



Revue d'anthropologie des connaissances

10-4 | 2016

Ce que les data font faire aux SHS (et vice-versa)

Eric T. Meyer & Ralph Schroeder, *Knowledge Machines: Digital Transformations of the Sciences and Humanities*

Dominique Vinck



Édition électronique

URL : <http://journals.openedition.org/rac/2367>

ISSN : 1760-5393

Éditeur

Société d'Anthropologie des Connaissances

Référence électronique

Dominique Vinck, « Eric T. Meyer & Ralph Schroeder, *Knowledge Machines: Digital Transformations of the Sciences and Humanities* », *Revue d'anthropologie des connaissances* [En ligne], 10-4 | 2016, mis en ligne le 01 décembre 2016, consulté le 31 janvier 2020. URL : <http://journals.openedition.org/rac/2367>

Ce document a été généré automatiquement le 31 janvier 2020.



Les contenus de la *Revue d'anthropologie des connaissances* sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

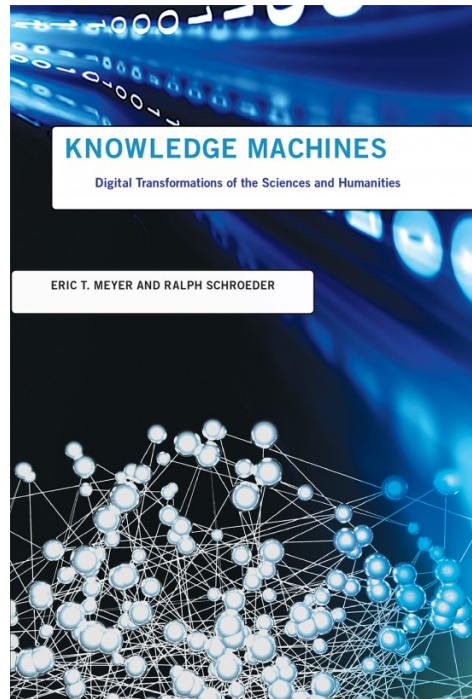
Eric T. Meyer & Ralph Schroeder,
*Knowledge Machines: Digital
Transformations of the Sciences and
Humanities*

Dominique Vinck

RÉFÉRENCE

Eric T. Meyer & Ralph Schroeder, *Knowledge Machines: Digital Transformations of the Sciences and Humanities*, Cambridge, Mass.: The MIT Press, 2015, 271 p.

- 1 L'ouvrage *Knowledge machines* est particulièrement intéressant pour l'analyse des transformations en cours dans les sciences du fait du recours aux technologies de l'information et de la communication. Il se penche tout particulièrement sur une série d'usages collaboratifs des outils numériques et couvre une ample variété de disciplines, de la physique des hautes énergies, de l'astronomie et des sciences de vie aux sciences sociales et aux humanités numériques. Il s'interroge sur les transformations en cours dans l'organisation et dans les pratiques de recherche mais aussi dans les contenus de connaissances : données, questionnements et approches. Il s'arrête notamment sur la production plus ou moins massive de données et l'accès à l'information ainsi que sur leur traitement faisant appel tantôt à d'importants réseaux d'ordinateurs, tantôt à d'innombrables collaborateurs, parfois volontaires et amateurs. S'intéressant à la façon dont les outils numériques reconfigurent la production des connaissances, les auteurs, Eric Meyer et Ralph Schroeder de l'*Oxford Internet Institute* se demandent en particulier si l'utilisation distribuée et collaborative d'outils et de données numériques ouvre de nouveaux horizons de recherche.
- 2 Après une substantielle introduction, l'ouvrage comprend neuf chapitres : un cadrage conceptuel, une qualification globale de l'émergence du numérique dans la recherche (financement, publication, visibilité), puis quatre chapitres offrant un très bon compte rendu d'études de cas réunis par thématiques (agrégation des personnes et d'ordinateurs, possibilités et limites du partage des données de recherche, différences entre disciplines, notamment les sciences humaines et sociales, et circulation des outils entre disciplines) et se termine par trois chapitres de lecture transversale portant sur l'ouverture (*open science*), les limites de la recherche collaborative numérique et les transformations et conséquences des *knowledge machines*.



Diversité des usages collaboratifs des outils numériques

- 3 Après avoir dressé une typologie des recherches portant sur l'Internet (phénomène social, source d'information, usages en tant qu'outil de recherche, amélioration des fonctionnalités), les auteurs posent Internet en tant que nouvelle infrastructure de recherche supportant l'*e-research* objet de l'ouvrage. Internet est ici défini comme un assemblage sociotechnique composé de dispositifs technologiques interconnectés, de protocoles, d'applications, de données et de bases de données mais aussi d'acteurs et d'usages. La question est d'identifier et de qualifier les collaborations de recherche qu'il

a rendues possibles et leurs conséquences pour les dynamiques scientifiques, qu'elles concernent les chercheurs promoteurs de nouvelles pratiques scientifiques ou les chercheurs accidentels ayant recours aux outils numériques parce qu'ils sont, de fait, répandus et utilisés par leurs pairs.

- 4 Bien que les transformations à l'œuvre avec les outils numériques fassent l'objet de nombreuses recherches, les auteurs les qualifient de fragmentées (*science and technology studies*, *media studies*, *infrastructure studies*, politiques de l'innovation, *web science*, droit des technologies numériques, sciences et technologies de l'information) mais surtout peu focalisées sur la transformation des pratiques scientifiques. Or, grâce aux impulsions institutionnelles en faveur des cyber-infrastructures pour la recherche mais aussi aux initiatives de certaines communautés scientifiques, nombreux sont les développements qui ont vu le jour. Les auteurs se penchent sur une série de cas, très bien décrits et documentés dans l'ouvrage, qui nourrissent la réflexion : la mise en réseau de nombreux ordinateurs dans le monde et l'organisation d'une bureaucratie associée afin de disposer d'une capacité de calcul suffisante pour traiter les masses de données issues des expériences en physique des hautes énergies réalisées au CERN à Genève (*Grid Particle Physics*) ; la mobilisation et la mise en réseau d'innombrables amateurs (*crowdsourcing*) d'astrophysique pour caractériser plus d'un million de photographies de galaxies (*Galaxy Zoo*) permettant aussi d'entraîner les algorithmes (*machine learning*) qui feront automatiquement ce même travail sur les milliards de photographies de l'astrophysique à venir ; la collecte, la mise en commun et le traitement des données expérimentales produites par de nombreux groupes de recherche en médecine et en génétique (*GAIN*) ; la mobilisation d'un vaste réseau d'ordinateurs personnels et de l'industrie pharmaceutique afin de disposer d'une capacité de calcul suffisante que pour analyser les masses de données en protéomique et pour réaliser un screening à grande échelle permettant de trouver des remèdes contre la dengue (*SwissBioGrid*) ; la collecte, de longue date, de données de grande qualité sur l'ensemble de la population suédoise et leurs mises en relation pour les besoins de la recherche, notamment en sciences sociales, supposant des régulations sociétales telles que la population accorde sa confiance à l'État et aux chercheurs ; la collecte et la mise en commun de photographies numériques d'ailerons de baleines, leur identification et leur association à des données environnementales, permettant de suivre et d'étudier leur population à l'échelle mondiale et à très long terme (*SPLASH*) ; la mobilisation de multiples amateurs pour interpréter, annoter et discuter tous les passages de l'œuvre d'un écrivain (*Pynchon Wiki*) ; la réorganisation des communautés de recherche pour rendre possibles de telles collaborations à grande échelle.
- 5 Les différentes études de cas mobilisées dans l'ouvrage permettent d'aborder une série de questions cruciales comme : les conditions du partage et de la mise en circulation des données, devant parfois être rendues anonymes dans le cas des sciences sociales ; la confiance dans la recherche et dans les institutions publiques ; la constitution d'un langage commun et de référentiels partagés ainsi que la mise en compatibilité des méthodes et des instruments, leur actualisation et leur maintien ; les créations d'un partenariat équitable privé-public ; le financement durable des infrastructures de recherche¹ ; le maintien dans la (très longue) durée de ces collaborations et des outils associés ; l'incitation à partager ses données et les mécanismes de reconnaissance des mérites des différents collaborateurs ; l'ouverture de l'accès des données de recherche

aux publics et la transformation des relations entre chercheurs professionnels et autres acteurs dans la société.

Styles de sciences et modélisation des réseaux de l'e-recherche

- 6 L'approche d'Eric Meyer et Ralph Schroeder est de s'appuyer sur les travaux de Ian Hacking pour identifier des « styles de science » (raisonnement mathématique, exploration expérimentale et mesure de relations, modélisation, construction de taxonomies et classification, analyse statistique et calcul des probabilités, explication de la genèse) ainsi que sur la théorie de l'acteur-réseau pour caractériser les réseaux sociotechniques correspondant à chacun des projets examinés. Dans le dernier chapitre, une modélisation de chaque réseau permet leur comparaison, prenant en compte les personnes et groupes impliqués, les types de machines utilisées, les matériaux, données et bases de données, le traitement des données et les styles de recherche à l'œuvre.
- 7 Ce cadrage conceptuel, ainsi que des références à Randall Collins, Richard Whitley, Terry Shinn et Bernward Joerges, permet aux auteurs de qualifier le rôle que jouent les outils et ressources numériques dans chacun des cas. Ils se demandent notamment s'ils sont le véhicule d'une scientification et d'une mathématisation croissante des sciences, s'ils favorisent l'émergence de consensus entre chercheurs en donnant plus d'objectivité aux phénomènes étudiés et s'ils conduisent à des transferts d'outils (méta-artefacts, instruments génériques), de méthodes (méta-méthodologies) et de formes d'organisation du travail d'une discipline à l'autre (par exemple, le recours au *crowdsourcing*, la mise en réseau d'ordinateurs, la mise en circulation de données). En s'appuyant sur la caractérisation organisationnelle, par Richard Whitley, des disciplines scientifiques en fonction du degré d'interdépendance stratégique et fonctionnelle de leurs membres, ils se demandent si les outils numériques s'implantent plus facilement dans certaines configurations que dans d'autres et s'ils donnent des impulsions à ces disciplines favorisant leur scientification. Puis, à la recherche d'un cadrage conceptuel qui aide à construire des analyses plus globales du phénomène, ils mobilisent les travaux de Rob Kling sur les interactions sociales, le développement des systèmes d'information et la configuration de leurs composants, portant attention aux usages et aux routines, ainsi que la notion de mouvement d'informatisation (prenant en compte les discours publics sur les technologies) et la notion de Scott Frickel et Neil Gross de mouvement intellectuel et scientifique (approche dominante d'un problème). Les auteurs se demandent si la e-recherche correspond plutôt à un mouvement d'informatisation (démonstration de l'intérêt des outils numériques) ou à un mouvement intellectuel et scientifique (réflexion sur l'amélioration des approches et pratiques scientifiques).
- 8 Ce faisant, Eric Meyer et Ralph Schroeder prétendent dépasser les *internet studies* et les *science and technology studies*, en expliquant ce qui change, y compris au niveau des connaissances. Les études sur Internet n'y auraient pas contribué parce qu'elles ne s'intéressent pas à ce qui se passe dans les sciences. Les études sur les sciences et les techniques ne sont limitées à des études de cas. Le présent ouvrage a au moins ce mérite d'avoir réuni une belle série d'études de cas, variées dans leur contenu et les disciplines auxquelles elles se réfèrent, de les avoir bien documentées et de les avoir

modélisées au sein d'un cadre conceptuel homogène rendant possible leur comparaison.

Pourquoi le soufflé retombe-t-il ?

- 9 La lecture de *Knowledge machines* est particulièrement intéressante. La problématique de départ est pertinente et les études de cas sont du plus grand intérêt. Or, des derniers chapitres, prétendant offrir des lectures transversales et plus globales, ressort l'impression d'un soufflé qui retombe.
- 10 S'agissant de l'ouverture des données de la recherche, le texte s'écarte des études de cas, pour mobiliser l'éthos des sciences qualifié par Robert Merton, les pressions en faveur et en défaveur de l'accès libre et ouvert *versus* d'un nouveau mouvement d'enclosure et de verrouillage de la propriété intellectuelle, l'impact finalement limité de l'ouverture à cause de l'économie de l'attention inévitablement elle-même limitée. Le chapitre n'est pas inintéressant mais il ne se nourrit pas des études de cas et n'y revient pas pour les éclairer sous un nouveau jour.
- 11 S'agissant de l'avant-dernier chapitre portant sur les limites de la recherche collaborative numérique, il mobilise de nouveaux cas, notamment celui de l'échec d'un outil de cartographie. Dans ce cas, tout semble bien parti pour constituer une *success story*; la technique est au point et les intermédiaires et utilisateurs finaux manifestement intéressés. Il n'y aurait même aucun problème financier. Le problème vient des barrières institutionnelles et légales et, surtout, de l'« indifférence institutionnelle » écrivent les auteurs, mobilisant pour le coup un registre explicatif ne faisant pas partie du cadre conceptuel posé en début d'ouvrage. Ils se penchent ensuite sur les archives du web pour questionner leur devenir et, notamment, le peu d'intérêt qu'elles suscitent chez les chercheurs. Ils se demandent si les archives du web ne deviendront pas poussiéreuses comme toute autre archive. Ils abordent ensuite une série de réflexions générales, pertinentes mais déconnectées des études de cas, sur les limites du partage des données, les enjeux éthiques, la fragilité de la confiance, les contraintes économiques.
- 12 Enfin, revenant sur la problématique des styles de science, Eric Meyer et Ralph Schroeder exposent leur modélisation des cas et les comparent avant d'en dégager des conclusions ou des questionnements généraux sur : le caractère plus ou moins cumulatif des disciplines ; la scientification de sous-disciplines qui adoptent les outils numériques et qui deviennent plus empiriques, positivistes, mathématiques ; le rôle des technologies qui, comme des passeports, facilitent le passage de méthodes d'une discipline à l'autre ; et la perpétuation de différences entre disciplines. Ils concluent en réaffirmant leur thèse (assez évidente mais pas vraiment démontrée), à savoir : les outils numériques ont transformé les sciences, y compris les sciences humaines et sociales. Ils indiquent toutefois que l'issue de ce mouvement est encore incertaine et réitèrent une série de questions, notamment sur le rôle relatif du mouvement d'informatisation *versus* mouvement intellectuel et scientifique, les acteurs inclus et exclus et la transformation des frontières entre sciences et publics, l'apprentissage, l'architecture plus ou moins solide selon les disciplines, le caractère non révolutionnaire des *knowledge machines* (peu de déplacement paradigmatiques) mais une machinisation croissante de la production des connaissances. Ils indiquent que toute conclusion générale a des limites et, de façon très relativiste, que le changement

dépend des disciplines et du niveau d'observation. Il serait également trop tôt pour proposer une vision complète des mécanismes à l'œuvre.

- 13 Fait étonnant, les auteurs reviennent peu sur les études de cas pour construire des bouts de connaissance ou pour traiter des grandes questions qui sont abordées. Étonnant aussi, ils ne mobilisent pas les travaux des chercheurs (Geoff Bowker, Susan L. Star notamment) qui ont finement analysé les pratiques de constitution de (cyber-)infrastructures de la recherche. Ils ne se réfèrent que marginalement à la théorie de l'acteur-réseau qu'ils réduisent à une approche constructiviste et qu'ils utilisent erronément pour soutenir l'idée d'une articulation entre la technique (réseau d'ordinateurs) et le social (communautés scientifiques, publics). Ils reprochent notamment à ces travaux de se limiter à des études locales ou de cas singuliers, de ne pas construire de vision d'ensemble du phénomène et de ne rien dire des effets de connaissance. Cette méconnaissance ou mécompréhension les conduit, dans les derniers chapitres, à monter en généralité sur les problématiques abordées dans l'ouvrage, de façon peu convaincante, alors que les études de cas qui constituent le cœur de l'ouvrage sont riches d'enseignement et méritent, à elles seules, la lecture de cet ouvrage qui devrait marquer la réflexion sur les transformations en cours dans les sciences.
-

NOTES

1. Cf. l'ouvrage de Brian Kleiner, Isabelle Renschler, Boris Wernli, Peter Farago et Dominique Joye, *Understanding Research Infrastructures in the Social Sciences*, dont un compte rendu a été publié dans la *Revue d'anthropologie des connaissances* en 2013, vol. 7, n° 4.
-

AUTEURS

DOMINIQUE VINCK

Adresse : STS lab, Institut des Sciences Sociales, Université de Lausanne.

<https://orcid.org/0000-0001-7835-7008>

Courriel : dominique.vinck[at]unil.ch