

et leur accès ouvert seront quasiment inévitables dès lors qu'on s'engagera dans une prolifération telle qu'elle est annoncée. Il ne sera plus possible de renvoyer à des personnes, à des entités sociales au sens des sciences sociales de première et deuxième génération. Il n'y a pas de raison pour que les sciences sociales ne s'emparent pas de ces nouvelles sources, lorsque l'on admet que les appuis techniques distribués d'une société, d'une opinion, peuvent prendre leur autonomie, comme dans le cas extrême de la robotique avancée. Dès lors que l'on admet l'abandon de la référence à des individus, à des sujets ou à des institutions, qui sont les présupposés de toute science sociale de la société ou de l'opinion et que l'on prend comme point de départ ces éléments, ces mêmes, ces vibrations qui se propagent et que l'on rend compte de leur pouvoir de circulation et de traduction, il devient possible de traiter tout le Big Data sur le même plan, en associant les traces produites par les plates-formes de réseaux sociaux comme celle des bases de données, celles générées par les capteurs présent dans tout objet comme celles des verbatims dans les corpus des forums sur le web.

Des traces aux répliques

Pour rendre plus solides les fondations de ces sciences sociales de troisième génération, il convient de donner statut scientifique à ces traces. Or, en préalable, il faut rappeler qu'il est fort probable que toutes ces traces qui pouvaient encore être connectées à des données personnelles ne seront plus accessibles dans les mêmes conditions dans quelques années. Le succès d'AdBlock Plus qui bloque les cookies et autres publicités intrusives s'amplifie constamment (200 M téléchargements, 40% d'installation sur Firefox). Le cryptage généralisé deviendra une nécessité face à l'incapacité des plates-formes et des services de renseignement à réguler leurs propres activités prédatrices de données personnelles. C'est pourquoi la prise en compte des traces, à la surface même des réseaux, et sans lien avec les données structurées et socio-démographiquement significatives, constitue une base solide de fondation des sciences sociales, au contraire de tous ceux qui continuent de vouloir appliquer les modèles de la société et de l'opinion à un univers qui ne leur est accessible qu'en raison d'un laxisme très provisoire. Travailler à la surface de ces traces, sans lien avec les données personnelles, permet aussi de réduire les contradictions éthiques dans

lesquelles se trouvent prises les sciences sociales de la société et de l'opinion qui veulent exploiter ces sources.

Nous l'avons souligné, la production des traces est directement dépendante de plates-formes qui génèrent elles-mêmes leurs analyses. Il est cependant nécessaire de fonder les sciences sociales de 3ème génération sur une proposition non captive des utilisations qui sont faites de ces traces, de la même façon que les sondages ne servent pas qu'aux médias ou que les recensements ne servent pas qu'aux états, comme nous l'avons indiqué précédemment en suivant les analyses de Desrosières. Aux couples registre/ enquête, puis audience/ sondages d'opinion, il faut parvenir à ajouter un couple traces/ X, X étant la place qui reste à définir pour la reprise des traces par les sciences sociales.

Nous proposons de parler alors de « répliques ». Le terme présente une parenté avec les mêmes (la mémétique considère le même comme un répliqueur) car les traces nous intéressent pour suivre des répliques (replicas), des imitations au sens tardien (et donc des oppositions et des adaptations). Il est aussi apparenté au monde de l'échange langagier pour désigner des réparties dans un dialogue (replies), dans une conversation, qui aurait dû être au centre des sciences sociales selon Tarde (*op.cit.*, 1901). Il permet ensuite de filer une métaphore suggestive avec les répliques des tremblements de terre (aftershock), sachant qu'il est possible de suivre sur le sismographe des répliques qui anticipent et d'autres qui suivent le choc lui-même (cf. P. Levy¹²). Sa traduction en anglais est cependant difficile et nous proposons de retenir pour la traduction le terme de « vibrations », qui reste proche de cet univers sémantique, qui a l'avantage de focaliser l'attention sur les ondes et moins sur les particules et qui fait écho au buzz qui obsède les marques et les médias mais qui n'est jamais théorisé.

D'autres concepts pouvaient prétendre à cette place, comme l'attention, l'influence, les « issues », l'acteur-réseau, les mêmes ou encore les conversations (Boullier, 2004). Cependant, le terme de « répliques » présente l'avantage d'être familier et polysémique et chacune des acceptations fait sens dans cette science qui traite les traces comme matière première. L'essentiel tient

dans le décentrement réalisé vis-à-vis des notions d'acteurs, de stratégies et de représentations, qui ont toutes leur légitimité dans le cadre des autres sciences sociales mais qui ne permettent pas de rendre compte du pouvoir d'agir des entités circulantes que sont les répliques. Nous ne pouvons pas dire a priori quelle est la taille ni le statut de ces entités, car ce sont seulement les investigations de corpus de masse qui peuvent nous les faire repérer dès lors que leur réplique émerge des capteurs que nous exploitons, certes à partir des plates-formes mais selon nos objectifs.

Le principe d'une sociologie des répliques repose sur l'impératif de suivre des éléments, sans pour autant savoir comment ils vont s'agréger pour faire des « tout » à géométrie variable. Le parti pris est donc « élémentariste » mais ne doit surtout pas devenir atomiste (car la géométrie variable reste une qualité que nous avons apprise de l'ANT) ni tendre vers la physique sociale, qui cherche des lois supposées transversales à tous ces flux. L'approche par les répliques permet de construire une combinatoire infinie, en suivant les extensions, les propagations, les répétitions, à la condition de rester centrée sur les « issues » que portent et font vivre les répliques, en ce sens bien différentes des traces « brutes ». L'objet d'étude n'est pas tant l'élément, qui peut avoir des attributs très variés, en étendue et en matérialité, ni seulement les agrégats, ce que l'on tend à faire avec les clusterisations des méthodes de graphes (Fouetillou, 2008), mais bien le processus de circulation et d'agrégation ou de désagrégation, au moment de bifurcation des courbes. Dans ces courbes, il faut alors plutôt se focaliser sur les moments d'émergence et non sur les pics qui fonctionnent, eux, comme des agrégats, comme le fait le memetracker de première génération. L'objet de cette science des répliques est bien l'agentivité des répliques qui se propagent et qui finissent par nous prendre. Car les individus sont en fait traversés par les idées et les idées nous agissent et non l'inverse comme l'avait bien indiqué Tarde (1893).

« Les rayons d'imitation d'abord et ensuite des êtres dont on induit l'existence à partir de la variation qu'ils font subir aux flux d'imitation » (Latour, 2011). Il est alors possible d'étudier les propriétés de ces répliques pour comparer éventuellement leurs chances de survie ou de contamination rendues possibles par ces différences de propriétés toujours directement liées aux «

12. « Nous sommes une « société de sismographes qui s'entrecaptent » . (Levy, P., 2000, p.46)

issues » qu'elles portent avec elles. Comme on le voit, l'approche par les répliques est alors une entrée vers une monadologie (qui se différencie radicalement d'une vision atomiste), ou vers une écologie, comme l'avait proposé Deleuze.

La traçabilité n'est cependant pas donnée telle quelle par les plates-formes et nécessite de produire les outils et les méthodes adaptées à une approche des répliques et non plus seulement des traces. Leskovec et Kleinberg ont fait figure de précurseurs sur ce plan en proposant leur memetracker. Ils sont en effet capables de restituer des flux de tous types, ce qu'ils appellent des streams et des cascades, ou des répliques. Le développement des méthodes de traçabilité des répliques devra tenir compte de cet existant, en veillant à les tester préalablement sur des corpora constitués à cette fin, quitte à en perdre le « réalisme ». Nous avons commencé ce travail en 1987 à partir du suivi des conversations télé transformées sur les lieux de travail pour en faire des opinions publiques locales (Boullier, 2004). Nous avons procédé de même pour le suivi des attributs d'une photo dans les bases de données Flickr (Boullier et Crépel, 2013) et des signes transposables dans un corpus de sites web en lien avec une région (Le Béhec et Boullier, 2014). Dans le premier cas, ce sont les attributs de la photo (ex : les bras croisés) qui deviennent des attracteurs de tags et qui connectent ainsi des comptes ou des photos qui n'auraient jamais été reliés a priori selon ces critères. Mais nous nous sommes arrêtés au principe de ce travail sans pouvoir le conduire empiriquement à une échelle suffisante. Dans le second cas, la propagation du drapeau breton sur le web devient un indicateur de connexion que l'on peut comparer avec d'autres entités constitutives de l'imagerie régionale qui, elles, ne parviennent pas à se propager. A partir d'un travail manuel sur près de 600 sites réalisé par M. Le Béhec, il fut possible d'esquisser ce que serait l'analyse des répliques générées par ce drapeau et de constater que ses propriétés sémiotiques n'étaient pas étrangères à sa capacité de circulation. Tags ou icônes sont ainsi des répliques que l'on peut suivre, quand bien même ils n'ont pas le caractère explicite de verbatims ou d'expressions comme dans le memetracker ni leur caractère massif. Potentiellement, toutes les traces que nous avons identifiées (telles que les likes, les tweets, les recommandations, etc.) sont susceptibles de faire l'objet de ces suivis : ils nécessitent cependant des outils de traçabilité spécifiques, qui existent

en grande partie pour Twitter seulement mais à la condition de faire un réexamen détaillé de ces outils pour vérifier qu'ils répondent au cahier des charges d'une traçabilité des répliques (et non seulement des traces pour elles-mêmes ou pour la réactivité des marques).

Conclusion

Un tableau synthétique des trois âges des sciences sociales permet de rendre perceptible la cohérence de l'approche mais oblige dans le même temps à schématiser et à éliminer des spécificités propres à chaque âge. Rappelons de plus que nous ne traitons pas ici des versants dits qualitatifs des méthodes des sciences sociales.

Tableau des trois générations de sciences sociales

[Voir le tableau page suivante]

La cohérence toujours abusive du tableau ne doit pas faire oublier que ce qui est en jeu est la construction d'une offre de sciences sociales de troisième génération qui n'est pas garantie. Certes la théorie de l'acteur-réseau en a posé les prémisses et Tarde en avait annoncé les principes, mais pour l'instant, la tendance à la fin de la théorie et l'occupation du terrain par les plates-formes du web (GAFA) qui elles-mêmes produisent, calculent, et publient sur ces traces, restent dominantes, et cela pour des visées commerciales avant tout, puisque les marques sont les grands demandeurs de ces approches. Rappelons que nous reconnaissons dans le même temps l'intérêt pour les marques d'apprendre à réagir en utilisant ces métriques. De même que nous reconnaissons l'intérêt pour les sciences sociales de la société et de l'opinion de continuer à développer leurs approches en utilisant ces sources. En ce sens, nous plaçons pour faire coexister ces approches, pour apprendre à changer de point de vue et pour admettre les conditions de possibilité de chaque génération, appuyées sur les Etats, les médias ou les marques. Chaque étude spécifique d'une question issue de l'expérience ordinaire ou posée par ces prescripteurs doit conduire à combiner les trois générations. A la condition que la recherche dispose d'un cadre spécifique pour ces traces qui envahissent notre monde. Notre intention est seulement de contribuer à poser les bases d'une convention permettant de faire émerger une théorie sociale et un objet, les répliques, qui ne rabattent pas le numérique sur les « méthodes

Tableau des trois générations de sciences sociales

	1 ^{ère} génération	2 nd e génération	3 ^{ème} génération
Concept du social	Société(s)	Opinion(s)	Réplique(s)
Dispositifs de collecte	Recensement	Sondage	Plateformes/Big Data
Principe de validation	Exhaustivité	Représentativité	Traçabilité
Co-construction institutions/recherche	Registre/enquête	Audience/sondage	Traces/méthodes numériques réorientées
Acteurs majeurs de référence (et financeurs)	Etats	Mass media	Marques
Acteurs opérationnels du calcul	Instituts nationaux	Instituts de sondage	Plates-formes du web (GAFA)
Auteurs fondateurs	Durkheim	Gallup Lazarsfeld	Callon Latour Law
Problèmes clés des approches scientifiques	Division du travail et État providence	Propagande et influence des médias (mesures d'audience)	Science et technologie (scientométrie)
Conjoncture technique	Machines de Hollerith (calcul mécanographique)	Informatique (Turing et Von Neuman)	Internet, web et Big Data
Formats sémiotiques	Tableaux croisés et cartes topographiques	Courbes et histogrammes/diagrammes circulaires (camemberts)	Graphes et timelines
Métriques	Statistique	Sampling	Topologie (Scores)
Critères techniques de qualité des données	Pertinence, précision, actualité, accessibilité, comparabilité, cohérence	Intervalle de confiance, probabilités	Volume, variété et vitesse (Big Data)
Modalités dominantes de compte-rendu	Explications	Corrélations descriptives puis prédictives	Corrélations prédictives

numériques » ni sur les « humanités numériques ». Il existe une nouvelle matière première qui mérite un examen pour elle-même et qui produit une troisième couche du social, mesurable selon d'autres principes, et non réductible à la société ou à l'opinion. La société a fini par exister, l'opinion a fini par exister, les répliques doivent finir par exister au même titre.

Références

- BLONDIAUX L., 1998, *La fabrique de l'opinion. Une histoire sociale des sondages*, Paris, Seuil. 1998. Plusieurs remarques prennent appui sur ce travail fondamental pour notre réflexion.
- BOULLIER D. et LOHARD A., 2012, *Opinion mining et sentiment analysis. Méthodes et outils*, Paris, Open Editions Press.
- BOULLIER D., 2013, « Plates-formes de réseaux sociaux et répertoires d'action collective » in NAJAR S. (ed.), *Les réseaux sociaux sur internet à l'heure des transitions démocratiques*, Paris, Editions Karthala, 492 p.
- BOULLIER D. ET LOHARD, A., 2013, « Médiologie des réputations », *Journées d'étude Association Française de Sociologie: Vers une sociologie des réputations ?*, Amiens.
- BOULLIER D., 2004a, *La télévision telle qu'on la parle. Trois études ethnométhodologiques*, Paris, L'Harmattan, 2004.
- BOULLIER D., 2004b « La fabrique de l'opinion publique dans les conversations télé », *Réseaux*, n°126, 57-87.
- BOULLIER, D. ET CREPEL, M., 2013, « Biographie d'une photo numérique et pouvoir des tags : classer/circuler », *Revue d'Anthropologie des Connaissances*, 2013/4 (vol.7, n°4) pp 785-813.
- BOURDIEU P., 1984, « L'opinion publique n'existe pas » in *Questions de sociologie*, Paris, Les Éditions de Minuit, pp. 222-235.
- BOWKER G. & LEIGH STAR S., 1999, *Sorting Things Out: Classification and its consequences*. Cambridge, MIT Press.
- BOWKER G., 2014, « The Theory/Data Thing. Commentary », *International Journal of Communication* 8, 1795-1799
- BOWLEY A.L., 1906, "Address to the Economic Science and Statistics Section of the British Association for the Advancement of Science", York, *Journal of the Royal Statistical Society*, Vol. 69, No. 3, pp. 540-558.
- BOYD D., & CRAWFORD K., 2011, "Six provocations for big data", Paper presented at the Oxford Internet Institute's *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, University of Oxford, Oxford, UK.
- BRUNO I. et DIDIER E., 2013, *Benchmarking. L'Etat sous pression statistique*, Paris, La Découverte, coll. Zones.
- CARDON D., 2013, « Du lien au like sur internet. Deux mesures de la réputation », *Communications*, 93, « La réputation », p. 173-186.
- COCHOY F., 1999, *Une histoire du marketing. Discipliner l'économie de marché*, Paris, La Découverte.
- COCHOY F., 2002, *Une sociologie du packaging ou l'âne de Buridan face au marché*, Paris, PUF, 2002.
- COLEMAN, J.S., 1990, *The Foundations of Social Theory*, Cambridge (Mass., É.-U.), Harvard University Press.
- COURTIAL JP, CALLON M. ET PENAN, H., 1993, *La scientométrie*, Paris, Presses universitaires de France.
- DAWKINS R., 1976, *The Selfish Gene*, Oxford, Oxford University Press.
- DESCOLA, P., 2005, *Par-delà nature et culture*, Paris, Gallimard.
- DESROSIERES A., 1993, *La politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte.
- DESROSIERES A., 2001, « Histoire de la raison statistique : le moment bayésien », *Courrier des Statistiques*, n° 100, décembre 2001.
- DESROSIERES A., 2008, *Gouverner par les nombres : L'Argument statistique II*, Paris, Presses de l'École des Mines.
- DESROSIERES A., 2014, *Prouver et gouverner : Une analyse politique des statistiques publiques*, La Découverte, 284 p. (Recueil posthume de textes choisis et rassemblés par Emmanuel Didier)
- DIDIER E., 2009, *En quoi consiste l'Amérique ? Les statistiques, le New Deal et la démocratie*, Paris, La Découverte.
- DRISCOLL K. AND S. WALKER, 2014, "Working Within a Black Box: Transparency in the Collection and Production of Big Twitter Data", *International Journal of Communication* 8, 1745-1764.
- DURKHEIM E., 2003, *Les formes élémentaires de la vie religieuse*, Paris, PUF, (1^{ère} éd. 1912).
- DURKHEIM, E., 1897, *Le suicide*, Paris, Alcan.

- EISENSTEIN, E., 1991, *La révolution de l'imprimé dans l'Europe des premiers temps modernes*, Paris, La Découverte.
- EYMARD-DUVERNAY F., FAVEREAU O., ORLEAN A., SALAIS R. et THEVENOT L., 2004, « L'économie des conventions ou le temps de la réunification dans les sciences sociales », *Problèmes économiques*, n° 2838, La Documentation française, Paris.
- FOUETILLOU G., 2008, « Le web et le traité constitutionnel européen : Ecologie d'une localité thématique compétitive », *Réseaux*, 2008, n°147, p. 229-257.
- GALLUP G., 1939, *Public Opinion in a Democracy*, Herbert L. Baker Foundation, Stafford Little lectures.
- GRANOVETTER M., 1995, *Getting a Job*, Chicago, University of Chicago Press. HA
- NERZ U., 1983, *Explorer la ville*, Paris, Editions de Minuit.
- HARAWAY, D., 2003, *The Companion Species Manifesto: Dogs, People, and Significant Otherness*, Chicago, Prickly Paradigm Press.
- JAMES W., 1892, *Précis de psychologie*, Les empêcheurs de penser en rond, Paris.
- KATZ E. and LAZARSFELD P., 1955, *Personal Influence: the part played by the people in the flow of mass communication*, Glencoe, Free Press.
- KLEINBERG J. , D. GIBSON, P. RAGHAVAN, 1998, "Inferring Web Communities From Link Topology", In Proc. of the 9th *ACM Conference on Hypertext and Hypermedia (HYPER-98)*, pages 225-234, New York, June 20-24.
- KLEINBERG, J., 2002, "Bursty and Hierarchical Structure in Streams", Proc. 8th *ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*.
- LATOUR B., JENSEN B., VENTURINI T., GRAUWIN S., BOULLIER D., 2012, « The Whole is Always Smaller Than Its Parts'. A Digital Test of Gabriel Tarde's monads », *British Journal of Sociology*, Volume 63, Issue 4, pages 590-615.
- LATOUR B., 2005, *Reassembling the Social - An Introduction to Actor-Network-Theory*, Oxford, Oxford University Press. Traduction française : Latour B., 2006, *Changer la société. Refaire de la sociologie*, Paris, La Découverte.
- LATOUR B., 2011, « Gabriel Tarde. La société comme possession. La preuve par l'orchestre » in DEBAISE D., *Philosophie des possessions*, Les presses du réel, 2011.
- LATOUR B., 1994, « Une sociologie sans objet ? Remarques sur l'interobjectivité », *Sociologie du travail*, n° 4, p. 587-607.
- LAZEGA E., 2001, *The Collegial Phenomenon : The social mechanisms of cooperation among peers in a corporate law partnership*, New York, Oxford University Press.
- LE BECHEC M. ET BOULLIER D., 2014, « Communautés imaginées et signes transposables sur un "web territorial" », *Etudes de communication*, n°42, pp.113-125.
- LESKOVEC J., L. BACKSTROM, J. KLEINBERG, 2009, "Meme-tracking and the Dynamics of the News Cycle", *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*.
- LEVY P., 2000, *World philosophie*, Paris, Odile Jacob, 2000, p.46
- LIPPMANN W., 2008, *Le public fantôme*, Paris, Démopolis (1ere ed. 1925).
- MARRES N. AND WELTEVREDE E., 2013, "Scraping the Social? Issues in live social research", *Journal of Cultural Economy*, 6(3), pp. 313-335
- MARRES N., 2007, "The Issues Deserve More Credit: Pragmatist Contributions to the Study of Public Involvement in Controversy", *Social Studies of Science*, 37, pp. 759-78.
- MAUSS, M. 1950, « Essai sur le don », in *Sociologie et Anthropologie*, Paris, PUF.
- MC LUHAN M., 1964, *Understanding Media: The Extension of Man*, Routledge.
- ORLEAN A., 2011, *L'Empire de la monnaie. Refonder l'économie*, Le Seuil, Paris, 2011.
- PENTLAND, A., 2014, *Social Physics. How good ideas spread. The lessons from a new science*, Penguin Press.
- QUETELET, A., 1846, *Lettre à S.A.R. le Duc régnant de Saxe Cabourg et Gotha, sur la théorie des probabilités, appliquée aux sciences morales et politiques*, Bruxelles, Hayez.
- ROGERS, R., 2013, *Digital Methods*, Cambridge, MA, MIT Press.

TARDE G., 1893, *Monadologie et sociologie*, Paris, Alcan, 55 p.

TARDE G., 1989, *L'opinion et la foule*, Paris, PUF, 1989 (1^{ère} éd. 1901).

TARDE, G., 2001, *Les Lois de l'imitation*, Paris, Les empêcheurs de penser en rond (1^{ère} éd. : 1895).

THEVENOT, L., 1986, « Les investissements de forme » in THEVENOT L. (ed), *Conventions économiques*, Paris, CEE-PUF, pp. 21-71.

TURING A., 1936, "On computable numbers, with an application to the Entscheidungsproblem", *Proceedings of the London Mathematical Society*, Vol. 42, N°2, Pages 230-265.

VIVEIROS DE CASTRO E., 2009, *Métaphysiques cannibales*, Paris, PUF. VON NEUMANN, 1945, First Draft of a Report on the EDVAC.

ZASK J., 2000, *L'opinion publique et son double; Livre I: L'opinion sondée; Livre II: John Dewey, philosophe du public*, Paris, L'Harmattan.

Working Papers : la liste

- Hervé Le Bras, Jean-Luc Racine & Michel Wieviorka, *National Debates on Race Statistics: towards an International Comparison*, FMSH-WP-2012-01, février 2012.
- Manuel Castells, *Ni dieu ni maître : les réseaux*, FMSH-WP-2012-02, février 2012.
- François Jullien, *L'écart et l'entre. Ou comment penser l'altérité*, FMSH-WP-2012-03, février 2012.
- Itamar Rabinovich, *The Web of Relationship*, FMSH-WP-2012-04, février 2012.
- Bruno Maggi, *Interpréter l'agir : un défi théorique*, FMSH-WP-2012-05, février 2012.
- Pierre Salama, *Chine – Brésil : industrialisation et « désindustrialisation précoce »*, FMSH-WP-2012-06, mars 2012.
- Guilhem Fabre & Stéphane Grumbach, *The World upside down, China's R&D and innovation strategy*, FMSH-WP-2012-07, avril 2012.
- Joy Y. Zhang, *The De-nationalization and Re-nationalization of the Life Sciences in China: A Cosmopolitan Practicality?*, FMSH-WP-2012-08, avril 2012.
- John P. Sullivan, *From Drug Wars to Criminal Insurgency: Mexican Cartels, Criminal Enclaves and Criminal Insurgency in Mexico and Central America. Implications for Global Security*, FMSH-WP-2012-09, avril 2012.
- Marc Fleurbaey, *Economics is not what you think: A defense of the economic approach to taxation*, FMSH-WP-2012-10, may 2012.
- Marc Fleurbaey, *The Facets of Exploitation*, FMSH-WP-2012-11, may 2012.
- Jacques Sapir, *Pour l'Euro, l'heure du bilan a sonné : Quinze leçons et six conclusions*, FMSH-WP-2012-12, juin 2012.
- Rodolphe De Koninck & Jean-François Rousseau, *Pourquoi et jusqu'où la fuite en avant des agricultures sud-est asiatiques ?*, FMSH-WP-2012-13, juin 2012.
- Jacques Sapir, *Inflation monétaire ou inflation structurelle ? Un modèle hétérodoxe bi-sectoriel*, FMSH-WP-2012-14, juin 2012.
- Franson Manjali, *The 'Social' and the 'Cognitive' in Language. A Reading of Saussure, and Beyond*, FMSH-WP-2012-15, July 2012.
- Michel Wieviorka, *Du concept de sujet à celui de subjectivation/dé-subjectivation*, FMSH-WP-2012-16, juillet 2012.
- Nancy Fraser, *Feminism, Capitalism, and the Cunning of History: An Introduction*, FMSH-WP-2012-17, august 2012.
- Nancy Fraser, *Can society be commodities all the way down? Polanyian reflections on capitalist crisis*, FMSH-WP-2012-18, august 2012.
- Marc Fleurbaey & Stéphane Zuber, *Climate policies deserve a negative discount rate*, FMSH-WP-2012-19, september 2012.
- Roger Waldinger, *La politique au-delà des frontières : la sociologie politique de l'émigration*, FMSH-WP-2012-20, septembre 2012.
- Antonio De Lauri, *Inaccessible Normative Pluralism and Human Rights in Afghanistan*, FMSH-WP-2012-21, september 2012.
- Dominique Méda, *Redéfinir le progrès à la lumière de la crise écologique*, FMSH-WP-2012-22, octobre 2012.
- Ibrahima Thioub, *Stigmates et mémoires de l'esclavage en Afrique de l'Ouest : le sang et la couleur de peau comme lignes de fracture*, FMSH-WP-2012-23, octobre 2012.
- Danièle Joly, *Race, ethnicity and religion: social actors and policies*, FMSH-WP-2012-24, novembre 2012.
- Dominique Méda, *Redefining Progress in Light of the Ecological Crisis*, FMSH-WP-2012-25, décembre 2012.
- Ulrich Beck & Daniel Levy, *Cosmopolitanized Nations: Reimagining Collectivity in World Risk Society*, FMSH-WP-2013-26, february 2013.
- Xavier Richet, *L'internationalisation des firmes chinoises : croissance, motivations, stratégies*, FMSH-WP-2013-27, février 2013.
- Alain Naze, *Le féminisme critique de Pasolini, avec un commentaire de Stefania Tarantino*, FMSH-WP-2013-28, février 2013.
- Thalia Magioglou, *What is the role of "Culture" for conceptualization in Political Psychology? Presentation of a dialogical model of lay thinking in two cultural contexts*, FMSH-WP-2013-29, mars 2013.
- Byasdeb Dasgupta, *Some Aspects of External Dimensions of Indian Economy in the Age of Globalisation*, FMSH-WP-2013-30, april 2013.
- Ulrich Beck, *Risk, class, crisis, hazards and cosmopolitan solidarity/risk community – conceptual and methodological clarifications*, FMSH-WP-2013-31, april 2013.
- Immanuel Wallerstein, *Tout se transforme. Vraiment tout ?*, FMSH-WP-2013-32, mai 2013.
- Christian Walter, *Les origines du modèle de marche au hasard en finance*, FMSH-WP-2013-33, juin 2013.
- Byasdeb Dasgupta, *Financialization, Labour Market Flexibility, Global Crisis and New Imperialism – A Marxist Perspective*, FMSH-WP-2013-34, juin 2013.
- Kiyomitsu Yui, *Climate Change in Visual Communication: From 'This is Not a Pipe' to 'This is Not Fukushima'*, FMSH-WP-2013-35, juin 2013.
- Gilles Lhuillier, *Minerais de guerre. Une nouvelle théorie de la*

- mondialisation du droit, FMSH-WP-2013-36, juillet 2013.
- David Tyfield, *The Coal Renaissance and Cosmopolitized Low-Carbon Societies*, FMSH-WP-2013-37, juillet 2013.
- Lotte Pelckmans, *Moving Memories of Slavery: how hierarchies travel among West African Migrants in Urban Contexts (Bamako, Paris)*, FMSH-WP-2013-38, juillet 2013.
- Amy Dahan, *Historic Overview of Climate Framing*, FMSH-WP-2013-39, août 2013.
- Rosa Rius Gatell & Stefania Tarantino, *Philosophie et genre: Réflexions et questions sur la production philosophique féminine en Europe du Sud au XX^e siècle (Espagne, Italie)*, FMSH-WP-2013-40, août 2013.
- Angela Axworthy *The ontological status of geometrical objects in the commentary on the Elements of Euclid of Jacques Peletier du Mans (1517-1582)*, FMSH-WP-2013-41, août 2013.
- Pierre Salama, *Les économies émergentes, le plongeon ?*, FMSH-WP-2013-42, août 2013.
- Alexis Nuselovici (Nous), *L'exil comme expérience*, FMSH-WP-2013-43, septembre 2013.
- Alexis Nuselovici (Nous), *Exilience : condition et conscience*, FMSH-WP-2013-44, septembre 2013.
- Alexis Nuselovici (Nous), *Exil et post-exil*, FMSH-WP-2013-45, septembre 2013.
- Alexandra Galitzine-Loumpet, *Pour une typologie des objets de l'exil*, FMSH-WP-2013-46, septembre 2013.
- Hosham Dawod, *Les réactions irakiennes à la crise syrienne*, FMSH-WP-2013-47, septembre 2013.
- Gianluca Manzo, *Understanding the Marriage Effect: Changes in Criminal Offending Around the Time of Marriage*, FMSH-WP-2013-48, GeWoP-1, octobre 2013.
- Torkild Hovde Lyngstad & Torbjørn Skarðhamar, *Understanding the Marriage Effect: Changes in Criminal Offending Around the Time of Marriage*, FMSH-WP-2013-49, GeWoP-2, octobre 2013.
- Gunn Elisabeth Birkelund & Yannick Lemel, *Lifestyles and Social Stratification: An Explorative Study of France and Norway*, FMSH-WP-2013-50, GeWoP-3, octobre 2013.
- Franck Varenne, *Chains of Reference in Computer Simulations*, FMSH-WP-2013-51, GeWoP-4, octobre 2013.
- Olivier Galland & Yannick Lemel, avec la collaboration d'Alexandra Frenod, *Comment expliquer la perception des inégalités en France ?*, FMSH-WP-2013-52, GeWoP-5, octobre 2013.
- Guilhem Fabre, *The Lion's share : What's behind China's economic slowdown*, FMSH-WP-2013-53, octobre 2013.
- Venni V. Krishna, *Changing Social Relations between Science and Society: Contemporary Challenges*, FMSH-WP-2013-54, novembre 2013.
- Isabelle Huault & Héléne Rainelli-Weiss, *Is transparency a value on OTC markets? Using displacement to escape categorization*, FMSH-WP-2014-55, janvier 2014.
- Dominique Somda, *Une humble aura. Les grandes femmes au sud de Madagascar*, FMSH-WP-2014-56, janvier 2014.
- Débora González Martínez, *Sur la translatio de miracles de la Vierge au Moyen Âge. Quelques notes sur les Cantigas de Santa Maria*, FMSH-WP-2014-57, janvier 2014.
- Pradeep Kumar Misra, *The State of Teacher Education in France: A Critique*, FMSH-WP-2014-58, janvier 2014.
- Naeem Ahmed, *Pakistan's Counterterrorism strategy and its Implications for domestic, regional and international security*, FMSH-WP-2014-59, janvier 2014.
- Anatole Fogou, *Histoire, conscience historique et devenir de l'Afrique : revisiter l'historiographie diopienne*, FMSH-WP-2014-60, janvier 2014.
- Pierre Salama, *Les classes moyennes peuvent-elles dynamiser la croissance du PIB dans les économies émergentes?*, FMSH-WP-2014-61, février 2014.
- Marta Craveri & Anne-Marie Losonczy, *Growing up in the Gulag: later accounts of deportation to the USSR*, FMSH-WP-2014-62, february 2014.
- Philippe Steiner, *The Organizational Gift and Sociological Approaches to Exchange*, FMSH-WP-2014-63, GeWoP-6, february 2014.
- Françoise Bourdarias, Jean-Pierre Dozon & Frédéric Obringer, *La médecine chinoise au Mali. Les économies d'un patrimoine culturel*, FMSH-WP-2014-64, février 2014.
- Ilan Bizberg, *The welfare state and globalization in North America*, FMSH-WP-2014-65, may 2014.
- Philippe Steiner, *Cartographie des échanges*, FMSH-WP-2014-66, GeWoP-7, mai 2014.
- Olga Stepanova, *Le roman, la pièce de théâtre et le film : traits communs et particularités*, FMSH-WP-2014-67, mai 2014.
- Flavia Buzzetta, *Adaptations de thèmes magico-cabalistiques juifs médiévaux par le Quattrocento italien*, FMSH-WP-2014-68, mai 2014.
- Frédéric Landy, *Quelle sécurité alimentaire en Inde ? Dilemmes économiques, socio-politiques et environnementaux. Une mise en miroir francilienne*, FMSH-WP-2014-69, juin 2014.
- Hafidha Chekir, *Le combat pour les droits des femmes dans le monde arabe*, FMSH-WP-2014-70, juin 2014.
- Géraldine Thiry, Philippe Roman, *The Inclusive Wealth Index. A Sustainability Indicator, Really?*, FMSH-WP-2014-71, juin 2014.
- Michael Cronin, *Représenter l'exil: le sujet du non-exil*, FMSH-WP-2014-72, juin 2014.

- Marc Goldschmit, *L'écriture de l'exil et l'hypothèse du Marrane (Kafka, Benjamin, Derrida et au-delà)*, FMSH-WP-2014-73, juin 2014.
- Boris Chukhovich, *Le street art, un genre exilique ?*, FMSH-WP-2014-74, juin 2014.
- Palanigounder Duraisamy, *Who Wins in the Indian Parliament Election? Criminals, Wealthy or Incumbents*, FMSH-WP-2014-75, august 2014.
- Denis Kondakov, *Francophonie en Biélorussie aux XVIII^e et XIX^e siècles*, FMSH-WP-2014-76, août 2014.
- Isabel Lustosa, *Le séjour de don Pedro 1^{er} à Paris et la presse française (1831/1832)*, FMSH-WP-2014-77, août 2014.
- Lucas Chancel, Géraldine Thiry, Damien Demailly, *Les nouveaux indicateurs de prospérité : pour quoi faire ? Enseignements de six expériences nationales*, FMSH-WP-2014-78, septembre 2014.
- Alex M. Nading, *Local Biologies and the Chemical Infrastructures of Global Health*, FMSH-WP-2014-79, september 2014.
- Maria Conterno, *"Storytelling" and "History writing" in Seventh-Century Near East*, FMSH-WP-2014-80, september 2014.
- Jean-Luc Racine, *Penser l'Inde émergente : de l'altérité orientaliste au post-postcolonialisme*, FMSH-WP-2014-81, septembre 2014.
- Brinda J. Mehta, *Fractures historiques, trauma et résistance dans l'écriture féministe algérienne : Maïssa Bey, Assia Djebar et Leïla Sebbar*, FMSH-WP-2014-82, novembre 2014.
- Sadia Chérif, *Construire la résilience au changement climatique par les connaissances locales : le cas des régions montagneuses et des savanes de Côte d'Ivoire*, FMSH-WP-2014-83, novembre 2014.
- Géraldine Thiry, Léa Sébastien, Tom Bauler, *Ce que révèle le discours des acteurs officiels sur un « au-delà du PIB »*, FMSH-WP-2014-84, novembre 2014.
- Sophie Roche, *The faithful assistant. Mubiddin Faizulloev's life and work in the light of Soviet ethnography*, FMSH-WP-2014-85, novembre 2014.
- Mala Singh, *Re-thinking Knowledge and Social Change in South Africa*, FMSH-WP-2014-86, novembre 2014.
- Georges Corm, Christiane Veauvy, *Proche-Orient et conscience historique, entretien*, FMSH-WP-2015-87, janvier 2015.
- Dominique Boullier, *Les sciences sociales face aux traces du big data ? Société, opinion et répliques*, FMSH-WP-2015-88, février 2015.

Position Papers : la liste

- Jean-François Sabouret, *Mars 2012 : Un an après Fukushima, le Japon entre catastrophes et résilience*, FMSH-PP-2012-01, mars 2012.
- Ajay K. Mehra, *Public Security and the Indian State*, FMSH-PP-2012-02, mars 2012.
- Timm Beichelt, *La nouvelle politique européenne de l'Allemagne : L'émergence de modèles de légitimité en concurrence ?*, FMSH-PP-2012-03, mars 2012.
- Antonio Sérgio Alfredo Guimarães, *Race, colour, and skin colour in Brazil*, FMSH-PP-2012-04, july 2012.
- Mitchell Cohen, *Verdi, Wagner, and Politics in Opera. Bicentennial Ruminations*, FMSH-PP-2012-05, may 2013.
- Ingrid Brena, *Les soins médicaux portés aux patients âgés incapables de s'autogérer*, FMSH-PP-2013-06, avril 2013.
- Thalia Magioglou, *Refaire l'Europe ou refaire le « monde » ? Un commentaire sur l'ouvrage : « Refaire l'Europe avec Jürgen Habermas »*, FMSH-PP-2013-07, septembre 2013.
- Samadia Sadouni, *Cosmopolitisme et prédication islamique transfrontalière : le cas de Maulana Abdul Aleem Siddiqui*, FMSH-PP-2013-08, septembre 2013.
- Alexis Nuselovici (Nouss), *Étudier l'exil*, FMSH-PP-2013-09, septembre 2013.

Retrouvez tous les working papers et les position papers sur notre site, sur hypotheses.org et sur les archives ouvertes [halshs](http://halshs.archives-ouvertes.fr)

<http://www.fmsh.fr/fr/ressources/working-papers>

<http://halshs.archives-ouvertes.fr/FMSH-WP>

<http://wpfmsh.hypotheses.org>