



**HAL**  
open science

# L'articulation difficile de deux démarches scientifiques pour étudier le web et les réseaux

Roger Bautier

► **To cite this version:**

Roger Bautier. L'articulation difficile de deux démarches scientifiques pour étudier le web et les réseaux. *Alliage: Culture - Science - Technique*, 2024, 84. hal-04754132

**HAL Id: hal-04754132**

**<https://hal.science/hal-04754132v1>**

Submitted on 25 Oct 2024

**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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## **L'articulation difficile de deux démarches scientifiques pour étudier le web et les réseaux**

Roger Bautier  
UR 1803 – Laboratoire des Sciences de  
l'Information et de la Communication,  
Université Sorbonne Paris Nord

Le principal responsable de la création du web en 1989, le physicien Tim Berners-Lee, a toujours prôné, avec son équipe, une analyse attentive du développement de ce gigantesque réseau, analyse qui, à partir de la physique, deviendrait pluridisciplinaire en intégrant les apports de différentes sciences sociales<sup>1</sup>. Cependant, si la collaboration qui s'est instituée, au cours des dernières décennies, entre les sciences de la communication et la sociologie, la science politique ou l'économie a permis de mieux connaître le réseau du web, il n'empêche que cette collaboration n'a souvent mis en relation que des disciplines relevant des sciences sociales. En même temps, les apports de celles qui n'en relèvent pas, notamment les apports de la physique, ont été tout à fait considérables. D'une part, les travaux effectués par les physiciens ne se sont pas cantonnés dans le repérage des contraintes physiques s'exerçant sur les réseaux techniques. Bien au contraire, ils se sont intéressés à tous les réseaux, quels qu'ils soient. Le développement des réseaux gigantesques comme celui du web ayant relancé l'intérêt pour les applications de la théorie des graphes, ils ont fourni d'abord, en recourant à celle-ci, des éléments précieux pour une meilleure connaissance de la structure des réseaux, soit au niveau macroscopique (celui de leur configuration générale et de leur évolution), soit au niveau microscopique (celui, par exemple, de sites web constituant des communautés, elles-mêmes liées à des communautés humaines). D'autre part, ces travaux n'ont pas traité que de l'aspect topologique : débordant vers la biologie, ils ont avancé fréquemment des analyses conceptuelles et empiriques qui imposent une analyse des pratiques de communication en termes de systèmes complexes, comme ceux que constituent les systèmes vivants. Ils ont insisté alors sur les caractéristiques auto-organisationnelles des réseaux étudiés, en les considérant seulement comme soumis à des règles de développement qu'il s'agit de découvrir, sans presque jamais considérer ces réseaux sous un angle proprement socio-politique. On comprend, dès lors, que ces deux démarches scientifiques, l'une ancrée dans les sciences sociales, l'autre dans les sciences de la nature, soient souvent restés indépendantes l'une de l'autre, d'autant plus qu'une certaine tension a pu apparaître entre les deux, à partir du moment où les confrontations, quand elles ont existé, ont posé la question d'une absorption éventuelle des sciences sociales par les sciences de la nature et celle des effets de l'informatisation sur les recherches.

### **Comment décrire le web ?**

Les premières recherches s'inscrivent dans un cadre de réflexion que l'on peut préciser en faisant référence aux différents aspects à distinguer lorsque l'on parle de réseaux<sup>2</sup>. Premier aspect : l'aspect topologique. On identifie alors le concept de réseau à celui de graphe au sens mathématique : un graphe est un couple de deux ensembles d'éléments, l'un étant formé de paires d'éléments (des arêtes si les couples ne sont pas ordonnés, des arcs s'ils le sont) qui

---

<sup>1</sup> Tim Berners-Lee, Wendy Hall, James Hendler, Kieron O'Hara, Nigel Shadbolt, Daniel Weitzner, "A framework for Web Science", *Foundations and Trends in Web Science*, vol. 1, n° 1, 2006, p. 1-130.

<sup>2</sup> Daniel Parrochia, « Les réseaux : de la mathématique à la métaphysique », *Art Press*, hors-série n° 2, « Internet all over », 1999.

appartiennent à l'autre (les sommets, qu'on appellera plutôt des nœuds lorsqu'on observera des réseaux réels). On définit le degré d'un sommet comme le nombre des arêtes qui lui sont incidentes et on parle de degré entrant ou de degré sortant quand le graphe est orienté. Deuxième aspect : on envisage des structures dans lesquelles circulent des flux ou des flots (le point de vue adopté est celui de la théorie des réseaux de transport). Troisième aspect : on s'intéresse à des réseaux complexes, difficiles à modéliser, notamment ceux qui sont associés aux systèmes vivants, et on parle de réseaux autonomes, de réseaux qui s'autorégulent.

Dès la fin des années 90, le premier aspect va donner lieu à des travaux décisifs (qui intégreront ensuite les autres aspects), effectués généralement par des chercheurs en physique qui s'intéressent au web, désormais le plus grand graphe jamais considéré. Ceux-ci montrent que les réseaux réguliers sont caractérisés par une forte densité de connexions entre des nœuds proches et par des longueurs de chemins considérables (pour aller d'un nœud à un autre éloigné, il faut passer par de nombreux nœuds intermédiaires) et que les réseaux aléatoires ont, au contraire, une faible densité de connexions entre des nœuds proches mais ont de faibles longueurs de chemins (le caractère aléatoire diminue la probabilité que des nœuds proches aient beaucoup de connexions, mais introduit plus de liens qui relient une partie du réseau à une autre partie). En comparaison, les réseaux réels que l'on peut étudier apparaissent comme ni totalement réguliers ni totalement aléatoires. D'où la proposition, pour en rendre compte, d'élaborer un modèle qui part d'une structure de treillis (graphe très simple et très régulier) et qui introduit de l'aléatoire dans la structure. Cependant, les modèles de réseaux habituels<sup>3</sup> ont deux défauts : premièrement, ils « démarrent » avec un nombre fixe de nœuds, qui sont ensuite reliés aléatoirement, sans qu'il y ait modification de leur nombre initial ; deuxièmement, ils supposent que la probabilité que deux nœuds soient reliés est à la fois aléatoire et uniforme. Il est donc nécessaire, d'une part, de tenir compte du fait que la plupart des réseaux réels sont ouverts, c'est-à-dire qu'ils se forment par l'addition continue de nouveaux nœuds, comme le montre, en particulier, l'expansion rapide et ininterrompue du web. Il faut, d'autre part, tenir compte du fait qu'il existe, dans les réseaux réels, une connexité préférentielle : dans le cas du web, il est clair qu'une nouvelle page inclura plus vraisemblablement des liens vers des documents populaires, qui ont déjà une forte connexité.

Il a ainsi pu être montré que les réseaux observables dans de nombreux systèmes, biologiques ou fabriqués par les humains, sont des réseaux de type « petit monde » (*small-world network*). Le web, en particulier, apparaît comme ayant une topologie de type<sup>4</sup>. Un réseau est dit « petit monde » lorsque la distance topologique moyenne, c'est-à-dire le nombre moyen de sauts pour aller d'un nœud à un autre, varie très lentement avec le nombre total de nœuds (de manière logarithmique), même si chaque nœud est relié à un faible nombre d'autres nœuds et même si la cohésion locale reste forte. Cette topologie a des implications concernant la manière dont les utilisateurs se servent du web et la facilité avec laquelle ils rassemblent l'information. De plus, la structure des liens fournit des indications sur les relations sous-jacentes entre les gens, leurs intérêts et les communautés existantes ou émergentes. Elle fournit notamment la possibilité d'envisager des applications au marketing commercial ou politique<sup>5</sup>. Par ailleurs, plusieurs études ont pu mettre en lumière que le réseau du web était marqué par des distributions « larges », très éloignées des caractéristiques des réseaux aléatoires<sup>6</sup>. Elles ont ainsi souligné

---

<sup>3</sup> Duncan Watts, Steven Strogatz, "Collective dynamics of 'small-world' networks", *Nature*, n° 393, 1998, p. 440-442.

<sup>4</sup> Lada Adamic, "The small world web", *ECDL '99: Research and Advanced Technology for Digital Libraries*, 1999, p. 443-452.

<sup>5</sup> Lada Adamic, Bernardo Huberman, "The nature of markets in the world wide web", *Quarterly Journal of Electronic Commerce*, vol. 1, n° 1, 2000, p. 1-10.

<sup>6</sup> Albert-László Barabási, Reka Albert, Hawoong Jeong, "Scale-free characteristics of random networks: the topology of the world wide web", *Physica A*, vol. 281, 2000, p. 69-77 ; Andrei Broder, Ravi Kumar, Farzin

(ce qui a été souvent confirmé ensuite) que, loin d'être constitué de nœuds dont la plupart auraient un nombre de liens proche d'une certaine moyenne (on aurait alors une représentation de la distribution en forme de cloche), le réseau du web est, au contraire, formé de nœuds dont certains, très nombreux, n'ont que très peu de liens, voire aucun, alors que d'autres (des super-nœuds), en petit nombre, en ont beaucoup ou même énormément. Plus précisément, il est apparu qu'un tel réseau est connecté suivant une loi de puissance : on constate que la probabilité qu'un nœud ait  $k$  liens est inversement proportionnelle à ce nombre  $k$  élevé à la puissance  $n$ , la valeur de cet exposant  $n$  variant entre 2 et 3. La courbe correspondante est souvent tellement écrasée que le repère linéaire est utilement remplacé par un repère logarithmique, faisant apparaître une droite emblématique de la loi de puissance.

On notera que le champ d'études qui en est issu est d'une fécondité telle que celle-ci a sans doute quelquefois obscurci des distinctions qui doivent pourtant être maintenues entre : ce qui relève, d'abord, de la démonstration mathématique (par exemple : le fait que, dans un graphe, le nombre de sommets de degré impair est pair) ; ce qui relève, ensuite, de l'observation empirique, nécessairement partielle, et qui va être, éventuellement, exprimé sous forme mathématique (par exemple : les distributions statistiques dont on peut donner une approximation par le recours à une loi de puissance) ; ce qui relève, enfin, de la simulation (par exemple : l'engendrement de réseaux par tel ou tel programme informatique, dont la fonction peut être de servir à prévoir une étape prochaine dans le développement d'un réseau existant). Il n'en reste pas moins que ce champ d'études a renouvelé profondément les perspectives de l'analyse des réseaux, aussi bien au sein des sciences sociales qu'au sein des sciences dites exactes.

### **La nouvelle science des réseaux et ses applications**

À partir de ces premières recherches, les travaux se sont donc multipliés, portant sur le web et sur d'autres réseaux, en visant à fournir des modélisations de leur développement. Ils ont pu mettre en évidence, d'une part, que les réseaux de ce genre existent en grand nombre et, d'autre part, qu'ils se différencient nettement des réseaux aléatoires : il s'agit de réseaux invariants d'échelle (*scale-free*), cette propriété étant liée à la présence de distributions statistiques dont on peut rendre compte par une loi de puissance. La modélisation de base<sup>7</sup> revient à considérer que le fonctionnement de ces réseaux suppose l'existence de deux caractéristiques qui, lorsqu'elles sont appliquées au web, se traduisent par deux hypothèses. La première est que le réseau de pages est en expansion continue ; la seconde est qu'une nouvelle page inclura des liens vers des pages déjà existantes avec une probabilité proportionnelle au degré entrant de ces pages, autrement dit à leur popularité. Il s'agit là de ce que l'on appelle un attachement préférentiel, qui est la traduction d'une tendance à ce que « les riches deviennent plus riches ». Une idée fondamentale apparaît alors, qui a été soulignée par Albert-László Barabási<sup>8</sup> : c'est l'idée que les pages web auxquelles on préfère se connecter ne sont pas des nœuds ordinaires (ce sont des plaques tournantes, qui ont d'autant plus de liens qu'elles sont mieux connues). Ce qui veut dire que, alors même que les comportements individuels sont largement imprévisibles, ceux des groupes observent des règles strictes. Une autre idée fondamentale s'y ajoute, qui consiste à estimer que ce réseau, parce qu'il établit des différences entre les nœuds, ne peut effectivement pas être qualifié de démocratique : la loi de l'attachement préférentiel paraît implacable.

---

Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, Janet Wiener, "Graph structure in the web", *Computer Networks*, vol. 33, 2000, p. 309-320.

<sup>7</sup> Albert-László Barabási, Reka Albert, "Emergence of scaling in random networks", *Science*, vol. 286, 1999, p. 509-512.

<sup>8</sup> Albert-László Barabási, *Linked: The New Science of Networks*, Cambridge, Perseus, 2002.

La référence à cette modélisation, ainsi qu'à ses variantes et à ses prolongements, implique l'acceptation de la popularité comme facteur essentiel de l'attractivité. Plus récemment, il a été mis l'accent sur le fait que l'attractivité comportait, en fait, deux dimensions : la popularité, certes, mais également la similarité<sup>9</sup>. Il s'agit alors de considérer que des nœuds marqués par une plus grande similarité ont plus de chances de se connecter, même en cas de popularité faible : quelqu'un qui crée une page personnelle nouvelle la reliera à des sites populaires, mais aussi à des sites relativement peu populaires qui sont proches de ses intérêts personnels. D'où une modélisation qui intègre ces deux dimensions et fait l'hypothèse d'un certain équilibre entre les deux. Premièrement, on admet que, toutes choses étant égales par ailleurs, les nœuds les plus anciens ont plus de chances de devenir populaires et que, s'ils sont créés un par un, ils vont être caractérisés par leur numéro d'ordre de naissance. Deuxièmement, afin de rendre compte de la similarité, on dispose les nœuds de manière aléatoire sur un cercle et l'on assimile les distances angulaires entre eux à une traduction de leur plus ou moins grande similarité. Dès lors, une modélisation simple de l'équilibre entre popularité et similarité peut être de supposer que chaque connexion nouvelle correspond à l'optimisation du produit de ces deux grandeurs, autrement dit au choix du produit minimal. Le nœud vers lequel sera établi un lien devra, d'une part, avoir le numéro d'ordre de naissance le plus petit et donc être le plus connecté, d'autre part, avoir la distance angulaire la plus petite et donc être le plus similaire.

Alors que le modèle classique de l'attachement préférentiel repose sur le hasard présidant à l'attribution initiale des liens, attribution qui détermine ensuite les nouvelles connexions correspondant au développement du réseau, le modèle qui introduit la prise en compte de la similarité fait appel, quant à lui, à une intervention humaine rationnelle. La comparaison de ces deux modèles<sup>10</sup> peut amener à la conclusion que l'existence du phénomène de l'attachement préférentiel dans de très nombreux systèmes profondément différents signifie qu'il dépend, en même temps, du hasard et de la raison. Mais n'est-ce pas, par là même, une invitation à se pencher plus précisément sur la question de la phase d'ascension vers la popularité ? Si l'on comprend bien que la popularité acquise à un certain moment donne la possibilité d'obtenir une popularité encore plus grande ensuite, il demeure que l'acquisition initiale d'une position favorable est marquée par la précarité et susceptible de rencontrer de multiples facteurs de risque. Une façon de le souligner est d'imaginer que l'histoire puisse être « rejouée<sup>11</sup> ». Si elle l'était plusieurs fois, il est probable que, à chaque fois, on rencontrerait une distribution de la popularité obéissant à une loi de puissance, mais il n'est pas du tout évident que cette popularité s'attacherait aux mêmes objets.

On constate que ces dernières décennies ont été marquées par l'intérêt qui s'est manifesté, chez de nombreux physiciens, à l'égard de l'analyse des réseaux sociaux en contexte numérique. Certains de leurs travaux ont suscité des réactions souvent favorables dans la presse et sur le web. En revanche, les sociologues ont été moins accueillants, notamment parce qu'il leur est apparu que les physiciens apportaient des observations intéressantes ainsi que des nouveautés théoriques, mais qu'ils ignoraient presque totalement les recherches plus anciennes sur les réseaux sociaux. Un spécialiste des réseaux sociaux comme John Scott<sup>12</sup>, par exemple, ne peut que s'étonner que Albert-László Barabási ait pu affirmer<sup>13</sup> qu'il avait révélé un ordre insoupçonné au sein des réseaux en montrant que les réseaux sociaux n'étaient pas le fruit de connexions aléatoires. L'ignorance des travaux sociologiques est difficile à interpréter :

---

<sup>9</sup> Fragkiskos Papadopoulos, Maksim Kitsak, M. Angeles Serrano, Marian Boguña, Dmitri Krioukov, "Popularity versus similarity in growing networks", *Nature*, vol. 489, n° 7417, 2012.

<sup>10</sup> Albert-László Barabási, « Network Science: Luck or reason », *Nature*, vol. 489, n° 7417, 2012.

<sup>11</sup> David Easley, Jon Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, Cambridge, Cambridge University Press, 2010.

<sup>12</sup> John Scott, « Social physics and social networks », in John Scott, Peter Carrington (eds.), *The Sage Handbook of Social Network Analysis*, Londres, Sage, 2011, p. 55-66.

<sup>13</sup> Albert-László Barabási, *Linked: The New Science of Networks*, op. cit.

signifierait-elle que les sociologues ne peuvent pas être de véritables scientifiques ? Pour les chercheurs qui appartiennent au courant de la science des réseaux, la différence entre l'ancienne science des réseaux et la nouvelle est que l'ancienne n'est pas allée très loin dans l'étude de la structure des réseaux. Il est vrai que la théorie des graphes, notamment, ne se rattache qu'aux mathématiques et que ses résultats, pourtant importants et élégants, soit n'ont pas apporté une vision particulièrement pertinente des réseaux existant dans le monde réel, soit se sont concentrés sur la conception et l'ingénierie des réseaux. Une des caractéristiques de la nouvelle science des réseaux, à cet égard, est d'avoir traité de ceux qui se comportent de manière non planifiée ou bien sont soumis aux actions d'agents très nombreux, sur de longues périodes de temps. L'approche sociologique classique, quant à elle, leur apparaît comme fortement empirique et évitant, sauf exception, la modélisation, pour simplement décrire les propriétés des réseaux observés au travers des données recueillies.

### **Comprendre le fonctionnement des réseaux complexes**

Ces travaux issus de la physique, plus particulièrement de la physique statistique, sont pourvoyeurs d'acquis extrêmement riches constituant une base à partir de laquelle il est possible d'étudier les réseaux complexes constitués par de nombreux systèmes artificiels ou naturels. Les réseaux en question sont complexes, en ce sens qu'ils sont en général le résultat d'une évolution décentralisée et non planifiée. D'où l'intérêt d'une étude de l'émergence de leurs propriétés statistiques macroscopiques, étude qui sera entreprise à partir d'une discipline habituée à traiter des systèmes physiques comportant un très grand nombre d'éléments en interaction<sup>14</sup>. Des présentations générales des travaux sur les réseaux complexes, toujours plus nombreux, ont été publiées plusieurs fois, soit sous une forme technique<sup>15</sup>, soit sous une forme vulgarisée, dans des ouvrages ou des magazines et même dans la presse quotidienne. Ces présentations soulignent l'importance d'une approche modélisatrice et transdisciplinaire de la complexité, qui permettrait de surmonter, à la fois, les limites des études seulement théoriques issues des mathématiques et celles des études issues des sciences humaines et sociales, purement descriptives et portant sur des réseaux relativement restreints. Quelle est la pertinence des représentations induites par ce type d'approche, qui tend à mettre en lumière une grande similarité de comportement des réseaux complexes ? Si l'on peut constater une grande similarité de comportement des réseaux complexes, il existe aussi des variations ; de plus, même s'il est tentant de parler d'universalité, il ne semble pas exister d'explication générale de la similarité.

Ce ne sont donc plus seulement l'internet et le web qui servent de terrain d'étude aux physiciens ou à ceux qui ont exploité leurs travaux, mais également toutes sortes de réseaux. Sont examinés ou réexaminés les réseaux sociaux tels qu'ils avaient déjà été étudiés généralement dans le cadre de la sociologie ou dans celui de la psychologie sociale (comme les réseaux d'individus travaillant ensemble), les réseaux techniques comme ceux de la distribution de l'énergie électrique et des lignes de transport aérien, les réseaux biologiques comme ceux des molécules interagissant à l'intérieur des cellules, les réseaux économiques comme ceux des entreprises, etc. C'est pourquoi les recherches sur les réseaux peuvent aboutir à la célébration d'une science des réseaux capable de traiter de l'ensemble des réseaux existants. Après une période marquée par une attitude jugée réductionniste, on aurait enfin la possibilité de traiter convenablement de la complexité de manière transdisciplinaire, avec pour guide les résultats

---

<sup>14</sup> Fabrizio Li Vigni, « Le projet épistémique des sciences de la complexité », *Philosophia Scientiae*, vol. 24, n° 1, 2020, p. 181-200.

<sup>15</sup> Albert-László Barabási, Reka Albert, « Statistical mechanics of complex networks », *Review of Modern Physics*, vol. 74, n° 1, 2002, p. 47-97 ; Romualdo Pastor-Satorras, Alessandro Vespignani, *Evolution and structure of internet*, Cambridge, Cambridge University Press, 2004.

qui ont été fournis par la physique puis étendus à la biologie. Une telle revendication<sup>16</sup> se fonde sur le caractère universel reconnu aux régularités qui sont censées gouverner l'évolution de l'ensemble des réseaux complexes entourant ou constituant les humains.

Ce qui est ainsi apparu, c'est que les réseaux invariants d'échelle sont, à la fois, robustes et vulnérables. Robustes, d'un côté, parce qu'ils ne sont pas fortement perturbés par une détérioration aléatoire de liens (dans ce cas, les nœuds détruits ne sont souvent que des nœuds à faible degré). Et vulnérables, d'un autre côté, parce qu'ils sont extrêmement sensibles à une détérioration capable d'affecter spécialement les nœuds à fort degré (dans ce cas, en effet, un très grand nombre de liens sont détruits). La robustesse et la vulnérabilité d'un réseau peuvent, dès lors, se manifester aussi bien dans le fonctionnement du web que dans les pratiques sexuelles des individus, la propagation des maladies sexuellement transmissibles faisant ainsi écho à celle des virus informatiques. Il est ainsi compréhensible que les recherches sur les réseaux débouche sur une science se voulant transdisciplinaire, une science élaborée par des physiciens, étant donné que ceux-ci se considèrent capables de traiter d'autres disciplines que la physique<sup>17</sup>, à partir du moment où celle-ci est passée de l'étude de la matière à celle de toutes les formes d'organisation<sup>18</sup>.

Il n'y a absolument rien de métaphorique dans cette conception : c'est d'ailleurs ce qui fait, à la fois, son intérêt et son danger. Son intérêt réside dans le fait qu'elle apporte des connaissances précieuses sur le fonctionnement des réseaux, alors que son danger est lié à la dissolution des sciences sociales dans les sciences de la nature qu'elle impose. En effet, si, au lieu de relever classiquement des sciences sociales, les deux premiers réseaux sont jugés comme relevant bien, comme le troisième, des sciences de la nature, ce n'est pas parce qu'ils pourraient, d'une certaine manière, être comparés au troisième, mais parce qu'ils n'en sont aucunement différents. Fondamentalement, leur évolution est vue comme soumise aux effets radicaux de la loi de puissance qui se manifeste par une domination de super-nœuds engendrée par leur croissance auto-organisée. La référence à l'auto-organisation, fréquemment soulignée, montre que l'insistance sur l'universalité de la théorie des réseaux peut se prolonger en une analyse plus normative de leur évolution. Ainsi, lorsqu'on assimile des entreprises à des nœuds et des relations financières et économiques à des liens, il devient difficile non seulement d'envisager la remise en cause de la domination des entreprises en position de super-nœuds, mais encore d'accorder de la valeur à des approches économiques et politiques, puisque les phénomènes observés prennent le statut de conséquences du fonctionnement interne du réseau considéré. Il en résulte la tendance à naturaliser les rapports de pouvoir et à écarter la question de la régulation, en promouvant une attitude ambiguë à l'égard de l'action, dont on ne sait pas très bien si elle peut avoir un sens ou non, quand il s'agit de contester des situations existantes ou d'intervenir sur des situations prévisibles.

### **L'essor du traitement informatisé des données**

Les progrès dans la connaissance des réseaux et le recours aux spectaculaires possibilités offertes par l'informatique ont favorisé un traitement de plus en plus efficace des données présentes sur le web, ce dont témoigne l'essor de l'exploitation des *Big Data*. Qu'il s'agisse de prédire les caractéristiques des liens dans un réseau social numérique, de repérer les communautés qui s'y forment et y évoluent ou bien d'analyser la production et la diffusion, sur le web, des informations (vraies ou fausses) comme celles des opinions, on voit bien que la mise en œuvre des outils disponibles désormais ne fait que renforcer la prégnance d'une représentation du traitement des données qui peut inquiéter autant qu'enthousiasmer : il s'agit

---

<sup>16</sup> Duncan Watts, "The 'new' science of networks", *Annual Review of Sociology*, vol. 30, n° 1, 2004, p. 243-270.

<sup>17</sup> Duncan Watts, *Six Degrees: The Science of a Connected Age*, New York, Norton, 2003.

<sup>18</sup> Mark Buchanan, *Nexus: Small Worlds and the Groundbreaking Science of Networks*, New York, Norton, 2002.

bien, dans tous les cas, de parvenir à la maîtrise de plus en plus grande de l'ensemble des données accessibles. L'accumulation des traces numériques est particulièrement mise en lumière, dans la mesure où celles-ci peuvent être compilées dans un tableau général des comportements individuels et sociaux, ce qui impliquerait la potentialité de transformer notre compréhension de nos vies, de nos organisations et de nos sociétés<sup>19</sup>. La transformation de la biologie et de la physique issue de la possibilité de recueillir et analyser des masses de données a été relativement rapide par rapport à la transformation des sciences sociales.

Mais apparaît également une science sociale informatisée, fondée sur l'exploitation d'une multitude de données. L'objectif de la création d'une sorte de système nerveux planétaire correspond exactement à cette évolution : l'attitude adoptée est celle de la mesure de l'état du monde et des interactions qui s'y opèrent. Pour cela, il est nécessaire de mettre en place une extraction de données en temps réel<sup>20</sup>, en utilisant les données du web et en récoltant des données supplémentaires obtenues par la mise en relation des différents capteurs existants (notamment sur les smartphones). Ces tentatives vont dans le sens de l'établissement d'une science qui se veut totale, en rassemblant les différentes disciplines et en les dépassant grâce à l'application de quelques principes unificateurs. La physique sociale ainsi prônée repose sur l'idée que l'humain et la nature qui l'entoure ont en commun d'être soumis à des mécanismes computationnels et que les comportements des individus sont fortement prévisibles, ce qui permet d'envisager l'orientation de ces comportements afin de garantir leur conformité avec ce qui est jugé souhaitable. Les travaux rattachés à la physique sociale n'ont pas tous cette caractéristique, mais leur diffusion habitue à considérer que l'accumulation des données numériques est toujours profitable, dans une perspective que Alex Pentland envisage très précisément<sup>21</sup>. Selon lui, il est possible, à partir de l'analyse des traces numériques, de décrire mathématiquement les relations entre l'information qui circule et le comportement des individus. D'où l'idée que les *Big Data* recueillies seraient capables de donner aux chercheurs et aux politiques une vision générale de la société qui leur permettrait d'agir sur elle.

Le développement du traitement systématique des données pose des questions difficiles concernant ses relations avec l'informatique, avec les sciences humaines et sociales et avec les sciences de la nature. L'essentiel, cependant, réside dans la prise en compte des difficultés liées à la modification des modalités du processus d'interprétation, comme le montre particulièrement les recherches relevant des humanités numériques. C'est pourquoi il faut souligner que le numérique, en se rattachant au calcul, n'a besoin que de la manipulation formelle d'éléments sans signification, ce qui, comme le montre Bruno Bachimont<sup>22</sup>, situe son domaine aux antipodes de l'herméneutique. Les humanités numériques peuvent être vues comme le résultat d'un recours à l'informatique en tant qu'outil permettant de mieux analyser les objets sémiotiques étudiés, mais elles peuvent être envisagées comme impliquant des données que les procédures informatiques vont traiter pour donner des résultats que le chercheur devra interpréter. Une opposition tout aussi forte apparaît au niveau de la conception de la place du sens. D'un côté, les objets étudiés sont fondamentalement porteurs de sens, alors que, de l'autre, il s'agit de travailler à fournir d'eux une représentation formelle qui débouchera sur la possibilité d'une validation empirique. Des ruptures radicales apparaissent ainsi, entre les données et leur origine, entre la simplicité de la présentation des résultats et la complexité des

---

<sup>19</sup> David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, Marshall Van Alstyne, "Computational Social Science", *Science*, vol. 323 n° 5915, 2009, p. 721-723.

<sup>20</sup> Dirk Helbing, Stefano Balietti, Steven Bishop, Paul Lukowicz, "Understanding, creating, and managing complex techno-socio-economic systems: Challenges and perspectives", *The European Physical Journal - Special Topics*, vol. 195, 2011, p. 165-186.

<sup>21</sup> Alex Pentland, *Social physics: How Good Ideas Spread - The Lessons from a New Science*, New York, The Penguin Press, 2014.

<sup>22</sup> Bruno Bachimont, *Le sens de la technique : le numérique et le calcul*, Paris, Les Belles Lettres, 2010.



traitements, entre la diversité des données et leur traitement unifié. L'incertitude qui pèse sur la fabrication de ces données fait que le calcul qui porte sur elles pose le problème de la construction du sens. Le langage n'est, dès lors, plus central et, comme les sciences expérimentales l'ont fait pour la connaissance de la nature, les humanités numériques cherchent à produire la connaissance par le calcul, avec des avantages, certes, mais aussi avec des inconvénients.

En effet, l'examen du rapport entre la manipulation des unités de sens et leur interprétation fait surgir une difficulté centrale : non seulement il n'y a pas de relation simple entre manipulabilité et signification, mais, lorsque l'on cherche à établir une telle relation simple, on en vient à faire régir la sémantique par la syntaxe<sup>23</sup>. Ce forçage est particulièrement apparent dans le cas de la construction du web sémantique, dont l'objectif est de parvenir à une utilisation perfectionnée des ressources du web par les machines aussi bien que par les humains<sup>24</sup>. Cela implique non seulement de donner une structure à l'information en discrétisant son contenant, mais également de lui donner un sens en discrétisant son contenu, puis de donner une structure aux unités de sens. Étant donné qu'il existe plusieurs procédés de discrétisation sémantique, il est pertinent de faire l'hypothèse que chacun d'eux ne se caractérise pas uniquement par des choix techniques, mais qu'il correspond aussi à une politique du sens, élaborée par des concepteurs et des développeurs dont les stratégies méritent sans doute une attention particulière. Au nom de la calculabilité, c'est une démarche de normalisation de la communication qui se met en place, par la mise au point d'un programme d'interopérabilité entre toutes sortes de données qui postule, d'une part, l'indépendance entre les connaissances représentées et les langues dans lesquelles elles sont produites et transmises, d'autre part, leur représentabilité systématique par un formalisme logique. Comme le montre particulièrement les débats sur les ontologies, les métadonnées, etc., on peut douter de la validité du projet. Il faut noter enfin que la normalisation impliquée par la recherche de la calculabilité n'est pas la seule. Elle se double d'une tentative qui semble bien tendre à l'imposition d'un consensus : avant même toute transformation ultérieure, les données se présentent sous forme d'unités textuelles de base marquées par une neutralité apparente, mais tout à fait contestable.

Il est certain que la richesse de l'apport des sciences exactes à la connaissance du web et des réseaux en général devrait susciter un plus grand intérêt de la part des chercheurs en sciences sociales. Et, bien évidemment, celle des travaux issus des sciences sociales pourrait retenir mieux l'attention des chercheurs qui ont emprunté le chemin tracé par les premiers analystes du web. Dans ce cas, la pluridisciplinarité initialement défendue par son inventeur serait peut-être atteignable. Mais des obstacles demeurent, même si l'on fait l'hypothèse de l'existence plausible d'une situation dans laquelle on trouverait une considération réciproque, un effort pour comprendre le langage de l'autre et des barrières institutionnelles plus aisément franchissables. En effet, si l'on peut estimer que l'opposition entre les sciences « dures » et les autres doit être dépassée, grâce à la prise en compte d'une globalité intellectuelle qui serait marquée par une rationalité dépassant le cadre du calcul, encore faut-il que cette globalité ne se révèle pas réductrice. Comme le montrent les réflexions sur l'analyse des réseaux de communication, il est nécessaire qu'elle n'impose pas la transformation de la pluridisciplinarité souhaitée en une prétendue transdisciplinarité visant l'élaboration d'une science totale qui risquerait, en fait, de supprimer les approches et les questions accordant un statut particulier

---

<sup>23</sup> Serge Bouchardon, Isabelle Cailleau, Stéphane Crozat, Bruno Bachimont, Thibaud Hulin, « Explorer les possibles de l'écriture multimédia », *Les Enjeux de l'information et de la communication*, vol. 12, n° 2, 2011, p. 11-24.

<sup>24</sup> Tim Berners-Lee, James Hendler, Ora Lassila, "The Semantic Web", *Scientific American*, vol. 284, n° 5, 2001, p. 34-43 ; Guillaume Sire, « Web sémantique : les politiques du sens et la rhétorique des données », *Les Enjeux de l'information et de la communication*, vol. 19, n° 2, 2018, p. 147-160.

aux sociétés humaines, à leurs comportements et à leurs créations. La poursuite du renouvellement des problématiques et des disciplines ne peut donc qu'être souhaitée, à condition, cependant, qu'elle soit accompagnée d'une vigilance accrue à l'égard de la conception de la société que ce renouvellement pourrait favoriser.

## Bibliographie

- ADAMIC, Lada, "The small world web", *ECDL '99: Research and Advanced Technology for Digital Libraries*, 1999, p. 443-452.
- ADAMIC, Lada, HUBERMAN, Bernardo, "The nature of markets in the world wide web", *Quarterly Journal of Electronic Commerce*, vol. 1, n° 1, 2000, p. 1-10.
- BACHIMONT, Bruno, *Le sens de la technique : le numérique et le calcul*, Paris, Les Belles Lettres, 2010, 185 p.
- BARABÁSI, Albert-László, "Network Science: Luck or reason", *Nature*, vol. 489, n° 7417, 2012.
- BARABÁSI, Albert-László, *Linked: The New Science of Networks*, Cambridge, Perseus, 2002, 280 p.
- BARABÁSI, Albert-László, ALBERT, Reka, "Statistical mechanics of complex networks", *Review of Modern Physics*, vol. 74, n° 1, 2002, p. 47-97.
- BARABÁSI, Albert-László, ALBERT, Reka, JEONG, Hawoong, "Scale-free characteristics of random networks: the topology of the world wide web", *Physica A*, vol. 281, 2000, p. 69-77.
- BARABÁSI, Albert-László, ALBERT, Reka, "Emergence of scaling in random networks", *Science*, vol. 286, 1999, p. 509-512.
- BERNERS-LEE, Tim, HALL, Wendy, HENDLER, James, O'HARA, Kieron, SHADBOLT, Nigel, WEITZNER, Daniel, "A framework for Web Science", *Foundations and Trends in Web Science*, vol. 1, n° 1, 2006, p. 1-130.
- BERNERS-LEE, Tim, HENDLER, James, LASSILA, Ora, "The Semantic Web", *Scientific American*, vol. 284, n° 5, 2001, p. 34-43.
- BRODER, Andrei, KUMAR, Ravi, MAGHOUL, Farzin, RAGHAVAN, Prabhakar, RAJAGOPALAN, Sridhar, STATA, Raymie, TOMKINS, Andrew, WIENER, Janet, "Graph structure in the web", *Computer Networks*, vol. 33, 2000, p. 309-320.
- BOUCHARDON, Serge, CAILLEAU, Isabelle, CROZAT, Stéphane, BACHIMONT, Bruno, HULIN, Thibaud, « Explorer les possibles de l'écriture multimédia », *Les Enjeux de l'information et de la communication*, vol. 12, n° 2, 2011, p. 11-24.
- BUCHANAN, Mark, *Nexus: Small Worlds and the Groundbreaking Science of Networks*, New York, Norton, 2002, 235 p.
- EASLEY, David, KLEINBERG, Jon, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, Cambridge, Cambridge University Press, 2010, 727 p.
- HELBING, Dirk, BALIETTI, Stefano, BISHOP, Steven, LUKOWICZ, Paul, "Understanding, creating, and managing complex techno-socio economic systems: Challenges and perspectives", *The European Physical Journal - Special Topics*, vol. 195, 2011, p. 165-186.
- LAZER, David, PENTLAND, Alex, ADAMIC, Lada, ARAL, Sinan, BARABÁSI, Albert-László, BREWER, Devon, CHRISTAKIS, Nicholas, CONTRACTOR, Noshir, FOWLER, James, GUTMANN, Myron, JEBARA, Tony, KING, Gary, MACY, Michael, ROY, Deb, VAN ALSTYNE, Marshall, "Computational Social Science", *Science*, vol. 323 n° 5915, 2009, p. 721-723.
- LI VIGNI, Fabrizio, « Le projet épistémique des sciences de la complexité », *Philosophia Scientiæ*, vol. 24, n° 1, 2020, p. 181-200.
- PAPADOPOULOS, Fragkiskos, KITSAK, Maksim, SERRANO, M. Angeles, BOGUÑA, Marian, KRIOUKOV, Dmitri, "Popularity versus similarity in growing networks", *Nature*, vol. 489, n° 7417, 2012.

- PARROCHIA, Daniel, « Les réseaux : de la mathématique à la métaphysique », *Art Press*, hors-série n° 2, « Internet all over », 1999.
- PASTOR-SATORRAS, Romualdo, VESPIGNANI, Alessandro, *Evolution and structure of internet*, Cambridge, Cambridge University Press, 2004, 267 p.
- PENTLAND, Alex, *Social physics: How Good Ideas Spread - The Lessons from a New Science*, New York, The Penguin Press, 2014, 300 p.
- SIRE, Guillaume, « Web sémantique : les politiques du sens et la rhétorique des données », *Les Enjeux de l'information et de la communication*, vol. 19, n° 2, 2018, p. 147-160.
- SCOTT, John, « Social physics and social networks », in John Scott, Peter Carrington (dir.), *The Sage Handbook of Social Network Analysis*, Londres, Sage, 2011, p. 55-66.
- WATTS, Duncan, «The 'new' science of networks», *Annual Review of Sociology*, vol. 30, n° 1, 2004, p. 243-270.
- WATTS, Duncan, *Six Degrees: The Science of a Connected Age*, New York, Norton, 2003, 368 p.
- WATTS, Duncan, STROGATZ, Steven, «Collective dynamics of 'small-world' networks», *Nature*, n° 393, 1998, p. 440-442.