

---

Jean-Christophe Plantin et Laurence Monnoyer-Smith

## Ouvrir la boîte à outils de la recherche numérique

### Trois cas de redistribution de méthodes

---

#### Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

**revues.org**

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

---

#### Référence électronique

Jean-Christophe Plantin et Laurence Monnoyer-Smith, « Ouvrir la boîte à outils de la recherche numérique », *tic&société* [En ligne], Vol. 7, N° 2 | 2ème semestre 2013, mis en ligne le 01 juin 2014, Consulté le 02 juin 2014.  
URL : <http://ticetsociete.revues.org/1527> ; DOI : 10.4000/ticetsociete.1527

Éditeur : ARTIC

<http://ticetsociete.revues.org>

<http://www.revues.org>

Document accessible en ligne sur : <http://ticetsociete.revues.org/1527>

Ce document est le fac-similé de l'édition papier.

Licence Creative Commons

# Ouvrir la boîte à outils de la recherche numérique

## Trois cas de redistribution de méthodes

**Jean-Christophe PLANTIN**

Laboratoire Costech  
Université de Technologie de Compiègne  
Jean-christophe.plantin@utc.fr

**Laurence MONNOYER-SMITH**

Laboratoire Costech  
Université de Technologie de Compiègne  
Laurence.monnoyer-smith@debatpublic.fr

Jean-Christophe PLANTIN est chercheur en sciences de l'information et de la communication. Il est l'auteur d'une thèse sur les pratiques de cartographie participative en ligne (laboratoire COSTECH, Université de Technologie de Compiègne), prenant la forme d'une enquête de terrain sur la réalisation et l'utilisation de cartes numériques de radiation après l'accident à la centrale de Fukushima-Daiichi le 11 mars 2011. Il est actuellement chercheur postdoctoral à l'Université du Michigan (*School of Information et Communication Studies department*), où il réalise une étude sur les nouvelles sources de données disponibles pour la recherche en sciences humaines et sociales (*big data, open data, données du web*), les débats méthodologiques et épistémologiques qu'elles suscitent et les infrastructures qu'elles nécessitent pour le partage et la réutilisation des données.

Laurence MONNOYER-SMITH est vice-présidente de la Commission nationale du débat public et professeur des universités. Elle a dirigé le laboratoire de recherche en sciences humaines et sociales COSTECH

Jean-Christophe PLANTIN et Laurence MONNOYER-SMITH

(Connaissance, Organisation, Systèmes TECHniques) de l'Université de Technologie de Compiègne dont les travaux portent sur les relations entre science, technique, et société. Les recherches menées concernent la démocratie électronique (analyse des conditions de mise en œuvre, fondements théoriques, études empiriques), les dispositifs électroniques de médiation citoyenne (forums participatifs, systèmes électroniques de débat public), et les conditions d'émergence de nouvelles formes de citoyenneté. Quant à ses travaux théoriques ils portent sur l'usage des technologies de l'information dans le cadre du processus décisionnel et sur les problèmes d'inclusion des populations dans les débats publics. Elle a publié de nombreux articles et ouvrages, parmi lesquels *Communication et délibération: Les enjeux des mutations techniques pour l'expression citoyenne*, Paris, Hermès Science Publishing (2010) et *Normative experience in Internet Politics*, Paris, Presses de l'École des Mines (2011), en collaboration avec Françoise Massit-Folléa et Cécile Méadel.

## Ouvrir la boîte à outils de la recherche numérique

### Trois cas de redistribution de méthodes

**Résumé :** Les méthodes numériques mobilisées dans le cadre de recherches en sciences humaines et sociales (SHS) sont souvent présentées comme des innovations radicales par rapport aux méthodes existantes. Selon Noortje Marres (2012a), ces dernières sont davantage « redistribuées » afin de s'adapter aux caractéristiques du support numérique. Si ces méthodes ne sont donc pas radicalement différentes, elles renouvellent toutefois des débats épistémologiques provenant du croisement entre informatique et SHS. Cet article analyse trois débats particulièrement importants et illustre chacun d'eux par une approche spécifique de recherche numérique. Les *Digital Humanities* interrogent la possibilité d'un dialogue entre sciences exactes et SHS, dont les épistémologies sont traditionnellement séparées ; les *Cultural Analytics* interrogent la possibilité de faire de la recherche en SHS à partir de grandes masses de données ; enfin, les *Digital Methods* interrogent la nécessité de créer des méthodes spécifiques pour appréhender un nouvel objet numérique. En conclusion, le potentiel applicatif et réflexif de ces travaux pour les recherches en sciences de l'information et de la communication (SIC) sera mis en évidence.

**Mots clés :** *Cultural Analytics*, *Digital Humanities*, *Digital Methods*, données, méthodes, recherche numérique, sciences de l'information et de la communication (SIC).

**Abstract:** Digital methods in humanities and social sciences are often presented as disruptive towards existing methods. On the opposite, Noortje Marres claims that online methods are « redistributed » in order to apply to the digital medium. If these methods are not radically new, they face important epistemological debates. This article will present three of these debates by focusing on three types of digital research. Digital humanities interrogate the possibilities of a dialogue between sciences and the humanities ; Cultural Analytics challenges the possibility to do research in humanities with large datasets ; digital methods adress the necessity to create new methods to apply to a digital research object. The conclusion will present how the reflexive and experimental nature of these works can benefit to information and communication sciences.

Jean-Christophe PLANTIN et Laurence MONNOYER-SMITH

**Keywords:** Cultural Analytics, data, *Digital Humanities*, digital methods, digital research, information & communication sciences, methods.

**Resumen :** Métodos numéricos movilizados en el marco de la investigación en humanidades y ciencias sociales a menudo se presentan como innovaciones radicales sobre los métodos existentes. Según Noortje Marres, son más "redistribuido" para adaptarse a las características de los medios digitales. Si estos métodos no son radicalmente diferentes, están, sin embargo, se enfrentan a desafíos epistemológicos importantes. Este artículo analiza tres debates de especial importancia e ilustra cada una con un enfoque específico para la investigación digital. Las humanidades digitales cuestionan la posibilidad de un diálogo entre las ciencias y SHS, cuya epistemologías son tradicionalmente separadas, los análisis culturales ponen en duda la posibilidad de hacer investigación en SHS a partir de grandes cantidades de datos y, por último, los métodos digitales pregunta la necesidad de la creación de metodologías específicas para entender un nuevo objeto digital. En conclusión, se destacó el potencial de aplicación de estos y reflexivo trabajo para la investigación de la CEI.

**Palabras claves :** Cultural Analytics, data, *Digital Humanities*, digital methods, digital research, information & communication sciences, methods.

## Ouvrir la boîte à outils de la recherche numérique Trois cas de redistribution de méthodes

L'expression recherche numérique désigne les recherches en sciences humaines et sociales (SHS) qui se fondent sur le traitement informatisé de données ; excluant les outils de bureautique ou de publication des résultats de la recherche (par exemple à travers un blogue ou une base de données). Elle désigne les recherches en SHS qui portent sur des objets numériques ou numérisés, tout en ayant recours à des méthodes elles-mêmes numériques.

Ce type de recherches fait actuellement l'objet de débats épistémologiques, méthodologiques et éthiques, notamment dans le champ des sciences de l'information et de la communication. Par exemple, Rieder et Rhöle (2012) mettent en avant les limites d'un fantasme de l'objectivité des chiffres et d'une « science totale » autorisée par le traitement informatisé, mais également les dangers et écueils de la constitution de corpus de données provenant d'applications en ligne (par exemple en passant par l'interface de programmation<sup>1</sup> de Twitter). Ces applications posent en effet des problèmes liés au fonctionnement de ce type d'interface : changements des conditions techniques ou financières d'accès aux données, ou encore représentativité des données obtenues à travers une interface de programmation (Gerlitz et Rieder, 2013). De leur côté, Boyd et Crawford (2012) soulignent les risques d'une nouvelle inégalité entre grands et petits laboratoires, mais aussi d'une perte d'indépendance de la recherche du fait de la multiplication des coopérations étroites avec de grands acteurs du web. En France, Franck Rebillard (2011) observe quant à lui que les outils informatiques migrent d'une discipline à l'autre sans toujours s'accompagner du regard critique nécessaire à leur mise en œuvre. Par ailleurs, le croisement entre méthodes numériques et méthodes traditionnelles nécessite une formation méthodologique qui n'est pas toujours assurée dans les laboratoires. De manière plus critique, Yves Jeanneret (2013) souligne l'imaginaire de l'exhaustivité que comportent les outils de représentation visuelle (comme les cartographies de réseaux de sites web<sup>2</sup>), et le danger subséquent d'une limitation de la recherche sur le social aux traces d'usages en ligne. Roger Bautier, pour sa part, met en cause les dérives quantitatives de ce type de recherche focalisée sur les réseaux (Bautier, 2008), qui privilégie l'analyse des phénomènes directement quantifiables. De manière

---

<sup>1</sup> Une interface de programmation, ou API, conserve sur le web le même principe que sur un ordinateur : elle sert à faciliter le dialogue et l'interopérabilité entre plusieurs programmes. Comme le définit Rieder : « En programmation, une API (*Application Programming Interface*) décrit un ensemble de fonctions qui peuvent être appelées par un autre logiciel. Pour créer une fenêtre sous Microsoft Windows par exemple, un développeur n'est pas obligé de le dessiner pixel par pixel sur l'écran ; il suffit d'appeler la classe *CreateWindow* qui fait partie de l'API de Windows et définit une fenêtre standard avec son comportement habituel » (Rieder, 2008, p. 2).

<sup>2</sup> Sur ce point, voir Ghitalla, 2008.

connexe, plusieurs publications ont récemment mis en avant la nécessité de s'emparer d'objets de recherche souvent ignorés par les SHS, par exemple les algorithmes (Cardon, 2013) ou les bases de données (Flichy et Parasie, 2013). Cet article propose de contribuer à ces débats sur la recherche numérique en s'appuyant sur l'analyse de trois approches spécifiques : les *Digital Humanities*, qui mobilisent les ressources numériques pour la recherche en sciences humaines ; les *Cultural Analytics*, qui analysent les formes numériques de la culture ; et les *Digital Methods*, qui portent spécifiquement sur les formes de la communication sur le web.

La recherche numérique contemporaine est souvent appréhendée à l'aune du caractère innovant de ses apports techniques, les outils informatiques ouvrant de nouveaux horizons de recherche en rendant notamment possible l'analyse de grands corpus de données. Ainsi, l'outil « Google Ngram Viewer » (Michel *et al.*, 2011), qui permet aux chercheurs d'interroger la base de données des livres numérisés par Google pour y extraire des termes (ici les *n-grams*), a fait l'objet de nombreux débats quant à sa nouveauté et à son utilité pour les chercheurs<sup>3</sup>. Au-delà des effets d'annonce, il apparaît plus heuristique d'analyser les usages de ces technologies numériques par les chercheurs en SHS. En ce sens, Noortje Marres propose, dans un cadre théorique des sciences, technologies et société (STS), de regarder les « redistributions de méthodes » qui ont lieu au sein de la recherche numérique (Marres, 2012a). L'auteure entend par ce terme le fait que les outils numériques et les sources de données en ligne actuelles transposent sur le web des méthodes traditionnelles en sciences sociales, tout en les adaptant aux nouvelles propriétés du support (Marres, 2012a, p. 141). Elle propose ainsi l'hypothèse selon laquelle les méthodes utilisées pour l'analyse des objets numériques constituent des adaptations de méthodes déjà existantes, et non des ruptures radicales. Les chercheurs en SHS mobilisent et adaptent donc des méthodes de recherche et les modifient au besoin pour faire face à de nouveaux défis épistémologiques. Cet article se concentre sur trois de ces défis, qui sont illustrés à travers trois approches de recherche numérique, en posant que chacune d'entre elles met en œuvre un type de redistribution spécifique qui pose des questions épistémologiques plus ou moins nouvelles :

- Les *Digital Humanities* interrogent la possibilité d'un dialogue entre sciences exactes et SHS, dont les épistémologies sont traditionnellement séparées ;

---

<sup>3</sup> Débats résumés par COHEN D., 19 décembre 2010, blogue personnel, « Initial thoughts on the Google Books Ngram Viewer and datasets », <<http://www.dancohen.org/2010/12/19/initial-thoughts-on-the-google-books-ngram-viewer-and-datasets/>>, dernière consultation le 20 mars 2014.

## Ouvrir la boîte à outils de la recherche numérique Trois cas de redistribution de méthodes

- Les *Cultural Analytics* interrogent la possibilité de faire de la recherche en SHS à partir de grandes masses de données, à travers l'exemple de l'analyse des productions culturelles ;
- Enfin, les *Digital Methods* interrogent la nécessité de la création de méthodes spécifiques pour appréhender un nouvel objet numérique.

Cet article est organisé en trois temps. Tout d'abord, nous présenterons l'hypothèse de Noortje Marres (*ibid.*) quant à une redistribution des méthodes de recherche numérique. Ensuite, nous mobiliserons cette perspective provenant des STS pour analyser trois types de redistributions à l'œuvre, à travers trois approches de recherche numérique. Enfin, nous mettrons en évidence le potentiel applicatif et réflexif de ces trois types de recherche pour les sciences de l'information et de la communication (SIC).

### 1. Recherche numérique et redistribution de méthodes

Les trois approches de la recherche numérique analysées ici se différencient par leurs objets spécifiques. En effet, les *Digital Humanities* portent sur des objets appartenant au champ des sciences humaines, tout en se concentrant sur leur forme numérique, par exemple les archives en ligne. Les acteurs du champ des *Cultural Analytics* étudient quant à eux les formes numériques de la culture à travers de grandes masses de données, par exemple les catalogues de musées en ligne ; enfin, les recherches en *Digital Methods* portent uniquement sur des « objets nativement numériques », c'est-à-dire des « objets, contenus, applications et environnements qui sont 'nés' sur le support [*qu'est le web*], par opposition à ceux qui y ont migré »<sup>4</sup> (Rogers, 2009, p. 5), par exemple *Wikipédia* ou les médias sociaux. En plus de leurs objets de recherche spécifiques, ces trois approches se différencient par leur positionnement institutionnel. En effet, les *Digital Humanities* désignent un ensemble de projets de recherche qui se sont dotés de relais institutionnels (par exemple des sociétés scientifiques), de structures de publication et de conférences, et s'appuient sur une large communauté de chercheurs. De leur côté, les *Digital Methods* et les *Cultural Analytics* sont d'une taille bien plus modeste et restent liées à des personnalités scientifiques spécifiques (le travail de Richard Rogers

---

<sup>4</sup> « (...) the objects, content, devices and environments 'born' in the new medium [*of the web*], as opposed to those which have 'migrated' to it » (Rogers, 2009, p. 5).

pour la première approche, celui de Lev Manovich pour la seconde). Toutefois, si ces trois approches se distinguent par leurs objets de recherche et leur structure, elles peuvent présenter des similarités au niveau des méthodes employées, comme les outils de visualisation ou de traitement de grandes masses de données.

Cet article ne prétend pas décrire de manière exhaustive l'ensemble de la recherche numérique. D'autres approches ont ainsi recours à des outils numériques ou à de grandes masses de données, comme les « *data-intensive science* » (Hey *et al.*, 2009), les « *computational social science* » (Lazer *et al.*, 2009), ou encore les « *data-driven science* » (Barabási, 2011). Pour leur part, les trois approches présentées dans cet article ont été choisies pour leurs apports heuristiques sur des points méthodologiques particulièrement débattus, qui peuvent par la suite être généralisés à tout le champ de la recherche numérique. En quoi chacune de ces approches interroge les conditions de réalisation de la recherche numérique ? Quels glissements méthodologiques y sont à l'œuvre ?

En suivant l'approche de Noortje Marres (2012a), on peut analyser les redistributions des méthodes qui ont lieu dans la recherche numérique. Cette hypothèse s'inscrit dans le champ des sciences, technologies et société, qui prend en considération l'ensemble des sujets et des objets qui participent au travail scientifique :

« [*Le champ des STS*] met en avant le fait que la recherche scientifique regroupe les contributions d'un large panel de contributeurs : chercheurs, objets de recherche, financeurs, fournisseurs de matériels scientifiques, créateurs d'infrastructures, amateurs intéressés, etc. La recherche scientifique, selon cette notion, doit être considérée comme l'accomplissement partagé d'un ensemble divers d'acteurs »<sup>5</sup> (Marres, 2012a, p. 140).

Cette perspective horizontale permet à l'auteure de partir des bases matérielles du travail scientifique afin de voir « comment la numérisation peut affecter et renseigner sur la *conduite* de la recherche en sciences sociales »<sup>6</sup> (*ibid.*, p. 147). Regarder la recherche numérique en train de se faire nécessite donc de prendre en compte le travail des chercheurs, mais également leur

---

<sup>5</sup> « [*The field of STS*] highlights how scientific research tends to involve contributions from a broad range of actors: researchers, research subjects, funders, providers of research materials, infrastructure builders, interested amateurs, and so on. Scientific research, according to this notion, must be understood as a shared accomplishment of a diverse set of actors. »

<sup>6</sup> « (...) how digitization may affect and inform the *conduct* of social research »

## Ouvrir la boîte à outils de la recherche numérique Trois cas de redistribution de méthodes

positionnement institutionnel, les compétences des membres des équipes de recherche, les logiciels utilisés, les outils développés, etc., pour voir si un changement dans la chaîne méthodologique entre ces différents éléments a lieu. Marres affirme ainsi que les outils numériques mobilisés dans le cadre de ces travaux « contiennent » déjà des « manières de faire » la science, héritées de précédentes traditions méthodologiques. Par exemple, écrire un script pour extraire un terme d'un ensemble de pages HTML suppose un traitement quantitatif de la langue, qui s'oppose par exemple à une analyse par son inscription dans la textualité de la page (Jeanneret et Souchier, 1999). Interroger les méthodes inscrites dans les outils est donc à la fois important pour expliciter la chaîne du travail scientifique, mais également pour circonscrire de manière réaliste la zone de pertinence de ces outils.

Alors que Noortje Marres réduit son analyse à la question des *Digital Methods*, cette hypothèse nous paraît également pertinente pour éclairer les débats épistémologiques qui ont lieu dans les trois approches de la recherche numérique présentées ici ; elle rejoint ainsi un ensemble de travaux émergents sur l'analyse de la production et de l'usage des données numériques de la recherche (Mabi, Plantin et Monnoyer-Smith, sous presse).

### **2. Digital Humanities : interroger le croisement entre épistémologies**

Les *Digital Humanities* recouvrent un ensemble de projets de recherche hétérogènes. Si une définition unique est difficile à produire<sup>7</sup>, il est néanmoins possible d'en résumer ainsi l'objectif : ce champ de recherche vise à croiser les possibilités du traitement informatique avec le travail de recherche en sciences humaines. Cette collaboration commence en 1949, lorsque le jésuite italien Roberto Busa initie une collaboration avec IBM afin de réaliser une fouille de textes de grande ampleur sur le corpus numérisé de Saint Thomas d'Aquin (*l'Index Thomisticus*). Cette collaboration a défini ce qui allait devenir le champ des « *humanities computing* » : dans ce cas, les informaticiens apportent leurs compétences computationnelles, c'est-à-dire ici informatiques, à un travail piloté par des chercheurs en sciences humaines. Les « *humanities computing* » ont produit des outils de fouille de données textuelles et picturales, notamment à travers le développement de standards communs pour la normalisation du

---

<sup>7</sup> Le site « Day of DH », qui organise la journée annuelle des *Digital Humanities*, recense près de 300 définitions accumulées au fur et à mesure des différentes éditions de cet événement. Archive des anciennes éditions : [TAPOR, 2012, Main page, <http://tapor.ualberta.ca/taporwiki/index.php/How\\_do\\_you\\_define\\_Humanities\\_Computing\\_-\\_Digital\\_Humanities%3F>](http://tapor.ualberta.ca/taporwiki/index.php/How_do_you_define_Humanities_Computing_-_Digital_Humanities%3F), dernière consultation le 20 mars 2014.

codage des données<sup>8</sup>, des infrastructures<sup>9</sup> et, enfin, des sociétés scientifiques<sup>10</sup> consacrées à la diffusion de ces croisements entre informatique et SHS.

Plus récemment, le passage du terme « *Humanities computing* » à celui de « *Digital Humanities* » constitue une tentative de dépassement de la séparation entre chercheurs et techniciens, les premiers se présentant comme la tête pensante et considérant les seconds comme le personnel exécutant. Les *Digital Humanities* s'attachent à montrer que ces deux rôles sont complémentaires dans le processus de recherche, davantage que successifs, et que cette collaboration peut être fructueuse. Enfin, le terme vise à susciter une reconnaissance de ce champ de recherche au sein des universités (Kirschenbaum, 2012b).

La phase la plus récente des travaux en *Digital Humanities* prend en compte les modalités de création, de partage et de publication de données sur le web. Presner résume ainsi le passage à cette tendance récente des *Digital Humanities* :

« Les '*Digital Humanities 2.0*' sont un champ fortement génératif, créent des environnements et des outils pour la production, la curation et l'interaction avec la connaissance sous forme numérique et provenant de plusieurs contextes. Alors que la première vague des *Digital Humanities* était concentrée, peut-être de manière un peu restreinte, sur l'analyse de textes (à travers les systèmes de classification, les langages de balisages, le codage de texte et l'édition scientifique) au sein de disciplines établies, les *Digital Humanities 2.0* apportent des paradigmes disciplinaires entièrement nouveaux, une convergence entre plusieurs champs, des méthodologies hybrides, et même de nouveaux modèles de publication qui ne sont souvent pas dérivés, ni limités à la culture de l'imprimé »<sup>11</sup> (Presner, 2010, p. 6).

---

<sup>8</sup> Comme le montrent les travaux de la *Text Encoding Initiative* ou plus récemment, de la *Text-Image Linking Environment*.

<sup>9</sup> Par exemple le Très Grand Équipement (TGE) Adonis en France, désormais Huma-Num.

<sup>10</sup> La Modern Language Association (MLA), l'Association for Computers and the Humanities (ACH), l'Association for literary and Linguistic Computing (ALLC) et la structure regroupant ces deux dernières: ADHO (Alliance of Digital Humanities Organizations) (Kirschenbaum, 2012a, p. 6-7).

<sup>11</sup> « 'Digital Humanities 2.0' – is deeply generative, creating the environments and tools for producing, curating, and interacting with knowledge that is 'born digital' and lives in various digital contexts. While the first wave of Digital Humanities concentrated, perhaps somewhat narrowly, on text analysis (such as classification systems, mark-up, text encoding, and scholarly editing) within

## Ouvrir la boîte à outils de la recherche numérique Trois cas de redistribution de méthodes

À la suite des travaux précurseurs en « *Humanities computing* », les « *Digital Humanities* » visent ainsi à mobiliser les ressources numériques en termes de calcul et de publication dans le cadre de recherches en sciences humaines.

Le projet *Mapping the Republic of Letters* (cf. figure 1) de l'Université de Stanford constitue un exemple d'application d'outils numériques dans le cadre d'une recherche dans le domaine des humanités. Ce projet se fonde sur la numérisation de la correspondance des intellectuels européens des Lumières, puis sur le traitement informatique de cette base de données en fonction de critères déterminés par le travail de recherche, comme ci-dessous, afin de visualiser les relations entre Jean-Jacques Rousseau et ses correspondants européens.

---

established disciplines, Digital Humanities 2.0 introduces entirely new disciplinary paradigms, convergent fields, hybrid methodologies, and even new publication models that are often not derived from or limited to print culture » (Presner, 2010, p. 6).



Figure 1. « Mapping Republic of letters », *Digital Humanities Stanford*<sup>12</sup>.

Ce projet de recherche illustre le croisement entre deux traditions épistémologiques différentes : l'histoire des textes et des idées, réalisée à partir de l'analyse des matériaux littéraires et des contextes socio-historiques (méthode également appelée « *close reading* ») *versus* une analyse statistique des flux de correspondances entre auteurs (relevant du « *distant reading* », cf. Moretti, 2008). Alors que la première approche se fonde sur une tradition propre aux humanités (croisant des ressources en littérature, en histoire, voire en sociologie), la seconde relève de l'analyse de réseaux sociaux (ou ARS, cf. Rieder, 2007) et des statistiques.

Or, Johanna Drucker reproche aux visualisations précédentes d'appliquer une conception anachronique du transport postal à la correspondance des

<sup>12</sup> STANFORD UNIVERSITY, 2010, *Voltaire and the Enlightenment*, <<http://republicofletters.stanford.edu/casestudies/voltaire.html>>, dernière consultation le 20 mars 2014.

## Ouvrir la boîte à outils de la recherche numérique Trois cas de redistribution de méthodes

Lumières : les flux de données y ressemblent davantage aux flux aéropostaux contemporains et ne reflètent pas la complexité du système d'acheminement du courrier de l'époque (Drucker, 2010). Les historiens s'approprient ici un peu rapidement des outils statistiques sans véritable réflexion sur les catégories d'analyse employées, et sur la représentation visuelle des résultats. Ce travail illustre alors le danger pointé par Rieder et Rhöle (2012) : les outils de traitement mathématique des phénomènes sociaux s'inscrivent dans un fantasme de l'objectivité dont peuvent se servir les humanistes afin de tordre le cou aux accusations de subjectivité dont ils sont la cible, en se dotant d'outils censés garantir l'objectivité de leur recherche, tout en réglant au passage un complexe d'infériorité par rapport aux sciences exactes (Rieder et Rhöle, 2012).

Si Johanna Drucker met en avant les limites de ce projet, elle affirme toutefois que les travaux des *Digital Humanities* constituent autant de tentatives pour les faire dialoguer (Drucker, 2011). Ainsi, ces travaux tentent-ils de croiser l'approche « subjective » de l'objet de recherche dans les sciences humaines – le scientifique faisant partie de son objet de recherche – avec la connaissance « objective », car davantage distanciée, des faits étudiés dans les sciences exactes, ce que l'auteure nomme l'« *aesthesis* » des sciences humaines (c'est-à-dire la connaissance partielle, située et subjective), et la « *mathesis* » des sciences exactes (recouvrant les « prétentions à l'objectivité » (Drucker, 2009, p. XIII)). Pour que l'« *aesthesis* » et la « *mathesis* » convergent, cela nécessite toutefois une posture de réflexivité de la part des chercheurs par rapport aux technologies utilisées. Ces précautions ont déjà été soulignées en histoire. La rencontre entre les ordinateurs et la discipline historique est en effet bien antérieure au web : il suffit, pour s'en convaincre, de rappeler la célèbre citation d'Emmanuel Le Roy Ladurie en 1968, pour qui « l'historien de demain sera programmeur ou il ne sera plus » (Le Roy Ladurie, 1968). Le projet *Mapping the Republic of Letters* (cf. figure 1) constitue dès lors une nouvelle étape de ce dialogue entre histoire et statistiques à partir de nouvelles technologies, tout en soulignant les précautions à adopter dans ce dialogue.

### **3. Cultural Analytics : interroger le traitement des grandes masses de données**

De son côté, le champ des *Cultural Analytics* se fonde largement sur les travaux de Lev Manovich. Suite à son ouvrage *The Language of New Media* (2001), il a ainsi contribué à la création du champ de recherche des « *Software studies* ». Il est également à l'origine de la « *Software Studies Initiative* » (reliée

à l'école d'ingénieur *Calit2* à l'UCSD<sup>13</sup> et, plus récemment, au *Graduate Center*, à la CUNY<sup>14</sup>). Le travail de Manovich agrège différentes traditions disciplinaires : l'histoire de l'art, les études visuelles et cinématographiques d'un côté, et l'analyse des réseaux sociaux, les sciences informatiques et l'étude des logiciels (« *software studies* ») de l'autre.

Les recherches de la *Software Studies Initiative* portent sur « les formes de la culture médiatisée par le numérique »<sup>15</sup> (Manovich, 2007). Le terme de « culture » est entendu ici au sens large, et comprend à la fois les arts visuels traditionnels (que l'on retrouve dans son analyse de l'évolution stylistique de Mondrian et de Rothko<sup>16</sup>) ainsi que les formes contemporaines de la culture populaire (comme l'analyse de la composition picturale des mangas, cf. figure 2). Toutefois, ces deux champs de recherche sont analysés à travers leur manifestation numérique, résultant soit de la numérisation d'archives, soit de l'utilisation de données numériques en ligne.

Manovich différencie ses travaux des *Digital Humanities* en mettant en avant des méthodes informatiques spécifiques (Manovich, 2007) : ses travaux mettent davantage l'accent sur l'analyse visuelle que sur l'analyse textuelle ; ils portent sur la culture populaire en ligne et non plus uniquement sur les canons littéraires ; enfin, ils s'appuient sur une méthodologie de fouille de données plus poussée que celle des *Digital Humanities*, en procédant notamment à un travail sur les métadonnées (Manovich, 2007).

Des trois champs de recherche numérique présentés ici (à savoir les *Digital Humanities*, les *Cultural Analytics* et les *Digital Methods*), les travaux de la *Software Studies Initiative* sont ceux qui se rapprochent le plus nettement du traitement algorithmique de grandes masses de données, également appelées « *Big data* ». L'étude de la culture selon Manovich passe ainsi par le recours à différents outils et méthodes développés en dehors des SHS, par exemple l'analyse exploratoire, la fouille et la visualisation de grandes masses de données. Celles-ci sont appliquées sur différents types de corpus :

- les sources de données provenant du concours « *Digging Into Data* »<sup>17</sup> ;

---

<sup>13</sup> University California, San Diego.

<sup>14</sup> City University of New York.

<sup>15</sup> « digitally-mediated forms of culture ».

<sup>16</sup> MANOVICH L., 2011, « Mondrian vs Rothko: footprints and evolution in style space », *Software Studies Initiative*, 29 juin 2011, <<http://lab.softwarestudies.com/2011/06/mondrian-vs-rothko-footprints-and.html>>, dernière consultation le 20 mars 2014.

<sup>17</sup> Relevant des organismes de financement suivant : la National Endowment for the Humanities (NEH) et la National Science Foundation (NSF).

### Ouvrir la boîte à outils de la recherche numérique Trois cas de redistribution de méthodes

- les données de musées qui ont rendu leur fonds accessible en ligne, par exemple sous la forme d'une API ;
- les API d'applications en ligne, telles que *Flickr*<sup>18</sup>.

Le traitement de ces données passe par la création d'outils de traitement d'image numérique et de logiciels de visualisation de médias<sup>19</sup> développés par la *Software Studies Initiative*, dont le plus célèbre se nomme *Imageplot*<sup>20</sup> (cf. figure 2). Il permet de visualiser de grands corpus de données, ici un million de pages de mangas, à partir de leurs caractéristiques visuelles.

---

<sup>18</sup> MANOVICH L., 2013, « Course schedule: Big Data, Visualization, and Digital Humanities. Spring GC CUNY », <<https://docs.google.com/document/d/1Mf6qIFNqwLLvVPIWAOy0eK3HfbgLLJgIJSyElmM4uSg/edit>>, dernière consultation le 20 mars 2014.

<sup>19</sup> SOFTWARE STUDIES INITIATIVE, « Software for digital humanities », <<http://lab.softwarestudies.com/p/software-for-digital-humanities.html>>, dernière consultation le 20 mars 2014.

<sup>20</sup> SOFTWARE STUDIES INITIATIVE, « ImagePlot visualization software: explore patterns in large image collections », <<http://lab.softwarestudies.com/p/imageplot.html>>, dernière consultation le 20 mars 2014.

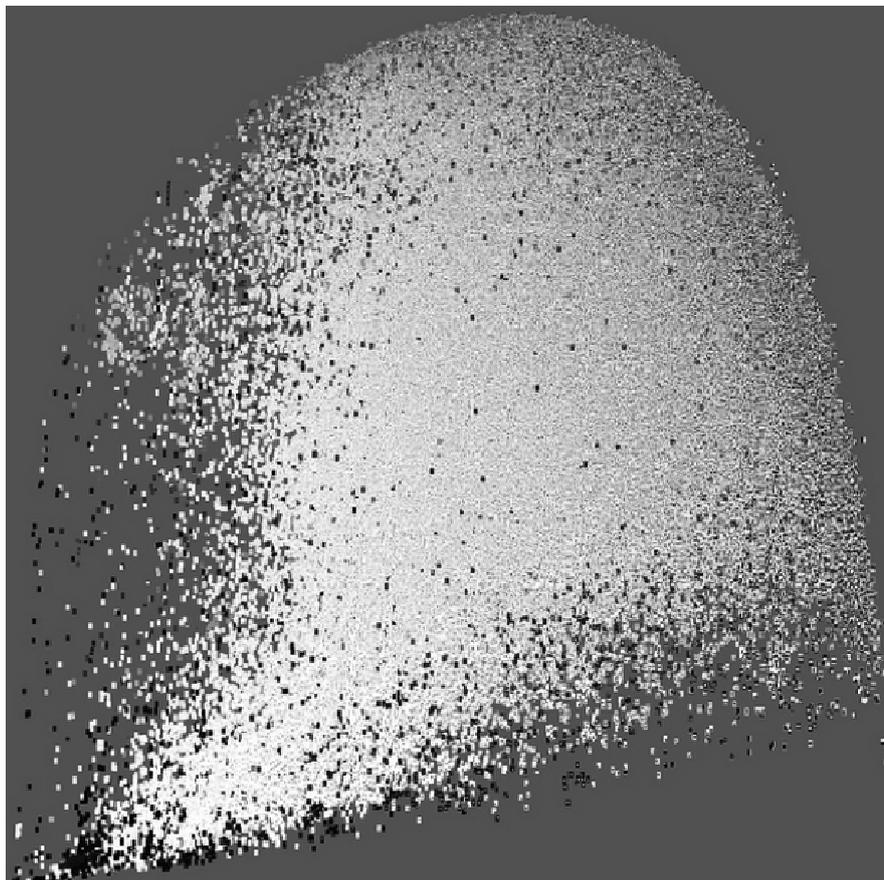


Figure 2. Le projet « 1 million Manga pages »<sup>21</sup>.

La figure ci-dessus est extraite du projet de recherche « 1 million Manga pages »<sup>22</sup>, proposant une analyse stylistique des bandes dessinées japonaises. Ce projet se fonde sur l'extraction de la totalité de la base de données d'un site web de publication de mangas, puis de leur visualisation et analyse. Cette figure

---

<sup>21</sup>FLICKR, « Software studies.com », <<http://www.flickr.com/photos/culturevis/4497385883/in/set-72157623280860377>>, dernière consultation le 20 mars 2014.

<sup>22</sup> MANOVICH L., 2010, « One million manga pages », *Software Studies initiative*, 14 novembre 2010, <<http://lab.softwarestudies.com/2010/11/one-million-manga-pages.html>>, dernière consultation le 20 mars 2014.

## Ouvrir la boîte à outils de la recherche numérique Trois cas de redistribution de méthodes

visualise ainsi l'ensemble des pages de ces bandes dessinées. L'axe horizontal présente l'écart-type (ici, les tonalités de gris) et l'axe vertical présente l'entropie (ici, la variation dans le détail de l'image).

Ce projet de recherche met en avant trois types de débats concernant l'exploitation de grandes quantités de données dans la recherche en SHS. Tout d'abord, l'étude de grandes quantités de productions culturelles en ligne nécessite le recours à des masses de données qui peuvent relever de sources privées, telles que les musées ou les entreprises du web (par exemple *Facebook*, *Amazon* ou *Netflix*). L'accès à ces données pour la recherche peut être rendu difficile du fait de la logique commerciale de protection des données. Le risque est alors grand de voir émerger un nouveau type de fracture numérique (Boyd et Crawford, 2012) entre grandes et petites structures de recherche : les chercheurs disposant de contrats de recherche avec des opérateurs privés leur garantissant un accès aux données, ou travaillant directement dans des centres de recherche privés (tels que Microsoft Research ou Orange Labs) auraient alors un avantage comparatif certain par rapport aux centres de recherche plus modestes. Plus largement, on peut même aller plus loin en s'interrogeant sur la représentativité des données mises à la disposition des chercheurs dans la mesure où les entreprises du web qui les mettent en ligne imposent régulièrement leurs conditions d'utilisation, en réglant par exemple la proportion des données accessibles ou le nombre et le rythme des requêtes envers leur API.

Ensuite, les recherches sur les grandes masses de données interrogent les modalités de partage des données entre chercheurs et leur utilisation, ces données devant *a minima* être réutilisables suivant un objectif de validation/vérification des travaux publiés. Ce débat, déjà amorcé en sciences exactes du fait d'une longue tradition de manipulation de grands ensembles de données qui a conduit à la mise en place d'infrastructures d'archive et de partage de données<sup>23</sup>, se pose aujourd'hui en SHS. La création de données s'y effectue encore souvent à l'échelle d'équipes de recherche, dont les membres ne tendent que rarement à partager leurs données une fois leur recherche terminée<sup>24</sup>. Toute une littérature sur l'étude des infrastructures de recherche a mis en évidence, à travers plusieurs enquêtes ethnographiques portant sur la constitution de bases de données scientifiques (Star et Ruhleder, 1996 ; Star,

---

<sup>23</sup> On pourra citer le *Sloan Digital Sky Survey*, pour les données en astronomie.

<sup>24</sup> On citera sur ce point des exceptions notables, tels que l'ICPSR (Inter-University Consortium for Political and Social Research), qui maintient une infrastructure de 500 000 archives en sciences humaines et sociales, comprenant par exemple des données administratives, des résultats d'étude ou de sondages.

1999), les défis rencontrés pour garantir les conditions de partage des bases de données, aussi bien au niveau du travail « des petites mains » sur les métadonnées (Dagiral et Peerbaye, 2012) que sur les modalités de pérennisation de la base de données (Millerand, 2011 ; 2012). Appliquées aux projets de recherche numérique, ces notions permettent de mettre en avant tout le travail de structuration des données nécessaire à leur traitement, autant d'étapes qui sont rarement explicitées lors de la communication des résultats.

Dernier débat : le traitement de masses de données en SHS appelle également une réflexion sur la structuration de ces données. Paul Edwards a ainsi montré que de nombreuses « frictions » jalonnent le processus de création des données (Edwards, 2010). Imprécisions, agrégations de données différentes, manque de robustesse, etc. échappent en effet régulièrement aux scientifiques soucieux d'apporter des preuves irréfutables dans le cadre de leur démonstration. Il est alors nécessaire, suivant en cela Drucker, de retracer ces « frictions » et leurs conséquences sur le processus scientifique. Partant du principe qu'il n'existe pas de données brutes (Gitelman, 2013) et que les données de la recherche sont toujours déjà une représentation construite du phénomène analysé (Kurgan, 2013), les multiples étapes de leur transformation, de leur extraction à leur utilisation, doivent être mises en avant pour en souligner la nature construite. Si les travaux de Manovich et de son équipe constituent un exemple de documentation du travail de recherche en la matière, puisqu'il retrace notamment la démarche scientifique s'agissant de l'accès aux données, du traitement et de l'analyse, ce dans un but pédagogique. Cette réflexivité est cependant loin d'être généralisée.

#### 4. *Digital Methods* : interroger la séparation entre anciennes et nouvelles méthodes

Les *Digital Methods* s'appuient quant à elles sur les travaux de Richard Rogers (2000 ; 2004) : le premier domaine d'application des *Digital Methods* a ainsi été l'analyse de controverses en ligne en ayant recours à la visualisation des relations entre les acteurs par leurs hyperliens. Cette méthode a par exemple été appliquée au cas d'étude du réchauffement climatique (Rogers et Marres, 2000), au conflit israélo-palestinien (Rogers et Ben David, 2008), avant de se porter sur plusieurs autres technologies de communication en ligne, comme les API (ou « interfaces de programmation » cf. note 1), les plateformes en ligne ou les boutons de « plussioiement » (par exemple « J'aime » pour Facebook). Ces recherches se différencient des deux approches précédentes par la restriction du domaine d'étude au web. En effet, les membres de la *Digital Methods Initiative* de l'université d'Amsterdam n'étudient pas les objets numérisés et importés sur le web, mais la manière dont les « objets nativement numériques » (Rogers, 2009), tels que ceux résumés ci-dessous par Rogers, structurent les pratiques de communication en ligne :

- L'hyperlien
- Le site web
- Les moteurs de recherche et les sphères
- Les réseaux
- Les réseaux sociaux et les mesures « post-démographiques »
- Wikipédia et les contenus en réseau (Rogers, 2009)<sup>25</sup>.

Le positionnement méthodologique des recherches menées par la *Digital Methods Initiative* est résumé par la formule suivante : « suivre le support »<sup>26</sup> (Rogers, 2009). Il s'agit dans ce cas de préserver les propriétés du milieu analysé, ici le web, en créant des outils fondés sur les objets et applications en ligne. Cette méthode se traduit par la reconfiguration des objets étudiés (« repurposing », cf. Rogers, 2009) pour prendre en considération les propriétés du support numérique. Contrairement aux « *virtual methods* » (Hine, 2005) qui appliquent des méthodes traditionnelles à des terrains de recherche en ligne<sup>27</sup>,

---

<sup>25</sup> The link  
The website  
The search engines & the spheres  
The webs  
Social networking sites & post-demographics  
Wikipedia and networked content (Rogers, 2009).

<sup>26</sup> « follow the medium » (Rogers, 2009).

<sup>27</sup> Par exemple, les entretiens par courriel ou l'observation participante.

les *Digital Methods* proposent d'ajuster la méthode à l'objet, nécessitant dès lors une méthodologie « sur-mesure » (Rogers, 2009), c'est-à-dire qui s'adapte à chaque objet étudié. Cette méthodologie se différencie des *Digital Humanities*, qui utilisent le web comme extension des travaux en humanités, et non comme objet d'investigation en tant que tel. Les *Digital Methods* se distinguent également des travaux de Manovich dans la mesure où la *Software Studies Initiative* défend un accès négocié aux données, en faisant par exemple des démarches auprès des musées. Les membres de la *Digital Methods Initiative*, à l'inverse, utilisent des méthodes qui passent par l'infrastructure du web, en développant par exemple des outils d'extraction de textes ou en passant par les API des services en ligne étudiés (Marres et Weltevrede, 2013).

Le projet « *What is Digital Humanities ?* » (cf. figure 3) illustre l'approche des *Digital Methods*. Il vise à étudier le site internet Amazon en tant qu'acteur du web qui intervient, à travers son algorithme de recommandation, dans la dissémination d'ouvrages portant sur les *Digital Humanities*. Toutefois, étant donné que cet algorithme de recommandation n'est *a priori* pas destiné à la recherche, il s'agit de reconfigurer ses propriétés pour en faire un objet d'étude. La méthode adoptée est donc la création d'un « script » qui va interroger l'API d'Amazon afin de récolter les données de suggestion de livres sur les *Digital Humanities*, qui seront ensuite visualisées et analysées (Berry *et al.*, sous presse).

Ouvrir la boîte à outils de la recherche numérique  
Trois cas de redistribution de méthodes

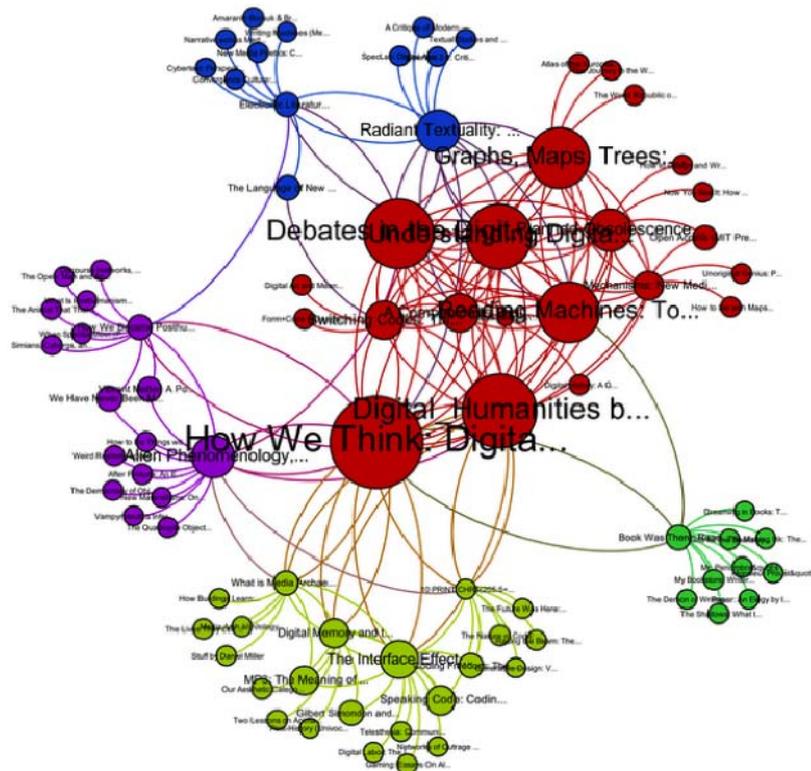


Figure 3. Projet « What is Digital Humanities ? » (Berry et al., sous presse)<sup>28</sup>.

Le graphe ci-dessus est une visualisation des liens de recommandations sur Amazon.com, récoltés à partir d'une sélection de dix livres de référence en *Digital Humanities*. Les paramètres de lecture sont les suivants : la taille des nœuds varie en fonction du nombre de liens entrants (c'est-à-dire que plus un livre – représenté par un nœud – est gros, plus il est suggéré) et les couleurs font ressortir les groupes de livres qui sont reliés par des liens de suggestion. Il est alors à la fois possible d'y voir les livres qui sont le plus souvent

<sup>28</sup> Voir : BERRY D. M., E. BORRA, A. HELMOND, J.-C. PLANTIN et J.-W. RETTBERG, *What is Digital Humanities ?*, 2013, <<https://sites.google.com/site/whatisdigitalhumanities/results>>, dernière consultation le 20 mars 2014.

recommandés par la plateforme Amazon, mais également à quel champ disciplinaire ces livres suggérés appartiennent. Par exemple, les livres appartenant au champ de la philosophie de la technique (en mauve à gauche) ont-ils tendance à recommander l'ouvrage *How we think: Digital media and contemporary technogenesis* de Katherine Hayles.

À travers ce projet relevant des *Digital Methods* – même s'il porte dans ce cas d'étude sur le champ des *Digital Humanities* – ce sont les modalités de la recherche à partir de boîtes noires qui sont interrogées. En effet, l'étude de la place de l'algorithme d'Amazon dans la dissémination d'ouvrages sur les *Digital Humanities* doit nécessairement passer par l'interrogation de son API, ce qui nécessite ainsi d'accepter les conditions d'accès du service à ces données. Or, tous les paramètres ne sont pas toujours clairement explicités : à propos de Twitter, Boyd et Crawford (2012) soulignent ce flou sur la proportion des données auxquelles on peut accéder lors de l'interrogation de l'API. En plus de cette opacité dans l'accès aux données, ce sont également les compétences des chercheurs pour mobiliser ces méthodes qui doivent être pensées. Les outils relevant du champ de l'analyse des réseaux sociaux numériques (Rieder, 2007), mobilisés dans le projet précédent (cf. figure 3), nécessitent effectivement, pour être utilisés, une connaissance pointue des statistiques, ce qui suppose un investissement fort de la part des chercheurs en SHS. D'un autre côté, la constitution d'équipes de recherche pluridisciplinaires, susceptibles de couvrir tout le panorama des compétences requises, nécessite des financements importants (Plantin, sous presse).

Si les *Digital Methods* courent le risque de faire de la science à partir de « boîtes noires », Noortje Marres (2012a) met également en garde contre la possibilité d'en créer de nouvelles. Si l'auteure reconnaît l'importance de « suivre le support », c'est-à-dire de développer des méthodes de recherche à partir des caractéristiques de l'objet de recherche, ce positionnement peut résulter en un manque de réflexivité par rapport aux objets étudiés. L'étude présentée en figure 3 porte au moins autant sur les *Digital Humanities* que sur la plateforme Amazon elle-même, avec pour inconvénient que c'est Amazon qui fixe les conditions d'accès aux chercheurs pour l'analyse, ici à travers son API. Il y a donc un fort risque « d'aveuglement » inhérent aux *Digital Methods*, (Marres, 2012b) : un effort de déconstruction réflexive des catégories de l'analyse s'avère nécessaire.

## 5. Les apports pour les sciences de l'information et de la communication

Les trois approches de recherche numérique présentées dans cet article – *Digital Humanities*, *Cultural Analytics*, et *Digital Methods* – portent chacune sur un objet précis, respectivement les sciences humaines, les grandes masses de données culturelles et les objets nativement numériques. Chacune de ces approches illustre la manière dont les SHS peuvent mobiliser les outils numériques ; de plus, chacune organise une forme de redistribution de la compétence technique au sein des collectifs de recherche : comment mobiliser des outils relevant d'épistémologies différentes ? Comment traiter les grandes masses de données en SHS ? Quelles méthodes adopter en fonction des spécificités de l'objet numérique ?

Ces méthodes émergentes en recherche numérique peuvent être fructueuses pour plusieurs champs d'étude en sciences de l'information et de la communication. On pense notamment aux mutations de la participation politique en ligne du fait du mouvement de libération des données publiques (*Open data*) (Plantin et Valentin, 2013) ; aux mutations du journalisme face aux nouveaux outils numériques et à l'inflation des ressources en ligne (le journalisme des données) (Parasie et Dagiral, 2013) ou encore à la place de la technologie dans la ville et dans les musées. Les apports de ces méthodes numériques peuvent prendre la forme d'un enrichissement des méthodes existantes. Pluri-méthodologique, ce positionnement permet de développer une approche à plusieurs échelles pour l'analyse des pratiques de communication en ligne, en passant de l'analyse des propriétés de la médiation (« *close-reading* ») à une prise en compte, plus macro, des propriétés des phénomènes étudiés (« *distant reading* »).

En outre, si ces approches présentent un potentiel heuristique pour les recherches en communication, elles renouvellent également, chacune à leur manière, des questionnements sur les types de redistribution qui ont lieu dans le processus de production scientifique : les différentes approches interrogent la place des outils, des données et des méthodes dans le cadre d'analyse des pratiques de communication. Selon une démarche réflexive, ces trois approches constituent donc non seulement des exemples de recherche, mais font aussi l'objet d'une interrogation sur la recherche numérique.

De plus, un regard sur ce tournant computationnel (Berry, 2011), qui désigne la multiplication d'outils et de méthodes numériques en SHS, permet d'analyser les évolutions conjointes des pratiques sociales et des dispositifs numériques

sur lesquels elles s'appuient pour se reconfigurer, sans négliger néanmoins la complexité des interactions qui se construisent et les conséquences sociopolitiques et sociotechniques qui en découlent. Il est ainsi possible à travers ces recherches de faire tenir dans l'analyse à la fois les propriétés du support de communication et les pratiques des acteurs. Enfin, la dimension interdisciplinaire, à la base des projets de recherche numérique décrits ici fait écho à l'interdisciplinarité inhérente aux sciences de l'information et de la communication (Ollivier, 2000 ; Bougnoux, 2002). En effet, les travaux évoqués pour ces trois approches regroupent une diversité de profils de chercheurs qui, appartenant à des disciplines variées, analysent les pratiques de communication en ligne.

Pour autant, on doit s'interroger sur les conséquences de la mobilisation de ces méthodes dans leurs domaines d'application respectifs, mais également, plus largement, sur les conséquences de la mobilisation de méthodes de recherche numérique en SIC. Assiste-t-on à l'émergence d'un champ de recherche dans lequel les hypothèses sont tributaires de la nature des données disponibles et des outils pour les traiter ? Les SIC présentent les ressources nécessaires pour prendre en compte les multiples dimensions de la place de la technique, mais également pour interroger le rôle des méthodes tout comme les étapes d'acquisition et de traitement des données numériques.

## Références bibliographiques

- BARABÁSI A. L., 2011, « The Network Takeover », *Nature Physics*, vol. 8, n° 1, pp. 14-16.
- BAUTIER R., 2008, « L'impérialisme des statistiques de réseaux », *Médiation & information*, n° 28, pp. 129-138.
- BERRY D. M., 2011, « The Computational Turn: thinking about the Digital Humanities », *Culture Machine*, n° 12, pp. 1-22, <<http://www.culturemachine.net/index.php/cm/article/view/440/470>>, dernière consultation le 20 mars 2014.
- BERRY D. M., E. BORRA, A. HELMOND, J.-C. PLANTIN et J.-W. RETTBERG, sous presse, « The Data Sprint Approach: Exploring the field of Digital Humanities through Amazon's Application Programming Interface », *Digital Humanities Quarterly*.
- BOUGNOUX D., 2002, *Introduction aux sciences de la communication*, Paris, La Découverte.

Ouvrir la boîte à outils de la recherche numérique  
Trois cas de redistribution de méthodes

- BOYD D. et K. CRAWFORD, 2012, « Critical Questions for Big Data », *Information, Communication and Society*, vol. 15, n° 5, pp. 662-679.
- CARDON D., 2013, « Dans l'esprit du PageRank », *Réseaux*, n° 177, vol. 1, pp. 63-95.
- COHEN D., 19 décembre 2010, « Initial thoughts on the Google Books Ngram Viewer and datasets », blogue personnel, <<http://www.dancohen.org/2010/12/19/initial-thoughts-on-the-google-books-ngram-viewer-and-datasets/>>, dernière consultation le 20 mars 2014.
- DAGIRAL É. et A. PEERBAYE, 2012, « Les mains dans les bases de données : connaître et faire reconnaître le travail invisible », *Revue d'anthropologie des connaissances*, Vol. 6, n° 1, pp. 191-216.
- DRUCKER J., 2009, *SpecLab: Digital Aesthetics and Projects in Speculative Computing*, Chicago, University Of Chicago Press.
- DRUCKER J., 2010, *Humanistic Approaches to the Graphical Expression of Interpretation*, Cambridge, MA, États-Unis, MIT Video, <<http://video.mit.edu/watch/humanistic-approaches-to-the-graphical-expression-of-interpretation-9596/>>, 91 min, 31 s, dernière consultation le 20 mars 2014.
- DRUCKER J., 2011, « Humanities Approaches to Graphical Display », *Digital Humanities Quarterly*, vol. 5, n° 1, <<http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html>>, dernière consultation le 20 mars 2014.
- EDWARDS P. N., 2010, *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*, Cambridge, MIT Press.
- FLICHY P. et S. PARASIE (dir.), 2013, « Sociologie des bases de données », *Réseaux*, n° 178-179, vol. 2-3.
- GERLITZ C. et B. RIEDER, 2013, « Mining One Percent of Twitter: Collections, Baselines, Sampling », *M/C Journal*, vol. 16, n° 2, <<http://journal.media-culture.org.au/index.php/mcjournal/article/view/620>>, dernière consultation le 21 avril 2014.
- GHITALLA F., 2008, « La 'Toile Européenne' Parcours autour d'une cartographie thématique de documents web consacrés au thème de l'Europe et à ses acteurs sur le web francophone », *Communication & langages*, n° 158, pp. 61-75.

- GITELMAN L. (dir.), 2013, « *Raw Data* » *Is an Oxymoron*, Cambridge, MIT Press.
- HEY T., TANSLEY S. et TOLLE K. (dir.), 2009, *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Redmond, Microsoft Research.
- HINE C., 2005, *Virtual Methods*, Oxford, Berg Publishers.
- JEANNERET Y. et E. SOUCHIER, 1999, « Pour une poétique de l'écrit d'écran », *Xoana*, n° 6, pp. 97-107.
- JEANNERET Y., 2013, « Les chimères cartographiques sur l'internet. Panoplie représentationnelle de la 'traçabilité' sociale » dans B. GALINON-MELENEC et S. ZLITNI (dir.), *Traces numériques. De la production à l'interprétation*, Paris, CNRS Éditions.
- KURGAN L., 2013, *Close up at a distance: Mapping, Technology, and Politics*, New York, Zone Books.
- KIRSCHENBAUM M., 2012a, « Digital Humanities As/Is a tactical term », dans M. K. GOLD (dir.), *Debates in the Digital Humanities*, Minneapolis, The University of Minnesota Press, pp. 415-429.
- KIRSCHENBAUM M., 2012b, « What is Digital Humanities and what it is doing in English department? », dans M. K. GOLD (dir.), *Debates in the Digital Humanities*, Minneapolis, The University of Minnesota Press, pp. 3-11.
- LAZER D., A. PENTLAND, L. ADAMIC, S. ARAL, A. BARABÁSI, D. BREWER, N. CHRISTAKIS, N. CONTRACTOR, J. FOWLER, M. GUTMANN, T. JEBARA, G. KING, M. MACY, D. ROY et M. VAN ALSTYNE, 2009, « Computational Social Science », *Science*, vol. 323, n° 5915, pp. 721-723.
- LE ROY LADURIE E., 1968, « La fin des érudits », *Le Nouvel Observateur* (Paris), 8 mai 1968.
- MABI C., J.-C. PLANTIN et L. MONNOYER-SMITH, sous presse, « Interroger les données scientifiques en SHS à partir de leur écosystème », dans V. SCHAFER (dir.), *Information et communication scientifique à l'heure du numérique*, Paris, Essentiel d'Hermès.
- MANOVICH L., 2001, *The Language of New Media*, Cambridge, MIT Press.
- MANOVICH L., 2007, « Cultural Analytics: analysis and visualization of large cultural data sets », pp. 1-18, <[http://www.manovich.net/cultural\\_analytics.pdf](http://www.manovich.net/cultural_analytics.pdf)>, dernière consultation le 20 mars 2014.

Ouvrir la boîte à outils de la recherche numérique  
Trois cas de redistribution de méthodes

- MANOVICH L., 2010, « One million manga pages », Software studies initiative, 14 novembre 2010, <<http://lab.softwarestudies.com/2010/11/one-million-manga-pages.html>>, dernière consultation le 20 mars 2014.
- MANOVICH L., 2011, « Mondrian vs Rothko: footprints and evolution in style space » Software studies initiative, 29 juin 2011, <<http://lab.softwarestudies.com/2011/06/mondrian-vs-rothko-footprints-and.html>>, dernière consultation le 20 mars 2014.
- MANOVICH L., 2013, « Course schedule: Big Data, Visualization, and Digital Humanities.Spring GC CUNY », <<https://docs.google.com/document/d/1Mf6qIFNqwLLvVPIWAOy0eK3HfbgLIJgJJSyElmM4uSg/edit>>, dernière consultation le 20 mars 2014.
- MARRES N., 2012a, « The redistribution of methods: on intervention in digital social research, broadly conceived », *The Sociological Review*, vol. 60, pp. 139-165.
- MARRES N., 2012b, « Digital Societies: between Ontology and Methods », communication à la journée d'étude *Digital Societies: between Ontology and Methods*, 7 mars 2012, Londres, Goldsmiths University of London, <<http://magiclantern.gold.ac.uk/movs/csisp/latour-rogers-questions.mp3>>, dernière consultation le 20 mars 2014.
- MARRES N. et E. WELTEVREDE, 2013, « Scraping the Social? Issues in real-time social research », *Journal of Cultural Economy*, vol. 6, n° 3, pp. 313-335.
- MICHEL J-B., Y. K. SHEN, A. PRESSER AIDEN, A. VERES, M. K. GRAY, W. BROCKMAN, J. P. PICKETT, D. HOIBERG, D. CLANCY, P. NORVIG, J. ORWANT, S. PINKER, M. A. NOWAK et E. LIEBERMAN, 2011, « Quantitative Analysis of Culture Using Millions of Digitized Books », *Science*, vol. 331, n° 6014, pp. 176-182.
- MILLERAND F., 2011, « Le partage des données scientifiques à l'ère de l'e-science: l'instrumentation des pratiques au sein d'un collectif multidisciplinaire », *Terrains & travaux*, vol. 1, n° 18, pp. 215-237.
- MILLERAND F., 2012, « La science en réseau », *Revue d'anthropologie des connaissances*, vol. 6, n° 1, pp. 163-190.
- MORETTI F., 2008, *Graphes, cartes et arbres: modèles abstraits pour une autre histoire de la littérature*, Paris, Les Prairies Ordinaires.

- OLLIVIER B., 2000, *Observer la communication : naissance d'une interdiscipline*, Paris, Éditions CNRS.
- PARASIE S. et E. DAGIRAL, 2013, « Des journalistes enfin libérés de leurs sources ? Promesse et réalité du 'journalisme de données' », *Sur le journalisme*, vol. 2, n° 1, pp. 52-63.
- PLANTIN J.-C., sous presse, « Les chercheurs en SHS rêvent-ils de code informatique ? Trois questions préliminaires à l'apprentissage du code » dans O. LE DEUFF (dir.), *Le temps des humanités digitales*, Limoges, FYP éditions.
- PLANTIN J.-C. et J. VALENTIN, 2013, « Données ouvertes et cartographie libre : autour du cas de Montpellier », *Les Cahiers du Numérique*, vol. 9, n° 1, pp. 85-110.
- PRESNER T., 2010, « Digital Humanities 2.0: A Report on Knowledge », <<http://cnx.org/content/m34246/1.6/?format=pdf>>, dernière consultation le 20 mars 2014.
- REBILLARD F., 2011, « L'étude des médias est-elle soluble dans l'informatique et la physique ? À propos du recours aux *digital methods* dans l'analyse de l'information en ligne », *Questions de Communication*, n° 20, pp. 353-376.
- RIEDER B., 2007, « Étudier les réseaux comme phénomènes hétérogènes : quelle place pour la 'nouvelle science des réseaux' en sciences humaines et sociales ? », *ARchiveSIC*, pp. 1-17, <[http://archivesic.ccsd.cnrs.fr/sic\\_00379526/en/](http://archivesic.ccsd.cnrs.fr/sic_00379526/en/)>, dernière consultation le 20 mars 2014.
- RIEDER B., 2008, « Entre marché et communauté : une discussion de la culture participative à l'exemple de Google Maps », *Actes de la conférence Ludovia 2008*, pp. 1-11, <[http://archivesic.ccsd.cnrs.fr/sic\\_00329899](http://archivesic.ccsd.cnrs.fr/sic_00329899)>, dernière consultation le 20 mars 2014.
- RIEDER B. et T. RHÖLE, 2012, « Digital Methods: Five challenges », dans D. M. BERRY (dir.), *Understanding Digital Humanities*, Basingstoke, Palgrave Macmillan, pp. 67-84.
- ROGERS R. et N. MARRES, 2000, « Landscaping climate change: a mapping technique for understanding science and technology debates on the World Wide Web », *Public Understanding of Science*, vol. 9, n° 2, pp.141-163.
- ROGERS R., (dir.), 2000, *Preferred Placement: Knowledge Politics on the Web*, Amsterdam, Jan Van Eyck Editions.

Ouvrir la boîte à outils de la recherche numérique  
Trois cas de redistribution de méthodes

- ROGERS R., 2004, *Information Politics on the Web*, Cambridge, MIT Press.
- ROGERS R., 2009, « The End of the Virtual: Digital Methods », Amsterdam, Amsterdam University Press.
- ROGERS R. et A. BEN-DAVID, 2008, « The Palestinian—Israeli peace process and transnational issue networks: the complicated place of the Israeli NGO », *New Media & Society*, vol. 10, n° 3, pp. 497-528.
- STANFORD UNIVERSITY, 2010, *Voltaire and the Enlightenment*, <<http://republicofletters.stanford.edu/casestudies/voltaire.html>>, dernière consultation le 20 mars 2014.
- STAR S. L., 1999, « The Ethnography of Infrastructure », *American Behavioral Scientist*, n° 43, pp. 377-391.
- STAR S. L. et K. RUHLER, 1996, « Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces », *Information Systems Research*, vol. 7, pp. 111-134.