

# Méthodes digitales: Approches quali/quanti des données numériques - Curation and Presentation of the Special Issue

Tommaso Venturini, Dominique Cardon, Jean-Philippe Cointet

▶ **To cite this version:**

Tommaso Venturini, Dominique Cardon, Jean-Philippe Cointet. Méthodes digitales: Approches quali/quanti des données numériques - Curation and Presentation of the Special Issue. Réseaux : communication, technologie, société, 2015, 188 (6), 10.3917/res.188.0009 . hal-01672302

**HAL Id: hal-01672302**
**https://hal.science/hal-01672302**

Submitted on 23 Dec 2017

**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.

CAIRN.INFO
Chercher, repérer, avancer.

# PRÉSENTATION

**Tommaso Venturini *et al.***

**La Découverte |** *Réseaux*

**2014/6 - n° 188**
**pages 9 à 21**

**ISSN 0751-7971**

Article disponible en ligne à l'adresse:

---

**http://www.cairn.info/revue-reseaux-2014-6-page-9.htm**

---

Pour citer cet article :

---

Venturini Tommaso *et al.*, « Présentation »,
*Réseaux*, 2014/6 n° 188, p. 9-21. DOI : 10.3917/res.188.0009

---

Distribution électronique Cairn.info pour La Découverte.

© La Découverte. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

# PRÉSENTATION

Tommaso VENTURINI
Dominique CARDON
Jean-Philippe COINTET

Après avoir souffert d'un manque chronique de données sur leurs objets d'étude, les sciences sociales se trouvent soudainement submergées par une avalanche de traces produites par le développement des médias électroniques (Giles, 2012). La numérisation des données offre de nouvelles voies d'accès aux terrains de recherche, en profitant de trois transformations largement interconnectées. La première est la réduction drastique des coûts de constitution de corpus documentaires, favorisée par la numérisation des archives qui rend de moins en moins nécessaire de se déplacer physiquement et d'engager de longues recherches pour constituer un corpus (presse, littérature scientifique, documents légaux, brevets…). La deuxième est la mise à disposition, dans le cadre du mouvement des données ouvertes, de jeux de données de plus en plus nombreux, récoltés par des institutions publiques et privées (y compris les données de la recherche elle-même), facilitant les possibilités de les interroger, de les croiser et de les réutiliser. Enfin, la troisième transformation est la possibilité offerte par les médias sociaux d'observer in vivo certains aspects de la vie collective (sociabilités, intérêts, goûts ou opinions), voire de mesurer à grande échelle des processus sociaux jusque-là réfractaires à la quantification (Rogers, 2013).

Si l'abondance succède à une certaine forme de pénurie, et au-delà de l'enthousiasme suscité par les promesses de sciences sociales « computationnelles » (Lazer et al., 2008), cette nouvelle « donne » engage les sciences sociales à une certaine vigilance quant à la façon dont les méthodologies et les pratiques se trouvent transformées. Certains auteurs (Boyd et Crawford, 2011) soulignent

DOI: 10.3917/res.188.0009

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

combien ces nouveaux jeux de données invitent à bien préciser le répertoire des interprétations avec lesquelles ils sont compatibles, celles-ci étant aussi bien susceptibles d'étendre que de réduire la gamme des analyses classiquement mobilisées en sciences sociales. L'arrivée des traces numériques suscite en effet des réactions ambivalentes de la part des chercheurs. Là où les optimistes saluent avec enthousiasme l'avènement d'une nouvelle source d'informations potentiellement plus riche et moins coûteuse, les pessimistes regardent avec méfiance la turbulence soulevée par une avalanche de traces confuses et peu éclairantes. Beaucoup de ceux qui se lancent avec transport dans l'expérience numérique reviennent déçus par la difficulté d'obtenir des résultats à la hauteur de ceux auxquels les méthodes traditionnelles, qualitatives ou quantitatives, les ont habitués.

Sans doute cette désillusion tient-elle aux malentendus qui entourent la mise en œuvre de méthodes numériques. Quatre d'entre eux retiendront notre attention : 1) une conception à la fois trop restreinte et trop ambitieuse des traces numériques ; 2) l'alternance de méconnaissance et de méfiance à l'égard de leurs conditions de production ; 3) une tendance à confondre méthode numérique et traitement automatique ; 4) l'espoir que les traces numériques se laissent facilement apprivoiser par les méthodes quantitatives classiques. Le dossier de ce numéro de Réseaux propose une série d'articles qui, d'une manière ou d'une autre, questionnent ces malentendus en esquissant des solutions originales aux difficultés méthodologiques de la recherche numérique dans les sciences sociales.

## QUE REPRÉSENTENT LES DONNÉES NUMÉRIQUES ?

Le premier malentendu relatif à l'usage des traces numériques par les sciences sociales porte sur la question de leur représentativité. De quelles réalités les traces numériques apportées par le « data deluge » sont-elles le référent ? Les données extraites des grandes plates-formes du web (Google, Facebook, Wikipédia, Twitter, Amazon, etc.) doivent-elles être employées pour l'analyse des seules pratiques numériques, ou peut-on en inférer des interprétations relatives à des phénomènes dont les données numériques ne seraient que le reflet (Marres et Weltevrede, 2013), à l'instar de ces études cherchant à prédire le succès d'un film ou le cours d'une action en bourse à partir des flux de conversations sur Twitter ? La vague de prophéties qui entourent aujourd'hui les big data entretient l'idée qu'en raison de leur volume exorbitant, les données numériques auraient suffisamment « absorbé » le réel pour le représenter plus efficacement

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

que toutes autres techniques d'objectivation. Le service Google Flu Trends qui mesure – assez mal (Lazer et al., 2014) – la progression géographique des épidémies de grippe prétend ainsi produire des indicateurs aussi bons que ceux des institutions sanitaires. De façon analogue, appliqués aux contenus des services web, les outils de sentiment analysis permettraient de mesurer et de qualifier les évolutions de l'opinion publique (Boullier et Lohard, 2012).

Ce malentendu dérive de la réduction du numérique à ses expressions les plus visibles (c'est-à-dire les grandes plates-formes du web citées) et de la prétention de substituer le volume des données à la représentativité de leur échantillonnage. Il est illusoire de penser que la taille des données circulant sur les flux des API (Application protocol interface) de Google ou de Twitter serait suffisante pour en faire de fidèles miroirs de la vie collective. La théorie statistique enseigne en effet que la représentativité d'un échantillon ne dépend qu'indirectement de sa taille : la qualité d'un échantillon dérive plutôt de sa ressemblance à la population de référence, de sa capacité d'en inclure la même variabilité.

Maîtriser les données numériques demande alors de développer des méthodes raisonnées d'échantillonnage et d'interprétation de ces données. Les travaux conduits à partir d'extraction d'informations issues des blogs, de Google ou de Facebook permettent d'apprendre beaucoup sur les usages de ces plates-formes, sans doute moins sur les phénomènes généraux que, par généralisation, elles voudraient représenter. En revanche, une interrogation sur les articulations entre les enregistrements numériques et d'autres techniques d'objectivation des pratiques sociales est souvent très féconde, à l'instar des travaux étudiant la correspondance entre sociabilité en ligne et sociabilité « réelle » (par exemple : Burke et Kraut, 2014). De nombreux travaux usant de « méthodes digitales » couplent le flux de données extraites des API du web avec des dispositifs d'interrogation ad hoc (questionnaire, échantillonnage, tirage aléatoire, etc.) permettant de contrôler la représentativité des flux capturés. Dans sa thèse sur la formation des opinions politiques sur Twitter, Julien Boyadjian (2014), par exemple, a procédé à un contrôle par questionnaire des propriétés sociales des utilisateurs sélectionnés afin de contrôler son échantillon.

Plutôt que de céder au gigantisme des big data, ce qui fait souvent l'originalité des digital methods tient à la manière dont les chercheurs parviennent à recomposer des jeux de données différents et à imposer un cadrage raisonné aux procédures de sélection qu'ils effectuent sur ces flux. Les nouveaux outils

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

numériques permettent aussi de constituer et de travailler des sources de données sans doute moins « big », mais sûrement plus spécifiques, pour aborder leur objet. Dans ce numéro, Thomas Tari étudie les discours autour des différentes générations de biocarburants en s'efforçant d'aller au-delà des données offertes par les bases bibliométriques classiques. En fabriquant une base de données à partir des projets de R&D en agronomie, il déplace l'attention des résultats de la recherche scientifique (publications scientifiques et brevets) vers ses dynamiques et ses financeurs. De manière analogue, Nicolas Baya-Laffite et Jean-Philippe Cointet analysent l'évolution thématique des négociations sur le climat en s'appuyant sur un ensemble de (quasi-)transcriptions des débats au sein de la Conférence des Nations unies sur le changement climatique. En raison de son extrême spécialisation, ce corpus de données numériques impose au chercheur un traitement particulier, mais offre aussi des résultats originaux.

## LES DONNÉES NUMÉRIQUES NE SONT PAS CONSTRUITES PAR ET POUR LES SCIENCES SOCIALES

Un deuxième malentendu relatif aux données numériques tient au fait que les méthodes numériques, à la différence des méthodes traditionnelles, s'appuient sur des données qui sont généralement collectées pour des finalités autres que la recherche scientifique. Il s'agit d'informations récoltées pour des besoins de marketing (cartes de fidélité ou cartes bancaires), de surveillance (déplacements aériens), d'optimisation technique (serveurs Internet), d'information (open data institutionnels). Comme le souligne Francis Chateauraynaud dans ce numéro, la sociologie (notamment celle sur les controverses sociotechniques) est aujourd'hui « débordée par la prolifération, sur la toile, de sites de discussion et de modes d'expression. Pour y faire face, elle doit sortir de la conception antérieure des données, selon laquelle un dossier était représenté par un corpus et un seul, dont la clôture était à la main du chercheur. » Puisque la plupart des nouveaux gisements de données numériques n'ont pas été construits pour ou par les chercheurs (en tout cas pas uniquement), leur utilisation demande toujours de s'interroger sur les conditions de leur production. Si la Banque mondiale met en ligne ses données statistiques (data.worldbank.org), il importe de savoir comment ces données ont été constituées et de saisir les raisons de leur divulgation. Si American Online publie par erreur vingt millions de requêtes vers son moteur de recherche, il faut se demander si l'utilisation de ces données est éthiquement correcte (Ess and AoIR ethics working committee, 2002). Lorsque Wikipédia met à disposition une API pour télécharger

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

tous ses articles (avec l'histoire complète de leur rédaction) (mediawiki.org/wiki/API), une réflexion sur le statut épistémique des traces de cette entreprise collective est nécessaire (Viegas, Wattenberg, Kriss et Van Ham, 2007).

Interroger les conditions de production des données numériques permet de résister à la tentation de les naturaliser. Comme toute source secondaire, la construction des données est toujours le résultat d'une longue chaîne d'actions dont certains maillons échappent au contrôle direct de l'expérimentateur. Aussi importe-t-il que l'expérimentateur et ses pairs puissent remonter tous les passages de la chaîne et vérifier leur solidité (Latour, 1993). Ce n'est pas toujours complètement le cas avec les données extraites des plates-formes numériques dont l'accès est parfois livré aux desiderata des propriétaires des sites ou soumis aux décisions d'algorithmes dont le principe de fonctionnement est obscur et secret (Pasquale, 2015). L'enquête en milieu digital suppose toujours une minutieuse investigation des conditions de production et de restitution des données numériques. Même lorsqu'il s'agit de traces mises à disposition par d'autres chercheurs, leur réutilisation reste problématique (Carlson et Anderson, 2007). En effet, dans un corpus numérique, la distinction entre bruit et information ne peut jamais être faite a priori. Elle dépend strictement des objectifs de recherche et du type d'analyse. Si, par exemple, on cherche à détecter les épidémies de grippe via l'analyse de requêtes Google (Ginsberg et al., 2008), on considérera les fonctionnalités d'autosuggestion comme un artefact. A contrario, c'est cette fonctionnalité même qui constituera le matériau empirique d'une analyse comparative des différences culturelles dans les recherches sur Google (Mazières et Huron, 2013).

Reconnaître l'artificialité des traces numériques n'interdit cependant pas d'en faire un usage scientifique. Comme voudrait en témoigner ce numéro de Réseaux, les producteurs de ces données ne complotent pas pour piéger les sciences sociales. Refuser systématiquement d'utiliser des données numériques au prétexte qu'elles sont produites dans des contextes industriels ou administratifs spécifiques risque de conduire les sciences sociales à manquer le tournant numérique de nos sociétés. Dans l'article qu'ils consacrent à la topographie de la renommée au sein des territoires de la blogosphère, Dominique Cardon, Guilhem Fouetillou et Camille Roth traitent de données issues du travail préalable de catégorisation du web effectué par la société Linkfluence. Il ne fait guère de doute qu'un classement thématique des sites et blogs français et allemands entrepris dans un contexte scientifique produirait des résultats différents de ceux effectués dans ce contexte industriel. Pour autant, les opérations statistiques réalisées à partir de ce jeu de données

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

permettent bien de produire des résultats interprétables et pouvant être contrôlés par un va-et-vient entre résultats macro et lecture micro des blogs.

## LE NUMÉRIQUE N'EST PAS AUTOMATIQUE

Un troisième malentendu réside dans la confusion fréquente dans les discours sur les big data consistant à penser que les traitements des données sont à la fois « automatiques » et « objectifs ». Une représentation ancienne de l'informatique dans les sciences sociales concevait l'ordinateur comme un outil destiné à soulager les chercheurs des tâches les plus « mécaniques ». Or, chaque fois qu'un nouvel instrument s'est imposé dans les laboratoires, il a fini par influencer profondément le cours des sciences. Le télescope n'a pas seulement permis de voir plus loin, il a permis de voir des choses différentes et de voir différemment. L'imprimerie n'a pas seulement permis l'accumulation et la circulation de plus d'ouvrages, elle a aussi rendu possibles de nouveaux rapports aux savoirs (Eisenstein, 1980). Ainsi les technologies numériques ne se limitent pas à se mettre au service des méthodes d'enquête traditionnelles, elles créent aussi de nouvelles méthodes et de nouvelles interrogations. Richard Rogers (2009) oppose en ce sens les méthodes « numérisées » aux méthodes « nativement numériques » (digitized/digital native methods).

L'expérience montre que les chercheurs désireux de se simplifier la vie ne trouveront pas leur bonheur avec les nouvelles méthodes digitales. Développer une enquête numérique demande toujours de remettre en question des habitudes et des présupposés de recherche. La simplification automatique des traitements n'est jamais présente au rendez-vous et la recherche informatisée n'est pas plus rapide que la recherche traditionnelle ; les données prétendument disponibles ne se laissent pas collecter aussi facilement qu'on l'avait espéré ; les données sont souvent sales et leurs catégorisations fragiles ; les algorithmes d'analyse prolifèrent, mais ils sont souvent mal documentés et doivent être adaptés aux données disponibles ; les techniques de visualisation sont souvent aussi séduisantes que trompeuses dans les modes de lecture qu'elles induisent ; etc. Il n'est pas rare que le travail harassant de récolte de la donnée soit décourageant (Bastard et al., 2013). Surtout, il n'économise en rien la nécessité pour le chercheur de formuler des hypothèses. Sur la fiction de l'automatisation des opérations de recherche s'est en effet greffée la croyance dans une neutralité des calculs parvenant, sans hypothèses préalables, à produire des résultats objectifs. À l'instar du modèle de « l'objectivité instrumentale » (Daston et Galison, 2012) des sciences

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

naturelles, la mécanisation de l'observation promet l'effacement de la subjectivité du chercheur. L'idée selon laquelle l'alliance des big data et des algorithmes donnerait naissance à un nouveau positivisme, formulée dans l'article provocateur de Chris Anderson, « The End of Theory: The Data Deluge Makes the Scientific Method Obsolete » a connu un large succès. Il y soutient notamment que :

> « Avec suffisamment de données, les nombres parlent d'eux-mêmes. Les petabytes nous permettent de dire : "la corrélation est suffisante". Nous pouvons arrêter de chercher des modèles. Il est possible d'analyser les données sans hypothèses à propos de ce qu'elles pourraient révéler. Il suffit de jeter les données dans les plus gros calculateurs que le monde n'a jamais connus et laisser les algorithmes trouver des phénomènes là où la science ne le peut pas. »

L'écho négatif de ce texte chez les chercheurs en sciences sociales trahit autant une inquiétude qu'une fragilité devant l'arrogance scientiste des promoteurs du big data. Tout data scientist sait qu'il n'est pas de méthode d'apprentissage qui ne soit « supervisée », ce qui veut dire qu'elle incorpore des hypothèses et des objectifs dans les calculs qu'elle commande (les techniques dites « non supervisées » en sont encore à leur balbutiement et servent principalement à la détection de formes dans des images). L'article de Thierry Poibeau dans ce numéro montre, par exemple, comment le Traitement Automatique du Langage (TAL) ne devient « automatique » qu'au prix d'une multitude de choix et d'opérations manuelles de la part des chercheurs. Dans la synthèse du domaine qu'il propose, Thierry Poibeau montre la diversité des outils à disposition du chercheur tout en soulignant que les choix méthodologiques effectués conduisent vers des représentations et des interprétations très différentes des données initiales : « Un texte ne saurait être assimilé à une masse de connaissances directement exploitable par la machine. Il faut dans un premier temps prévoir des traitements complexes pour identifier l'information pertinente, la normaliser, la catégoriser et éventuellement la mettre en contexte. Alors seulement l'ordinateur ou l'expert sera capable d'en tirer parti pour mener à bien ses analyses. » Le débat entourant le texte de Chris Anderson montre à quel point un effort redoublé pour conduire à la fois une analyse critique des données et une réflexion exigeante sur la fabrication des hypothèses de recherche constitue un objectif essentiel pour les digital methods. En effet, le risque qu'elles doivent d'abord affronter est celui de produire, au terme d'un gigantesque effort, des résultats évidents, triviaux et qui n'ont souvent pour eux que la séduction des nouvelles visualisations interactives permises par les outils numériques.

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

## GAGNER EN QUANTITÉ SANS PERDRE EN QUALITÉ

Un dernier malentendu tient au fait que si la nouvelle traçabilité de la vie sociale produit davantage de données, elle provoque aussi plus de bruit. De la fouille de données (datamining), on n'extrait souvent quelques pépites qu'après avoir évacué d'énormes quantités de terre. Dans le cas des traces numériques pour les sciences sociales, le travail de nettoyage est à la fois décisif et complexe. Ce labeur demande une attention d'autant plus minutieuse que les flux de données numériques deviennent plus vastes et divers. L'augmentation de la quantité de traces à disposition des sciences sociales ne peut donc pas toujours être exploitée par la simple utilisation des méthodes quantitatives classiques. Basées sur l'agrégation et la corrélation, ces méthodes ont été développées pour s'adapter à des données relativement simples, dont le chercheur peut contrôler le nombre de variables. Appliquées à une foule de traces riches, mais très incertaines, ces méthodes peuvent facilement entraîner des erreurs d'interprétation. Plus vaste est la quantité des données manipulées et meilleure doit être la qualité de leur compréhension. L'arrivée des traces numériques n'assure pas le triomphe des méthodes quantitatives sur les méthodes qualitatives. À l'inverse, plus le chercheur dispose de données, plus leur agrégation, leur découpage et leur articulation doivent être validés par la possibilité d'observer chaque information dans son contexte de production (ce qui est rendu justement plus facile par les possibilités de navigation numérique dans les corpus). Les traces numériques peuvent offrir davantage de quantification, mais seulement au prix d'une meilleure qualification. Cependant, souligne Francis Chateauraynaud dans ce numéro, « il ne s'agit pas seulement de lier une échelle macro et une plongée dans des détails micro, mais aussi de repérer les processus par lesquels des jeux d'acteurs et d'arguments se déplacent, changent de cadres et de modalités en produisant des effets loin de leur contexte d'apparition ». Son article montre comment, afin de comprendre les discours sur les futurs de nanotechnologies, il est impossible de se fier exclusivement à une analyse classique de cooccurrence lexicale. Les arguments dans ce champ se jouent plus sur les modalités et les adverbes que sur les mots clés. Leur analyse doit donc rester sensible à ces subtilités de l'argumentation, tout en préservant le lien avec un corpus très vaste de textes rassemblés en dossiers historicisés (6 000 documents dans un corpus qui court sur une période de 25 années).

Les nouvelles traces numériques demandent non seulement d'être traitées successivement par des méthodes quantitatives et par des méthodes qualitatives, elles requièrent aussi et surtout d'être traitées par de nouvelles

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

méthodes quali-quantitatives (Venturini, 2012). S'il est absurde de céder aux sirènes du « correlation is enough » et de jeter aux orties l'expérience accumulée avec les méthodes quantitative et qualitative classiques (Savage et Burrows, 2007), il serait pour autant imprudent d'imaginer que les méthodes des sciences sociales ne soient pas profondément transformées par les traces digitales (Marres, 2012). Et s'il est un aspect des méthodes sociologiques que les traces numériques peuvent transformer, c'est précisément la séparation entre approches qualitatives et quantitatives. Cette coupure est, en quelque sorte, une conséquence de la carence de données qui a jusqu'ici caractérisé l'analyse des phénomènes collectifs. Habituées à payer très cher chaque fragment d'information sur la vie collective, les sciences sociales étaient forcées à un choix cornélien entre acquérir peu d'informations sur beaucoup d'acteurs (avec les méthodes quantitatives telles que les sondages et les questionnaires) ou beaucoup d'informations sur peu d'acteurs (avec les méthodes qualitatives telles que les entretiens et l'ethnographie), mais rarement les deux en même temps : survoler la surface des phénomènes collectifs et identifier les régularités globales, ou examiner en détail les interactions locales et apprécier la richesse de leur singularité. Les chercheurs peuvent, dans certains contextes, collecter aujourd'hui beaucoup d'informations sur beaucoup d'acteurs et à un prix raisonnable. Pour la première fois, au lieu de sauter du local au global, ils peuvent retracer comment des milliers d'interactions s'enlacent pour tresser le tissu de la vie collective.

Toutefois, les méthodes quali-quantitatives permettant d'associer qualité et quantité dans l'exploitation des traces numériques restent encore à développer. Les outils qui permettraient de doser agrégation et contextualisation, d'alterner entre l'appréciation des motifs plus larges et les détails les plus minuscules, sans négliger tout ce qui est entre les deux, restent encore à construire. Selon des approches différentes, plusieurs travaux récents ont proposé des voies originales pour penser autrement et différemment ces articulations : analyse des trajectoires « balistiques » des arguments se déplaçant l'espace du débat public (Chateauraynaud, 2011) ; alignement sur un même plan des acteurs et des agrégats (Latour et al., 2012) ; suivi des trajectoires de notoriété dans la blogosphère (Cardon et al., 2011) ; phylogénies des disciplines scientifiques (Chavalarias et Cointet 2013) ; identification d'effets de style et de genre par l'analyse de millions d'images individuelles (Manovitch, 2012b) ; « géo-médiatique » comme projection de données médiatique dans l'espace (Grasland et al., 2012). Déjà habituées à alterner grand angle et téléobjectif, les sciences sociales doivent maintenant s'équiper d'autofocales pour accommoder leur perspective à toutes les échelles.

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

L'objectif de ce numéro spécial de la revue Réseaux dédié aux méthodes quali-quantitatives est de rendre compte de la diversité de ces nouvelles approches. Jour après jour de nouvelles méthodes rétives à la dichotomie classique entre approches qualitatives et quantitatives prennent forme. Ces méthodes circulent entre micro et macro, local et global, permettant aux chercheurs de traiter des larges quantités de données sans perdre en finesse d'analyse. Ce sont précisément ces expériences que nous avons souhaité rassembler dans ce numéro de Réseaux, sans exclusion de disciplines, d'objets ou d'approches.

On trouvera également dans ce numéro deux articles en Varia. Jonathan Chibois s'est intéressé à la manière dont fonctionnait à l'Assemblée nationale la salle des quatre colonnes, espace dédié aux relations entre députés et journalistes depuis que Twitter a créé un nouveau canal de communication. Il constate que les rapports entre les politiques et la sphère médiatique ne se sont pas fondamentalement transformés.

De son côté, Marie Pruvost-Delaspre interroge les archives pour qualifier ce qu'ont été les modèles de production dans le cinéma d'animation japonais des années 1950-1960. Elle le fait autour de deux hypothèses, la concurrence avec les studios américains et la manière dont le studio de la Tôei Dôga, « le Disney de l'Orient », aura formaté le paysage de l'industrie du dessin animé japonais.

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

## RÉFÉRENCES

ANDERSON, C. (2008), « The End of Theory: The Data Deluge Makes the Scientific Method Obsolete », *Wired*, vol. 16, n° 7.

BARNES, T. J. (2014), « What's old is new, and new is old: History and geography's quantitative revolutions », *Dialogues in Human Geography*, vol. 4, n° 1, pp. 50-53.

BASTARD, I., CARDON, D., FOUETILLOU, G., PRIEUR, C., RAUX, S. (2013), « Travail et travailleurs de la donnée », *Internet.actu*, 13 décembre 2013. En ligne : http://www.internetactu.net/2013/12/13/travail-et-travailleurs-de-la-donnee/

BOLLIER, D., FIRESTONE, C. (2010), *The promise and peril of big data*, Washington, The Aspen Institute, 2010. En ligne : http://www.ilmresource.com/collateral/analyst-reports/10334-ar-promise-peril-of-big-data.pdf.

BOULLIER, D., LOHARD, A. (2012), *Opinion Mining et Sentiment Analysis*. Paris, Open Edition Press.

BOYD, D., CRAWFORD, K. (2011), « Six Provocations for Big Data », *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, doi:10.2139/ssrn.1926431.

BOYADJIAN, J. (2014), *Analyser les opinions politiques sur Internet. Enjeux théoriques et défis méthodologiques*. Thèse de Science politique, Université de Montpellier 1.

BURKE, M., KRAUT, R. (2014), « Growing Closer on Facebook: Changes in Tie Strength Through Social Network Site Use », *CHI 2014*, Toronto.

CARDON, D. (2013), « Dans l'esprit du PageRank », *Réseaux*, n° 177.

CARDON, D., FOUETILLOU, G., ROTH, C. (2011), « Two Paths of Glory », *Proceedings AAAI ICWSM'11 Conference Weblogs and Social Media*.

CARLSON, S., ANDERSON, B. (2007), « What Are Data? The Many Kinds of Data and Their Implications for Data Re-Use », *Journal of Computer-Mediated Communication*, vol. 12, n° 2, pp. 635-651.

CARR, N. (2014), « The Limits of Social Engineering », *MIT Technology Review*, 16 avril. En ligne : http://www.technologyreview.com/review/526561/the-limits-of-social-engineering/

CRESSWELL, T. (2014), « Deja vu all over again: Spatial science, quantitative revolutions and the culture of numbers », *Dialogues in Human Geography*, vol. 4, n° 1, pp. 54-58.

CHATEAURAYNAUD, F. (2011), *Argumenter dans un champ de forces. Essai de balistique sociologique*. Paris, Éditions Petra.

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

CHAVALARIAS, D., COINTET, J.-P. (2013), « Phylomemetic patterns in science evolution. The rise and fall of scientific fields », *PloS one*, 11 février. En ligne : http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0054847.

DALTON, C., THATCHER, J. (2014), « What does a critical data studies look like, and why do we care ? Seven points for a critical approach to "big data" », *Society and Space Open Site*. En ligne : http://societyandspace.com/material/commentaries/craig-dalton-and-jim-thatcher-what-does-a-critical-data-studies-look-like-and-why-do-we-care-seven-points-for-a-critical-approach-to-big-data/

EISENSTEIN, E. L. (1980), *The Printing Press as an Agent of Change*. Cambridge, Cambridge University Press.

GAYO-AVELLO, D. (2012), « I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper. A Balanced Survey on Election Prediction using Twitter Data », *arXiv Preprint* arXiv:1204.6441, pp. 1-13. En ligne : http://arxiv.org/abs/1204.6441.

GILES, J. (2012), « Making the Links », *Nature*, n° 488, pp. 448-450.

GINSBERG, J., MOHEBBI, M. H., PATEL, R. S., BRAMMER, L., SMOLINSKI, M. S., BRILLIANT, L. (2008), « Detecting influenza epidemics using search engine query data », *Nature*, n° 457, pp. 1012-1014. doi:10.1038/nature07634.

GRASLAND, C., GIRAUD, T., SEVERO, M. (2012), « Un capteur géomédiatique d'événements internationaux », in P. BECKOUCHE, C. GRASLAND, F. GUÉRIN-PACE, J.-Y. MOISSERON, *Fonder les sciences du territoire*. Paris, Karthala.

HARFORD, T. (2014), « Big data: are we making a big mistake? », *Financial Times Magazine*.

LATOUR, B. (1993), « Le Topofil de Boa Vista ou la Référence Scientifique – montage Photo-Philosophique », *Raison pratique*, n° 4, pp. 187-216.

LATOUR, B., JENSEN, P., 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*, vol. 63, n° 4, pp. 591-615.

LAW, J., RUPPERT, E., SAVAGE, M. (2011), « The double social life of methods ». En ligne : http://www.cresc.ac.uk/sites/default/files/The Double Social Life of Methods CRESC Working Paper 95.pdf.

LAZER, D. et al. (2009), « Life in the network: the coming age of computational social science », *Science* (New York, NY) 323, No. 5915.

LAZER, D., KENNEDY, R., KING, G., VESPIGNANI, A. (2014), « Big data. The parable of Google Flu: traps in big data analysis », *Science* (New York, N.Y.) 343, No. 6176, pp. 1203-1205.

MANOVICH, L. (2012a), « Trending: the promises and the challenges of big social data », in M. K. GOLD, *Debates in the Digital Humanities*, Minneapolis, University of Minnesota Press, pp. 460-475.

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte

MANOVITCH, L. (2012b), « How to Compare One Million of Images », in D. BERRY (dir.), *Understanding Digital Humanities*, Houndmills and New York, Palgrave.

MARCUS, G., DAVIS, E. (2014), « Eight (No, Nine!) Problems With Big Data », *The New York Times*, 6 avril.

MARRES, N. (2012), « The redistribution of methods: on intervention in digital social research, broadly conceived », *The Sociological Review*, n° 60, pp. 139-165.

MARRES, N., WELTEVREDE, E. (2013). « Scraping the Social? Issues in live social research », *Journal of Cultural Economy*, 6(3), pp. 313-335.

MAZIERES, A., HURON, S. (2013), « Toward Google Borders », *Proceedings of the 5th Annual ACM Web Science Conference*, ACM, pp. 244-247.

PASQUALE, F. (2015), *The Black Box Society. The Secret Algorithms That Control Money and Information*. Cambridge, Harvard University Press.

RIEDER, B., RÖHLE, T. (2012), « Digital Methods: Five Challenges », in D. M. BERRY (dir.), *Understanding Digital Humanities*. Houndmills, Palgrave Macmillan, pp. 67-84.

ROGERS, R. (2009), *The End of the Virtual: Digital Methods*. Amsterdam, Amsterdam University Press.

ROGERS, R. (2013), *Digital Methods*. Cambridge, The MIT Press.

ROUVROY, A., BERNS, T. (2013), « Gouvernementalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individualisation par la relation ? », *Réseaux*, n° 177.

SAVAGE, M., BURROWS, R. (2007), « The Coming Crisis of Empirical Sociology », *Sociology*, vol. 41, n° 5, pp. 885-899.

UPRICHARD, E. (2013), « Big Data, Little Questions? », *Discover Society*, 1. En ligne : http://www.discoversociety.org/focus-big-data-little-questions/

VENTURINI, T. (2012), « Great expectations : méthodes quali-quantitatives et analyse des réseaux sociaux », in J.-P. FOURMENTRAUX (dir.), *L'Ère Post-Media. Humanités digitales et Cultures numériques*. Paris, Hermann, vol. 104, pp. 39-51.

Document téléchargé depuis www.cairn.info - Institut d'Etudes Politiques de Paris - - 84.253.47.156 - 04/03/2015 18h41. © La Découverte