

# le tissu social/ the social fabric

## traces numériques et méthodes quali-quantitatives

### Tommaso Venturini et Bruno Latour

#### Introduction

*Tite-Live raconte qu'en 493 av. J.-C., la plèbe de Rome se retira sur le mont Sacré pour protester contre la politique du Sénat. Envoyé par les patriciens pour parlementer avec les insurgés, le consul Agrippa Menenius convainquit les plébéiens de revenir dans la cité en haranguant la foule avec son célèbre apologue, plus tard repris par La Fontaine, dont voici la version : (Les Fables III, 2) :*

*« De travailler pour lui [Gaster ou l'Estomac] les Membres se lassant, chacun d'eux résolut de vivre en gentilhomme, sans rien faire, alléguant l'exemple de Gaster. Il faudrait, disaient-ils, sans nous, qu'il vécût d'air. Nous suons, nous peinons, comme bêtes de somme ; Et pour qui ? Pour lui seul, nous n'en profitons pas ; Notre soin n'aboutit qu'à fournir ses repas. Chômons, c'est un métier qu'il veut nous faire apprendre. Ainsi dit, ainsi fait. Les Mains cessent de prendre, les Bras d'agir, les Jambes de marcher. Tous dirent à Gaster qu'il en (4) allât chercher. Ce leur fut une erreur dont ils se repentirent. Bientôt les pauvres gens tombèrent en langueur ; Il ne se forma plus de nouveau sang au coeur : Chaque Membre en souffrit : les forces se perdirent ; Par ce moyen, les Mutins virent Que celui qu'ils croyaient oisif et paresseux, A l'intérêt commun contribuait plus qu'eux. »<sup>1</sup>*

*À l'époque où Agrippa Menenius narrait son apologue, la médecine était encore loin de se doter de la technique de la scintigraphie, qui permet de suivre le déplacement et de mesurer l'accumulation dans l'organisme d'isotopes radioactifs précédemment administrés au patient. Grâce à la scintigraphie, les membres mis en scène dans l'apologue auraient pu savoir non seulement comment l'estomac répartit les aliments, par quels canaux et dans quelles proportions, mais aussi comment chaque membre coopère avec les autres dans le corps humain. Si les membres avaient connu tout cela, la conclusion de l'histoire d'Agrippa aurait pu être tout autre...*

Les technologies numériques sont supposées être en train de révolutionner les sciences sociales, comme elles ont antérieurement révolutionné les sciences naturelles. Or cela n'est pas le cas. La situation des sciences sociales ressemble plutôt à celle des sciences naturelles au XVème siècle, suite à l'introduction de la

---

<sup>1</sup> « Tempore quo in homine non ut nunc omnia in unum consentiant, sed singulis membris suum cuique consilium, suos sermo fuerit, indignatas reliquas partes sua cura, suo labore ac ministerio uentri omnia quaeri, uentrem in medio quietum nihil aliud quam datis uoluptatibus frui; conspirasse inde ne manus ad os cibum ferrent, nec os acciperet datum, nec dentes quae acciperent conficerent. Hac ira, dum uentrem fame domare uellent, ipsa una membra totumque corpus ad extremam tabem uenisse. Inde apparuisse uentris quoque haud segne ministerium esse, nec magis ali quam alere eum, reddentem in omnes corporis partes hunc quo uiuimus uigemusque, diuisum pariter in uenas maturum confecto cibo sanguinem » *Ad Urbe Condita*, II, 32.

presse typographique. Comme le décrit Elizabeth Eisenstein (1979), la presse a été l'une des causes principales de la révolution scientifique, mais non pas une cause immédiate. Plusieurs dizaines d'années après l'invention de Gutenberg, les naturalistes imprimaient encore les mêmes erreurs que celles qu'ils avaient copiées à la main.

Les sciences sociales sont aujourd'hui dans une situation comparable. Au lieu de se renouveler par le numérique, elles s'efforcent encore d'appliquer les anciennes méthodes aux nouvelles données. *Cyberculture, communautés virtuelles, identités en ligne, computer-mediated communication*, toutes ces notions ont été développées pour mettre en quarantaine la nouveauté des médias électroniques. Cette inertie, pourtant, devient chaque jour plus intenable face à la vitesse à laquelle le numérique infiltre la modernité. Grâce à leur diffusion capillaire, les médias numériques offrent beaucoup plus qu'une application pour les méthodes existantes : ils offrent une chance de refonder l'étude des phénomènes sociaux (Rogers, 2004).

Jusqu'à présent, l'accès aux phénomènes de la vie collective s'est avéré à la fois très partiel et très coûteux. À la différence de leurs collègues des sciences naturelles, les chercheurs en sciences sociales ont toujours été relativement mal équipés. Contrairement aux physiciens qui peuvent suivre des milliards d'atomes dans leurs accélérateurs ou aux biologistes qui peuvent cultiver des millions de microbes sous les lentilles de leurs microscopes, les chercheurs en sciences sociales ne peuvent pas maintenir à la fois le focus et l'étendue de leur observation. Les méthodes et les outils dont ils disposent ne leur permettent d'examiner que quelques interactions spécifiques ou de survoler la surface des structures globales.

Cela explique, entre autres, pourquoi on a été attribué à la notion d'émergence dans les sciences sociales un sens tout à fait différent de celui des sciences naturelles. En physique, en biologie, en médecine, la notion d'émergence est utilisée pour décrire les phénomènes d'auto-organisation dûs à la coordination spontanée entre une pluralité d'agents. Dans les sciences sociales, les phénomènes émergents deviennent des structures, des grilles d'organisation se surimposant aux échanges individuels, des champs de forces qui accueillent et influencent les interactions. Cette vision structuraliste et dualiste provient en grande partie du fait que les sciences sociales n'ont jamais disposé des méthodes permettant d'observer comment les phénomènes globaux sont construits par l'assemblage des interactions locales. Le numérique promet de révolutionner cette situation et d'offrir aux sciences sociales la possibilité de suivre chaque fil d'interaction et de montrer comment la trame de la vie collective est tissée par leur assemblage.

## **Le problème de la vie collective**

Le problème de l'existence sociale est que toute situation qui implique plus de deux personnes peut rapidement devenir incroyablement compliquée. Prenez la complexité d'un être humain et multipliez-la par le nombre de personnes en interaction puis par le nombre d'interactions nécessaires pour organiser le plus simple des phénomènes sociaux. Rajoutez le fait que les interactions peuvent affecter et être affectées par tout type d'agents non-humains et que ces agents peuvent transporter les effets d'autres interactions venant de très loin dans l'espace et dans le temps. Si l'on parvient à

composer avec cette complexité, c'est parce que nos sociétés peuvent compter sur un grand nombre d'outils de simplification tel que les institutions, les normes, les standards, les procédures, etc. Contrairement à ce que l'on pense, les structures sociales dites 'émergentes' sont en réalité plus simples que les interactions qui les composent.

L'*émergence* est une notion fort utile, une idée qui profite tant aux acteurs sociaux qu'aux chercheurs en sciences sociales. Elle permet aux acteurs de présupposer la stabilité des phénomènes collectifs et aux chercheurs de présumer que ces phénomènes ont un sens (Johnson, 2001). Dans les deux cas, pourtant, la notion d'émergence recouvre une signification tout à fait différente. Pour les acteurs sociaux, il n'y a pas d'autres moyens d'accéder aux propriétés émergentes que de les construire. Tous ceux qui ont essayé de monter un groupe de rock ou d'organiser un tournoi d'échecs, d'initier une association de broderie peuvent en témoigner : les phénomènes sociaux n'existent pas par eux-mêmes. Les phénomènes collectifs émergent grâce au travail de leurs participants, s'étendent jusqu'où les acteurs les transportent et persistent tant qu'ils les maintiennent.

Bien sûr, les faits sociaux peuvent être matérialisés : les décisions peuvent être transcrites, les frontières renforcées par des barrières, les relations scellées par des symboles. Les structures sociales peuvent être stabilisées par des infrastructures matérielles (Latour, 2005), mais ce n'est que par le labeur de coordination des acteurs que les phénomènes collectifs peuvent être générés et maintenus (Weisman, 2007). Ce labeur, pourtant, n'est que très rarement reconnu par les sciences sociales qui, depuis leur fondation, n'ont jamais pu disposer des ressources nécessaires pour explorer la construction des structures sociales. Elles peuvent suivre quelques interactions spécifiques, mais décomposer les faits sociaux dans les milliers d'interactions qui les constituent a toujours été au delà de leurs capacités. Ainsi, elles ont préféré contourner la question.

Au XIX<sup>ème</sup> siècle, la naissante sociologie a trouvé dans la statistique le contournement qu'il lui fallait. En combinant la tradition du calcul des probabilités avec celle de l'estimation des erreurs de mesure, la statistique a fourni aux sciences sociales un raccourci bien pratique (Desrosières, 1993). Au lieu de retracer la manière dont les structures émergentes sont construites par la simplification de milliers d'échanges locaux, les chercheurs pouvaient accéder aux phénomènes sociaux par un *saut mathématique*. Là où les acteurs devaient *assembler*, les chercheurs pouvaient *agréger*. Au lieu, par exemple, d'examiner la manière dont les organisations criminelles choisissaient et socialisaient leurs recrues, les chercheurs pouvaient calculer le profil du « criminel moyen ». Au lieu d'étudier la manière dont les réseaux criminels différenciaient leurs investissements grâce à la coordination d'une multitude de malfaiteurs individuels, les chercheurs pouvaient estimer 'l'incidence moyenne' de chaque crime. L'efficacité gagnée était prodigieuse. Non seulement les sciences sociales pouvaient finalement s'adresser aux structures sociales, mais elles pouvaient même contribuer à leur solidification.

Bien sûr, la dépendance des sciences sociales à la statistique a eu un coût. En contournant le travail de construction collective par le raccourci de la statistique, les chercheurs ont introduit une distinction fictive entre micro-interactions et macro-structures. Sans jamais avoir été autre chose qu'une illusion d'optique, cette distinction a toujours été très pratique pour les sciences sociales. Grâce à l'opposition

micro/macro, les chercheurs intéressés par les interactions locales peuvent les observer sans se mêler à l'imbroglie d'influences qui assiège chaque événement social. D'autre part, ceux qui souhaitent s'intéresser au dessin global de la vie collective, peuvent le faire sans se préoccuper de chaque coup de pinceau. En séparant micro et macro, les sciences sociales valident à la fois les méthodes qualitatives *et* les méthodes quantitatives.

L'approche dualiste fondée sur la séparation micro/macro, toutefois, a un désavantage majeur : elle ne fonctionne qu'une fois que les acteurs ont achevé leur travail de simplification et tant qu'ils arrivent à le maintenir. On ne définit des échantillons que si l'on connaît les limites d'une population. On ne retrouve des courbes normales que si les acteurs se conforment aux normes. On n'observe des comportements déviants que si la déviance est découragée ou sanctionnée. Les méthodes quantitatives peuvent décrire les structures globales seulement si celles-ci sont stabilisées par les acteurs sociaux. Les méthodes qualitatives peuvent se focaliser sur les interactions locales uniquement si celles-ci sont circonscrites par les participants. L'approche dualiste arrive à décrire assez bien les phénomènes *émergés*, mais elle est démunie face aux phénomènes *émergents*.

## **La leçon des controverses**

La faiblesse de l'approche dualiste apparaît particulièrement évidente si, au lieu de se concentrer sur les zones de consensus, on se tourne vers l'étude des controverses, pour observer la vie collective à l'endroit où la coordination n'est plus (ou pas encore) achevée. Dans l'étude des disputes, on ne peut faire confiance ni aux méthodes quantitatives (puisque les positions locales sont tellement variées que cela n'a aucun sens de les agréger dans une moyenne générale) ni aux méthodes qualitatives (car la plus spécifique des controverses fait toujours partie d'un réseau d'influences global). Là où les disputes font rage, où les acteurs n'arrivent pas à s'accorder, où les liens sociaux se dénouent ou, tout simplement, là où les changements se déroulent si vite que les anciennes institutions se décomposent avant que les nouvelles ne soient créées, les méthodes opposant quantitatif et qualitatif atteignent leurs limites.

C'est la raison pour laquelle les sciences sociales ont toujours été mal à l'aise face aux controverses : leur méthodes les obligent à les délaisser comme des phénomènes résiduels ou, tout au plus, comme des intervalles de transition entre périodes de stabilité. Cela est particulièrement évident dans la sociologie des sciences et des techniques. Longtemps, les disputes entre savants ont été considérées comme des incidents négligeables sur le chemin du développement scientifique. Même les auteurs qui, comme Thomas Kuhn (1962), ont pris au sérieux les controverses scientifiques, ont fini pour le définir comme des phases de passage entre deux périodes de « science normale ».

Le mérite des *Sciences and Technologies Studies* initiées par l'école d'Édimbourg (Bloor, 1976) et de Bath (Collins et Pinch, 1993) est d'avoir montré que, loin d'être des phénomènes résiduels ou transitoires, les controverses sont le véritable moteur des sciences. Dans les sciences et les techniques (Bijker et Law, 1992), le désaccord et le conflit sont la norme et le consensus ne peut être atteint que par un formidable labeur de coordination et de stabilisation (Callon, Lascoumes et Barthe, 2001). Les

paradigmes scientifiques ne sont pas des systèmes cohérents qui canalisent le travail des savants, mais le résultat fragile de ce travail. Leur maintien nécessite une négociation ininterrompue et leur existence est constamment menacée par le risque que cette négociation puisse échouer (Latour, 1988).

Travaillant avec les élèves de Sciences Po dans le cours de *Cartographie de controverses* ([controversies.sciences-po.fr](http://controversies.sciences-po.fr)), nous avons appris à reconnaître que, loin d'être limitée aux sciences, cette situation est commune à tout phénomène collectif. Les controverses nous ont enseigné que rien dans la vie sociale ne perdure par le changement. La vie collective ressemble moins à un long fleuve tranquille, qu'à une mer agitée par des milliers de courants qu'aucun contrat social ne pourra jamais réconcilier définitivement. Ainsi, les institutions sociales ne sont pas des structures capables de s'imposer aux individus, mais une sorte de « droit de la guerre » constamment sujet à la négociation des parties prenantes. Certes elles existent, certes elles ont des effets, mais ces effets sont le résultat des interactions, non leurs prémisses.

Le mélange entre stabilité et changement qui caractérise les controverses est impossible à démêler avec une approche dualiste (Venturini, 2009). Empêchant de décomposer les disputes sociales dans le réseau d'oppositions multiples qui les caractérisent, les méthodes quantitatives finissent par s'adresser au peu d'éléments qui font consensus entre les acteurs (et qui sont typiquement les parties les moins intéressantes d'une controverse). Empêchant de montrer comment chaque contraste est lié à un réseau d'opposition plus large, les méthodes qualitatives empêchent de comprendre comment des équilibres peuvent être temporairement négociés tout en laissant la controverse ouverte.

Étudier les controverses demande une nouvelle série de méthodes « quali-quantitatives » permettant de retracer les phénomènes sociaux tout au long des processus de construction, déconstruction et reconstruction qui les constituent.

## **La promesse de méthodes numériques**

Les méthodes quali-quantitatives ne sont pas la simple juxtaposition de l'analyse statistique avec l'observation ethnographique. Effacer la frontière micro/macro implique de se préparer à recueillir et gérer une quantité de données sans précédent. S'intéresser à la construction des phénomènes sociaux implique de suivre chacun des acteurs concernés et chacune des interactions entre eux. Un objectif impossible à atteindre jusqu'à il y a une dizaine d'années, mais qui commence à devenir de plus en plus réaliste grâce à la diffusion des technologies numériques.

L'intérêt des médias électroniques est que toute interaction qui les traverse laisse des traces qui peuvent être facilement enregistrées, massivement stockées et aisément récupérées. Jour après jour, les chercheurs découvrent de nouveaux réservoirs de données numériques: les archives publiques et privées sont avalées par la mémoire des ordinateurs, les transactions économiques migrent en ligne, les réseaux sociaux s'enracinent dans le web. La médiation numérique s'étale comme un immense papier-carbone offrant aux sciences sociales plus de données qu'elles n'en ont jamais rêvées.

Grâce à la traçabilité numérique, les chercheurs ne sont plus obligés de choisir entre la précision et l'ampleur de leurs observations : il est désormais possible de suivre une multitude d'interactions et, simultanément, de distinguer la contribution spécifique que chacune apporte à la construction des phénomènes collectifs. Nées dans une époque de pénurie, les sciences sociales entrent dans un âge d'abondance. Face à la richesse et à l'extension des nouvelles données, rien ne justifie le maintien des anciennes distinctions. Équipées avec une quantité de données comparable à celle des sciences exactes, les sciences sociales peuvent finalement corriger le strabisme de leurs méthodes : maintenir à la fois le focus et l'étendue de leur observation.

Le potentiel de cette approche quali-quantitative a été mis en lumière par le projet MACOSPOL (MAPPING Controversies on Science for POLitics). Financé par l'Union Européenne, ce projet a réuni huit laboratoires européens dans l'effort d'inventorier les sources de données numériques pour la cartographie de controverses et de tester les outils disponibles pour les traiter (Venturini, 2010).

Quoiqu'encourageants du fait la quantité de ressources repérées, les résultats du projet ont aussi révélé la relative immaturité des méthodes numériques. De plus en plus d'outils sont aujourd'hui disponibles pour l'extraction et l'analyse des données numériques et nombreux sont les chercheurs qui commencent à les employer (Lazer et al., 2009). Toutefois, le succès de ces initiatives reste limité par leur caractère intermittent. Dans la plupart des cas, les chercheurs se sont limités à numériser un passage ou l'autre de leurs recherches sans remettre en discussion la chaîne de la recherche en sciences sociales. Les résultats de cette approche sont certainement intéressants et offrent l'avantage d'être facilement comparables avec les méthodes classiques. Toutefois, une véritable transition vers les méthodes quali-quantitatives ne sera possible que lorsque la chaîne de la recherche en sciences sociales sera entièrement numérique. Pour que les nouvelles méthodes puissent déployer leur potentiel d'innovation, il faut que chaque étape du processus de recherche soit renouvelée d'une façon cohérente :

1. L'identification des silos de données devrait privilégier les archives numériques, le web, les médias et les réseaux en ligne, la documentation et la littérature numérisées et, de manière plus générale, toute sorte de traces numériques.
2. L'extraction des données devrait se baser sur la navigation assistée ou sur les différentes techniques de *collecte de données* (*crawling, querying, scraping, parsing...*).
3. L'intégration des données provenant de sources diverses (*data mashing*) devrait se réaliser grâce à l'indexation dans une base de données relationnelles.
4. L'analyse et la modélisation des données devrait s'appuyer sur des outils web ou numériques et, si possible, open source.
5. Les résultats des recherches devraient être publiés sur le web, de préférence dans des archives ouvertes et dans un format standard afin de faciliter leur réutilisation.

Comme l'on peut facilement voir, la chaîne que nous décrivons est, en effet, une boucle. À disposition sur le web, les résultats d'une enquête numérique pourraient devenir les données d'une nouvelle analyse. Encore plus important : les données numériques aussi bien que leurs outils d'extractions et d'analyse pourraient être publiés à côté des résultats, en offrant la possibilité de manipuler non seulement les conclusions d'une recherche, mais tous les passages qui y mènent. C'est là tout l'intérêt des méthodes numériques : la possibilité de remplacer la description des

procédures expérimentales avec un accès direct aux étapes qui les constituent. Dans une chaîne de recherche numérique, aucun passage n'est figé, aucune transformation n'est irréversible, aucun résultat n'est donné (et d'ailleurs aucune donnée n'est donnée non plus). Tout élément de la chaîne est mis à disposition du public et ouvert à son scrutin. Dans une recherche numérique, la navigation dans les données et la navigation dans la chaîne de traitement des données finissent par coïncider.

## L'oligoptique quali-quantitative

Dans cette possibilité d'ouverture réside la promesse des méthodes numériques. Mais pour que cette promesse soit tenue, il faut que toute la chaîne de la recherche soit numérisée. Tant que les méthodes numériques seront confinées à une partie de la chaîne, la navigation dans les paysages de données numériques ne pourra déployer tout son potentiel d'innovation. Les chaînes mixtes qui caractérisent la plupart des recherches actuelles ne permettent pas de profiter à fond des potentialités du numérique et génèrent tout sorte d'effet d'incompatibilité.

La question de la représentativité, par exemple, illustre parfaitement la difficulté de mélanger méthodes numériques et méthodes traditionnelles. Une critique que l'on adresse souvent à l'analyse des traces numériques, est que ces traces concernent un échantillon non représentatif de la société. Du point de vue des sciences sociales traditionnelles, cela est certainement vrai. On sait très bien que la population des internautes ne correspond pas à la composition de la société (les hommes, les jeunes et hauts niveaux de scolarisation sont en général surreprésentés sur internet). Toutefois, cette disproportion ne risque de biaiser la recherche que si l'on s'obstine à traiter les données numériques comme on traitait les données des sondages. L'avantage des nouvelles méthodes est qu'elles nous permettent de tracer l'assemblage des phénomènes collectifs au lieu de les d'obtenir par agrégation statistique.

La question de la représentativité se pose donc tout à fait différemment pour les données numériques. Les données numériques sont représentatives si chaque maillon de leur chaîne de traitement (identification, extraction, intégration, analyse, publication) reste adhérent au travail de construction des acteurs sociaux. C'est pourquoi il est si important de numériser toute la chaîne de la recherche en sciences sociales. Une chaîne entièrement numérisée nous permettra enfin de décrire le tissu social par la description du tissage de chacun de ses fils.

Attention : nous ne disons pas que les méthodes quali-quantitatives permettront d'aplatir toute la complexité de la vie collective. Au contraire, leur avantage est d'être assez souples pour suivre *certain*s phénomènes sociaux dans chacun de leurs plis. Aucune méthode de recherche n'offre une vision panoptique de l'existence collective et les méthodes quali-quantitatives ne font pas exception. Les méthodes numériques ne peuvent offrir qu'une vision oligoptique de la société (Latour et Hermant, 1988), ni plus ni moins partielle de celle des méthodes traditionnelles, et toutefois, pour la première fois dans l'histoire des sciences sociales, cette vision sera au moins continue.

C'est la raison pour laquelle nous avons ouvert ce chapitre par l'histoire de *Agrippa Menenius* et par la technique de la *scintigraphie*. L'avantage de la scintigraphie n'est

pas de permettre l'observation de tout phénomène biologique, mais d'en suivre quelques uns d'une façon continue, sans sauts et sans fractures. L'existence sociale ne se joue pas sur deux niveaux comme nos méthodes nous ont habitué à croire. Les interactions micro et les structures macro ne sont que deux façons différentes de regarder la même toile collective, comme les fils et la trame du tissu social. Là est toute la merveille de l'existence collective, dans cette unité générée par multiplication des différences, cette stabilité produite par accumulation de mutations, cette harmonie naissant des controverses, cet équilibre reposant sur des milliers de fractures. Les méthodes qualitatives et quantitatives nous ont trop longuement caché ce spectacle, les méthodes numériques nous ouvriront les yeux.

## **Bibliographie**

Bijker, W. E. et Law, J. (ed.) (1992), *Shaping Technology, Buiding Society*. Cambridge, Massachusetts: MIT Press.

Bloor, David (1976), *Knowledge and Social Imagery*. Chicago: Chicago University Press.

Callon, M., Lascoumes, P., et Barthe, Y. (2001), *Agir dans un monde incertain. Essai sur la démocratie technique*. Paris: Le Seuil.

Collins, H. et Pinch, T. (1993), *The Golem: What Everyone Should Know about Science*. Cambridge: Cambridge University Press.

Desrosières, A. (1993), *La politique des grands nombres: histoire de la raison statistique*. Paris: La Découverte.

Eisenstein, E. (1979), *The Printing Press as an Agent of Change*. Cambridge: Cambridge University Press.

Johnson, S. (2001), *Emergence*. New York: Touchstone.

Kuhn, T. (1962), *The Structure of Scientific Revolutions*. Chigaco: University of Chicago Press.

Latour, B. (1988), *Science in Action: How to Follow Scientists and Engineers Through Society*. Boston: Harvard University Press.

Latour, B. et Hermant, E. (1998), *Paris ville invisible*. Paris: La Découverte.

Latour, B (2005), *Reassembling the Social*. Oxford: Oxford University Press.

Lazer, D. et al. (2009), "Computational Social Science" in *Science*, Vol 323: 721-723.

Rogers, R. (2004), *Information Politics on the Web*, Cambridge, Mass: MIT Press.

Venturini, T. (2009), "Diving in Magma: How to explore controversias with actor-network theory", in *Public Understanding of Science*, May.

Venturini, T. (2010) "Building on Faults: How to represent controversias with digital methods", in *Public Understanding of Science* (under review).

Weisman, A. (2007), *The World without Us*. New York: St Martin Press.