

Spatialités algorithmiques

Boris Beauce

« Un poète doit laisser des traces de son passage, non des preuves. Seules les traces font rêver »

René Char, *La Parole en archipel*, Editions Gallimard, 1962.

Les traces numériques sont éminemment spatiales. Elles rendent compte de la dimension spatiale de l'action, comme autant de témoignages de ce qui a lieu. Plus que l'espace, ce sont d'ailleurs les spatialités qui intéressent les sciences sociales, c'est à dire non seulement « les caractéristiques de la dimension spatiale d'une réalité sociale », mais aussi « l'ensemble des actions spatiales réalisées par les opérateurs d'une société » [Lussault, 2003]. Car, toujours, nous faisons *avec* l'espace [Lussault & Stock, 2010].

En nous informant sur ce qui fut, les traces numériques nous informent aussi sur ce qui est en puissance. La virtualité recouvre ici toute sa portée, celle du monde en devenir, et non celle de l'artificialité ou de l'irréalité [Deleuze, 1968]. Les traces numériques nous informent sur ce qui a lieu, sur les pratiques spatiales individuelles, sur la spatialité d'une société. Et, de plus en plus, quelle que soit notre action spatiale, territoriale ou réticulaire, dans la rue ou sur Internet, nous laissons des traces de notre passage.

En cela, la distinction entre espace réel et espace virtuel relève d'une aporie préjudiciable, héritée d'une conception positionnelle et absolue de l'espace, en plus d'être matérialiste. Elle confond ce qui situe et ce qui est situé. L'espace, depuis Leibniz, n'est pas une chose, mais l'ordre des choses [Robinet et al., 1957; Lévy, 1994]. Internet, mais aussi l'ensemble des dispositifs de transmission numérique contemporains, s'inscrivent ainsi dans une perspective historique de réagencement continue de notre environnement, comme autant de dispositifs de canalisation distribuée de l'énergie, permettant potentiellement l'interaction immatérielle à l'échelle de la planète.

Selon cette perspective, qui appréhende l'espace comme n'étant rien d'autre que l'agencement particulier du réel, la distance apparaît comme l'un des concepts les plus

fondamentaux pour appréhender les phénomènes sociaux [Lévy, 1994; Lévy, 1999]. C'est effectivement la distance qui permet de penser le singulier, l'ordre des choses, la relation et l'interaction. La distance, mais aussi les moyens de la réduire ou de l'augmenter, se trouvent ainsi au cœur de l'action, qui est toujours localisée (absolue) et plus encore de l'interaction, qui est toujours située (relative), conçue comme l'ensemble des relations établies entre les réalités qui la rende effective.

C'est pourquoi les moyens de transport ou de transmission constituent de puissants vecteurs de transformations sociales, à la fois comme projet, mais aussi comme modalités pratiques de l'interaction sociale. L'évolution de ces moyens réagence effectivement l'ordre potentiel des choses, dont l'actualisation change non seulement l'espace, mais aussi la société, c'est-à-dire l'ensemble des réalités sociales qui partagent activement un monde commun [Beaude, 2012].

Ainsi, après la *synchronisation* du monde, nous assistons à sa *synchronisation*, c'est-à-dire à l'élaboration non seulement d'un temps, mais aussi d'un lieu en commun pour l'humanité [Beaude, 2012; Beaude, 2013]. La circulation et le contrôle de l'information numérique sont à présent si rapides, qu'il est possible d'être en contact, d'interagir, de collaborer et de partager à l'échelle de l'humanité [Mitchell, 1996; Castells, 2001; Benkler, 2006]. Les traces numériques s'inscrivent dans cette spatialité complexe, qui associe des individus et des objets, à des vitesses tout à fait inédites, produisant un environnement dont les virtualités sont profondément transformées.

Visualiser les spatialités

La composante territoriale de ces traces est effectivement particulièrement décisive pour renouveler la connaissance que nous avons des espaces tels qu'ils sont habités [Gonzalez-Bailon, 2013]. L'enjeu principal des traces numériques est un enjeu de visibilité. L'individualisation des traces permet en effet de désagréger la spatialité pour en saisir les composantes individuelles, mais aussi leur temporalité et leur déploiement.

La spatialité peut ainsi être saisie dans toute sa puissance, comme agrégation d'actes élémentaires aux intentionnalités multiples. La précision et la quantité des traces numériques constituent en effet une opportunité stimulante de saisir l'espace comme un agencement dynamique de réalités sociales, octroyant une visibilité inédite des pratiques spatiales selon des perspectives novatrices.

Les traces numériques complètent en cela les données plus conventionnelles, issues des recensements ou d'enquêtes spécifiques, sélectives et ponctuelles¹. Car si les traces numériques ne respectent pas de nombreuses exigences de la statistique officielle, elles présentent en revanche un potentiel suffisamment important pour que leur usage s'impose

¹ L'exemple de *Fragile Success*, projet de visualisation développé par la Regional Plan Association, présente la région métropolitaine de New York avec une précision remarquable, agrégeant de très nombreuses données publiques, relatives à l'emploi, les transports, la performance des écoles publiques ou la criminalité (<http://fragile-success.rpa.org/>). Ce projet s'inscrit dans une perspective de valorisation innovante de données conventionnelles, soulignant la qualité et l'importance de telles données.

dans le marketing, dans la recherche scientifique, dans l'urbanisme et plus généralement dans toutes les activités qui exigent de connaître précisément un environnement².

Explorer les traces

Parmi les traces numériques exploitées, nombreuses sont celles qui ont trait à l'espace urbain en particulier. Beaucoup ne sont que des démonstrations de compétences, ne faisant pas l'objet de recherches approfondies, à l'exemple de *World Touristness map* proposée en 2009 par Bluemoon Interactive, qui exploitait les données de Panoramio. C'est aussi le cas de *Vizualizing Facebook Friends*, la célèbre carte des 10 millions « d'amitiés » que Paul Bulter réalisa en 2010. Notons aussi les *New City Landscape Maps* de Fabian Neuhaus, « tweetography » qui propose des cartes de densité de tweets au sein de quelques villes, dont Londres, New York, Munich et Paris.

D'autres projets relèvent plus de la recherche artistique, à l'image des travaux d'Esther Plolak, dont *Real Time Amsterdam* (2002) et *Elsatic Mapping* (2009) furent particulièrement novateurs. Notons aussi les surprenantes cartes de *biomapping* de Christian Nold, qui propose depuis 2004 des représentations de la variation des émotions dans le contexte urbain.

Le design, enfin, occupe une place non négligeable de ces productions visuelles, à l'image d'Anil Bawa-Cavia, qui utilisa les données de Foursquare pour *Archipelago* (2010), ou des prolifiques et stimulantes cartes d'Éric Fischer (*Local et Tourist*³, 2010 ; *A Day of Muni*, 2011 ; *See Something or Say Something*, 2011), qui utilisèrent respectivement les données de Flickr, NextBus, et une association de Flickr et de Twitter.

Enfin, d'autres productions furent plus commerciales, à l'image d'*UrbanMobs* (2008), présenté par Orange et faberNovel, dont les représentations animées exploitaient l'activité continue des téléphones cellulaires. Mais dans l'ensemble, tous ces travaux relèvent principalement d'un exercice de style ayant trait à la visualisation esthétisante de données et non à la production de représentations de l'espace s'inscrivant dans un processus de découverte scientifique, qui suppose une caractérisation plus systématique de ce qui est donné à voir. Ces artistes, infographistes ou designers, ont néanmoins largement stimulé l'exploitation des traces numériques et souligné leur potentiel remarquable.

² L'opposition entre *big data* et *small data*, mais aussi entre *soft data* et *hard data* illustre très bien ce débat. Elle questionne la valeur relative des traces numériques dans un environnement déjà largement pourvue d'informations spatiales.

³ Ce travail remarquable portait initialement sur Flickr (<https://flic.kr/s/aHsjqXbTjG>), avant d'être adapté à Twitter à l'échelle du monde dans le cadre d'une collaboration avec Mapbox. (<https://www.mapbox.com/labs/twitter-gnip/locals/>), Le concept fut emprunté à *Tracing the Visitor's eye* (2006) de Fabien Girardin, qui portait sur la localisation distincte des photos publiées sur Flickr par les touristes et les habitants de Barcelone (<http://www.girardin.org/fabien/tracing/>).

À la recherche des traces

À l'exception de quelques pionniers, la recherche académique a tardé à s'intéresser à ce sujet. À présent, les traces numériques font l'objet d'une attention majeure, qui ne fait que s'accroître ces dernières années. Dès 2008, quelques travaux ont particulièrement fait sensation, lorsqu'ils furent publiés dans *Nature* et *Science*, alors qu'ils se limitaient pourtant à constater que les mobilités individuelles sont très largement prévisibles. Portant sur de très nombreuses données (100 000 personnes pendant six mois [González et al., 2008] et 50 000 individus pendant 3 mois [Song et al., 2010]), ces recherches traitent les individus comme des atomes, produisant des analyses certes convaincantes, mais dont les limites sont évidentes : le travail, l'école et le logement sont des pratiques si structurantes, et si localisées, qu'il n'est pas surprenant que notre localisation soit si prévisible.

Un ensemble de recherches plus vastes porte sur des analyses détaillées des pratiques, ne s'intéressant pas tant à la prévision des mobilités qu'à leur inégale répartition dans l'espace. Une large impulsion fut lancée en 2006, avec quelques projets menés par des groupes de recherche structurés autour de Carlo Ratti et Sandy Pentland, tous deux chercheurs au MIT [Eagle & Pentland, 2006; O'Neill et al., 2006; Ratti et al., 2006]. Ces recherches, portant respectivement sur des étudiants du MIT, des passants de la ville de Bath et des habitants de Milan, ont initié un renouveau de la visualisation des mobilités urbaines.

Avec les années, les traces exploitées lors de telles recherches se sont émancipées de la contrainte relative à l'utilisation de dispositifs dédiés, dont les limites et les biais furent nombreux. Aussi, il apparaît que les chercheurs ayant publié sur le sujet ont quasi systématiquement exploité le même type de traces. Jusqu'à 2011, avant que les recherches ne se multiplient, les six sources les plus importantes furent les suivantes :

1. les données relatives aux *téléphones cellulaires*, de loin les plus nombreuses [Reades et al., 2007; Shoval & Isaacson, 2007; Girardin, Calabrese, Fiore et al., 2008; Girardin, 2009; Bayir et al., 2009; Eagle & Pentland, 2009; Reades et al., 2009; Vaccari et al., 2009; Calabrese et al., 2010; Ahas et al., 2010; Olteanu et al., 2011].
2. les méta-données (dénominations de lieux, coordonnées géographiques et temporelles) associées à des *photos* présentes sur des sites internet tels que Flickr. [Rattenbury et al., 2007; Girardin, Calabrese, Fiore et al., 2008; Girardin, 2009; Crandall et al., 2009; Vaccari et al., 2009; Hollenstein & Purves, 2011; Purves & Edwardes, 2011]
3. les données *GPS* (dispositifs embarqués, taxis, bus) [Ratti et al., 2006; Froehlich et al., 2006; Shoval & Isaacson, 2007; Vaccari et al., 2009; Neuhaus, 2010].
4. les données des capteurs *Bluetooth* détectant la présence de dispositifs munis de cette fonctionnalité [Eagle & Pentland, 2006; Bayir et al., 2009; Eagle & Pentland, 2009].
5. les données fournies par les opérateurs de *vélos* en libre service [Froehlich et al., 2008; Borgnat et al., 2009] et *Revealing Paris Through Velib' Data*, projet expérimental de Fabien Girardin présenté en 2008.

6. les données présentes sur des *réseaux sociaux* à composante spatiale [Phithakkitnukoon & Olivier, 2011].

Ces traces furent parfois complétées par d'autres informations :

- des bases de données relatives à la présence de lieux spécifiques (commerces, cafés...), afin de contextualiser les traces [Cortright, 2009; Reades et al., 2009].

- des enquêtes (souvent limitées) auprès des individus [Eagle & Pentland, 2006; Froehlich et al., 2006; Bayir et al., 2009; Ahas et al., 2010].

Dans l'ensemble, ces recherches présentent des travaux très innovants qui contribuent significativement à une meilleure lisibilité de l'environnement urbain et plus précisément de son inégale pratique par les individus. Elles proposent des représentations originales et souvent inédites de l'espace urbain, le révélant au travers de pratiques singulières (déplacements, prise de photos, partage ou historique de lieux fréquentés...).

Ces dernières années, d'autres démarches peuvent être soulignées, tant elles renouvellent les modalités pratiques de la connaissance et de la représentation de l'espace. Dans le prolongement de la notion de *Cellular Census* proposée dès 2007 par Jonathan Reades, Francisco Calabrese, Andres Sevtsuk et Carlo Ratti [Reades et al., 2007], certaines initiatives proposent ouvertement de se substituer aux recensements, jugés trop longs, peu représentatifs et trop coûteux, en exploitant systématiquement et massivement les données relatives à la localisation des téléphones cellulaires [Deville et al., 2014].

Partant de l'hypothèse selon laquelle les communications téléphoniques se font d'autant plus rares que les interlocuteurs sont éloignés [Lambiotte et al., 2008; Krings et al., 2009], des chercheurs de l'Université catholique de Louvain, autour de Vincent Blondel, et du MIT, autour de Carlo Ratti, ont par ailleurs suggéré de reconsidérer les délimitations régionales sur la base des appels téléphoniques [Blondel et al., 2010; Ratti et al., 2010].

À une échelle plus fine, des chercheurs se sont appuyés sur les réseaux socio-spatiaux tels que Foursquare pour identifier des quartiers intra-urbains [Cranshaw et al., 2012], et ont proposé d'éprouver le concept au sein de quelques villes nord-américaines dans le cadre du projet *Livehood*. Plus récemment, des recherches ont été menées à partir de bornes Bluetooth afin d'analyser précisément les parcours de touristes dans le Musée du Louvre [Yoshimura et al., 2014].

Perspectives des traces

Du fait de leur technicité avancée, ces recherches sont quasi exclusivement menées hors des sciences sociales, ce qui limite grandement les moyens d'en qualifier les données et plus généralement la pertinence, dès lors qu'elles rendent compte de pratiques éminemment sociales. Or, ces recherches sont limitées par les nombreux biais inhérents aux méthodes utilisées, ainsi que par l'inégale fiabilité des données primaires.

Ces projets furent souvent réalisés dans le cadre d'événements, biennales, fêtes de la musique, et n'ont pas été développées au-delà, accentuant la tendance de ces initiatives à proposer un

instantané de la ville qui en montre l'exceptionnel plutôt que l'usuel (*Amsterdam RealTime*, *Real Time Rome*, *UrbanMobs*).

Des recherches s'inscrivant plus ouvertement dans les sciences sociales tendent néanmoins à émerger, mais avec beaucoup de retard, et un accès aux données souvent plus difficile. Des recherches innovantes sont parfois menées, à l'exemple du projet *Imagitour*, qui exploite les données de Flickr, Panoramio, Facebook et Tripadvisor, associant informatique et ethnologie, afin de mieux comprendre les pratiques touristiques françaises, leurs spatialités et leurs temporalités. De telles recherches restent néanmoins relativement rares.

Pourtant, les traces numériques se multiplient avec la mise à disposition de données relatives aux transports par de nombreuses villes, dont l'exemple des coordonnées GPS des taxis à New York compte parmi les plus spectaculaires⁴. Plus exceptionnellement, des données relatives aux téléphones cellulaires sont elles aussi valorisées, à l'image du projet *D4D* (Data for Development) en Côte d'Ivoire puis au Sénégal.

Cette prolifération des traces numériques se traduit aussi par des partenariats de plus en plus nombreux entre des villes et des entreprises telles que Waze⁵ et Uber⁶, à l'image des nombreuses initiatives de Boston. Enfin, les capteurs d'activités constituent de plus en plus de traces des mobilités piétonnes, autorisant des visualisations particulièrement fines⁷. Depuis peu, l'exploitation des traces numériques est aussi envisagée pour hiérarchiser l'importance des voies de circulation, identifier automatiquement leurs sens, mais aussi leurs limitations de vitesse⁸.

Panoptique, catoptique, oligoptique, trois configurations de la visibilité.

La pluralité des traces numériques produit ainsi de nouvelles représentations de l'espace, qui informent non seulement sur l'espace, mais aussi sur les individus. Elles renouvellent puissamment les conditions de la visibilité, en partant d'actes isolés, ponctuels et individuels. Les vues synoptiques ainsi produites exigent un processus de transformation, d'agrégation, d'anonymisation, inégalement efficace, mais remarquablement signifiant. Cette dynamique se traduit par une forme d'*hypervisibilité* et de traçabilité généralisée [Lussault, 2009 :186-203], accentuée par la *synchorisation* des traces numériques, alors que leur circulation, leur traitement et leur représentation profitent pleinement de l'*hyperspatialité* contemporaine,

⁴ Le projet HubCab du *Senseable City Lab* témoigne parfaitement de ce potentiel.

⁵ Dès 2013, Google a racheté Waze 1,13 milliards de dollars et a investi dans Uber par l'intermédiaire du fond de placement Google Venture.

⁶ Uber communique explicitement sur ce point (<http://blog.uber.com/city-data>).

⁷ Les nombreuses cartes proposées par Human Data témoignent parfaitement de ce potentiel (<http://cities.human.co>).

⁸ Pour plus de détail, se référer à la collaboration entre OpenStreetMap et Mapbox (<https://www.mapbox.com/blog/mining-probe-data/>).

c'est-à-dire de la systématisation du principe de connexion entre réalités spatiales [Lussault, 2013].

Ce processus d'*hypervisibilité* autorise une *omniveillance*, dont les modalités pratiques recouvrent les qualités respectives du *panoptique*, du *catoptique* et des *oligooptiques*. Le panoptique, depuis la figure architecturale suggérée par la perspective utilitariste de Jeremy Bentham à la fin du 18^e siècle, et plus encore depuis sa déclinaison dans les années soixante-dix à l'ensemble des structures spatiales de contrôle par Michel Foucault [Foucault, 1975], permet effectivement de reconsidérer la puissance marginale des hétérotopies, comme la normalité des spatialités contemporaines. Avant même la prolifération des traces numériques, dans *Post-scriptum sur les sociétés de contrôle*, Gilles Deleuze souligna ce glissement vers un contrôle spatialement distribué [Deleuze, 1990]. Aux prémices des espaces numériques, l'architecte William Mitchell soulignait déjà la capacité du bracelet électronique à étendre la prison à l'ensemble de l'espace [Mitchell, 1996 :77], tout en y substituant la puissance des traces.

Nous retrouvons ici la lecture surplombante de l'autorité, *surveillante*, qui veille sur ses sujets pour mieux les contrôler [Graham, 2005]. Rares, néanmoins, sont ceux qui jouissent réellement d'une position suffisamment centrale, pour éprouver la puissance du panoptique. L'*hyperspatialité* se traduit ainsi par une *hypercentralité*, c'est-à-dire une concentration de la connexité et par continuité de la spatialité numérique, au profit d'un nombre très limité d'acteurs [Beaude, 2012; Beaude, 2014]. Google, Facebook et le gouvernement des États-Unis sont effectivement de très loin les mieux placés pour observer le Monde contemporain, disposant d'une quantité d'information tout à fait inédite, qu'aucun État et qu'aucune entreprise n'aurait imaginé il y a à peine trente ans. Ce constat souligne à quel point coexistent deux conceptions opposées et en partie convergentes de la surveillance individuelle et légitime, l'une et l'autre étant néanmoins de plus en plus contestée. En quelques années à peine, la surveillance s'est effectivement généralisée à tout individu, toujours conjointement criminel et client potentiel, enjoignant à une transparence que seuls les suspects devraient craindre⁹.

Cette perspective peut néanmoins être inversée, dès lors qu'Internet décentralise autant qu'il centralise les points de vue sur le Monde. Internet, au moins virtuellement, permet à chacun de voir le Monde avec une rare intensité. La surveillance n'est effectivement plus exclusivement surplombante, mais potentiellement en tout individu, telle une *sousveillance*, distribuant horizontalement la veillance, déployant une appréciation collective et généralisée de ce qui convient. Ainsi, tel un panoptique inversé, nous assistons à l'émergence d'un *catoptique* [Ganascia, 2009], un dispositif de médiation qui redistribue si puissamment les observations, qu'il place chacun en position de voir chaque autre. Les plateformes telles que Yelp ou Tripadvisor, bien que limitées à des pratiques tout à fait circonscrites, ont

⁹ Eric Schmidt, alors dirigeant de Google, résuma parfaitement cette logique lorsqu'il suggéra en 2009 à une journaliste de CNBC que si nous avons un fait à cacher, peut être que nous n'aurions tout simplement pas dû le réaliser.

effectivement radicalement renouvelé les modalités pratiques de la qualification de l'espace, redistribuant l'acte d'évaluation, de représentation et de consultation. Le *catoptique*¹⁰ résume cette visibilité redoutable de chaque instant, de chaque acte qui, surpris par quiconque, peut être instantanément exposé au Monde.

Aussi complexes que puissent être de tels dispositifs de visualisation des pratiques individuelles et collectives, il convient de souligner à quel point leur puissance et leur intensité n'a d'égal que leurs lacunes. De même que l'omniprésence des caméras de surveillance initiée au Royaume-Uni pouvait donner l'illusion d'une complétude [Thrift & French, 2002; Dodge & Kitchin, 2005; Graham, 2005; Klauser, 2010], leurs images ne sont pourtant qu'une multitude de cadres, bornés et limités par leur champ. Toujours, ces visions sont partielles, comme autant d'oligoptiques [Latour & Hermant, 1998], comme autant « d'étroites fenêtres qui permettent de se relier, par un certain nombre de conduits étroits, à quelques aspects seulement des êtres (humains et non humains) dont l'ensemble compose la ville » [Latour, 2007].

Google ou la NSA s'évertuent à croiser les oligoptiques, mais leur association ne constitue jamais un panoptique total. Aussi, la carte n'est jamais le territoire, et l'on sait depuis Jorge Luis Borges que même à l'échelle 1:1, la carte serait insatisfaisante. C'est ainsi que « l'utopie de la représentation du monde du personnage bourgeois s'éloigne » [Nova, 2009 :150].

Big brother, big data

Malgré leur prolifération, l'illusion de réalité que de telles représentations du Monde octroient à ceux qui les maîtrisent comporte effectivement le risque d'un réductionnisme redoutable, qui réduirait le Monde à ses traces [Beaude, 2010; Gordon & Silva, 2011].

La critique acerbe de George Orwell à l'égard de la dérive totalitaire des dispositifs de surveillance est datée. Elle peut néanmoins être reconsidérée à l'aune de ce renouvellement de la surveillance [boyd & Crawford, 2012 :664]. Le pouvoir de propagande de Big Brother, parfaitement incarné par la figure du grand frère, repose sur l'ambivalence du regard, comme attention et comme contrôle. Or, au-delà de la veillance, il importe de rappeler que Big Brother s'assure conjointement de l'appauvrissement du monde symbolique, pour mieux le contrôler.

Or, la généralisation des dispositifs de production, de collecte et de traitement des traces numériques associe étroitement ces trois composantes (attention, contrôle et appauvrissement). Elle s'inscrit dans des logiques de surveillance (NSA) et de services (Google et Facebook), en partie consentis, tout en comportant un risque significatif de réduction des pratiques individuelles à leurs simples traces.

Les traces numériques constituent sans aucun doute un potentiel inédit pour les sciences qui s'intéressent à la dimension spatiale du social. Elles permettent d'accéder à une connaissance des pratiques spatiales, plus individualisée, plus continue et plus intense. Ainsi, la

¹⁰ Ce terme est emprunté à la catoptrique, discipline qui étudie la réflexion de la lumière.

prolifération des traces numériques et la généralisation de leur traitement automatique répondent plus précisément aux motivations de l'empirisme, du pragmatisme, de l'utilitarisme et du positivisme.

Héritages

L'exploitation systématique et générale des faits, par opposition aux idées, divisait déjà Platon et Aristote. Cette opposition traversa d'ailleurs l'ensemble de la philosophie, discutant inlassablement la pertinence de ce que Leibniz appelait les vérités de raison et les vérités de fait. La prévalence des faits sur la raison, la critique du rationalisme, la valorisation de l'expérience et la primauté de l'empirisme émergèrent vraiment à la fin du 17^e siècle et au début du 18^e siècle en Irlande et en Angleterre, avec John Locke, puis George Berkeley. Dès la fin du 18^e siècle, cette distinction fut néanmoins remise en cause, avec Kant, soulignant plus précisément l'intrication des faits et de leurs interprétations, et l'inefficacité de leur distinction. Le constructivisme, et plus précisément le co-constructivisme proposé par Edgar Morin à la fin du 20^e siècle, clarifia cette opposition, que Bruno Latour résume habilement par la distinction entre les données et les obtenus [Latour 1999; Carmes & Noyer 2014], soulignant le processus actif de co-construction des faits scientifiques.

C'est probablement ce pan de la philosophie et de l'épistémologie contemporaines qui fait actuellement défaut au paradigme des *big data* et de la *social physics*¹¹ proposée par Alex Pentland [Pentland, 2014], l'un des chercheurs les plus prolifiques dans l'exploitation des traces numériques. Le paradigme émergent avec les *big data* articule ensemble une pensée particulière (des faits, des réseaux, des individus...) et la mobilisation de dispositifs techniques spécifiques (base de données, corrélations statistiques...).

Ce paradigme s'inscrit dans un tout autre héritage, qui trouve ses fondements dans le saint-simonisme, qui valorisait déjà la puissance des réseaux [Musso, 1997], le positivisme d'Auguste Comte et son aspiration à l'émergence d'une physique sociale [Comte, 1995 (1842)], la monadologie de Gabriel Tarde, et son « épidémiologie » des idées [Tarde, 2001 (1890)], la cybernétique de Norbert Wiener, plus particulièrement l'enjeu de sa transposition à la société comme modalité idéale de gouvernance [Wiener, 1950], et le « village global » de McLuhan, en particulier l'hypothèse de la régulation des pratiques avec l'avènement d'un réseau mondial de capteurs [McLuhan, 1969].

Pourtant, dans le dernier ouvrage d'Alex Pentland, *Social Physics*, mais aussi dans la majeure partie des travaux portant sur le potentiel des *big data* [Mayer-Schonberger & Cukier, 2013],

¹¹ Afin de ne pas créer de confusion entre la physique sociale de la fin du 19^e et celle du début du 21^e siècle, l'expression anglophone sera maintenue pour la proposition d'Alex Pentland. « Social physics is a quantitative social science that describes reliable, mathematical connections between information and idea flow on the one hand and people 's behavior on the other hand ». « Social physics helps us understand how ideas flow from person to person through the mechanism of social learning and how this flow of ideas ends up shaping the norms, productivity, and creative outputs of our companies, cities, and societies » [Pentland, 2014 :4].

les références à cet héritage sont inexistantes. Pierre Bourdieu, Antony Giddens, Michel Foucault ou Jacques Derrida, penseurs du social parmi les plus reconnus au 20e siècle, sont aussi absents de tels édifices théoriques. Alex Pentland semble construire son projet scientifique en réaction à Adam Smith et à Karl Marx, auquel il se réfère abondamment, réduisant les sciences sociales à la faiblesse des structures qui réduiraient et déshumaniseraient les individus, ce que la sociologie contemporaine a pourtant déjà abondamment critiqué ces dernières décennies.

Cette omission peut s'expliquer par le fonctionnement disciplinaire de la pratique scientifique, mais la transposition de la physique à la société mériterait peut-être plus d'attention. Ce paradigme naissant a en effet non seulement l'ambition de permettre aux sciences sociales d'accéder à l'efficacité des autres sciences, mais aussi de renouveler profondément la politique [Pentland, 2014 :176-192]. Plus subtilement, et non sans une certaine incohérence, cette disposition favorable aux traces numériques apparaît aussi dans les textes récents de Bruno Latour [Latour, 2010 151-183; Latour et al., 2012], à l'intersection d'une approche relationnelle et computationnelle du social, puisant aussi aux prémices de la sociologie, et plus précisément de la monadologie de Gabriel Tarde.

La rupture suggérée suppose que la quantité tend vers l'exhaustivité, et que l'accès à de telles données renouvelle profondément les sciences sociales [Lazer et al., 2009; Song et al., 2010; Pentland, 2014], permettant de « matérialiser » le social, de l'observer enfin [Venturini & Latour, 2010]. L'exhaustivité constituerait en quelque sorte une étape ultime, qui remettrait en cause deux principes fondamentaux de la science : la représentativité et la reproductibilité ($n=all$) [Mayer-Schonberger & Cukier, 2013].

Ce qui est communément appelé *big data* constitue en cela une opportunité pour les sciences sociales qui encourage à repenser les conditions empiriques de la recherche [Kitchin, 2013; Ruppert, 2013; Graham & Shelton, 2013; Severo & Romele, 2014]. Cette approche permettrait de désagréger le social en évitant les catégories réductrices telles que les classes, tout en reliant les individus à l'ensemble des réalités qui les constituent [Pentland, 2014; Latour et al., 2012].

Précautions

Il convient néanmoins d'être prudent à l'égard d'un tel ensemble de faisceaux convergents, tant l'histoire des sciences et en particulier de la statistique fut déjà riche de telles aspirations [Barnes, 2013]. Les approches quantitative, physique et informatique du monde social ont effectivement été largement éprouvées et critiquées, à l'image de la première loi de la géographie proposée par Waldo Tobler, selon laquelle tout est relié à tout, mais les choses proches sont plus reliées que les choses distantes¹² [Tobler, 1970].

¹² Alex Pentland arrive à la même conclusion après avoir étudié la localisation des transactions de cartes de crédit de la moitié de la population active des États-Unis [Pentland, 2014 :160].

Après la provocation de Chris Anderson, lorsqu'il annonçait la fin des théories avec l'avènement des *big data* [Anderson, 2008], les critiques furent d'ailleurs vives à l'égard d'une telle hypothèse [boyd & Crawford, 2012; M. Graham & Shelton, 2013]. Aussi, le potentiel des *big data* doit être d'autant plus discuté que l'une de ses figures les plus emblématiques, Google Flu Trends, se révèle être déjà inopérant à prédire la grippe¹³ [Butler, 2013; Lazer et al., 2014], alors qu'il fut régulièrement invoqué comme exemple de la puissance du traitement massif de données sans a priori [Ginsberg et al., 2009; Mayer-Schonberger & Cukier, 2013]. L'exemple de Google Flu Trends n'est pourtant pas anodin, car lors de sa première version, il était particulièrement conforme aux exigences des *big data* : quantité considérable de traces numériques territoriales¹⁴ passives associées à des corrélations massives, sans hypothèses et sans a priori¹⁵. En 2014, alors que l'algorithme ne fonctionnait plus depuis 2009, Pentland y fait pourtant encore référence comme une réussite qui ne ferait qu'effleurer le potentiel de ces données [Pentland, 2014].

Force est de constater que malgré leurs limites, ce type de recherches abonde dans les revues scientifiques les plus prestigieuses (Nature et Science en particulier). Le plus surprenant est certainement que l'échec de Google Flu Trends était pourtant largement prévisible. Les causes énoncées pour expliquer les erreurs récurrentes de ce projet, qui ne s'appuie que sur des corrélations, ont d'ailleurs été décriées par des chercheurs en sciences sociales avant même que les écarts ne deviennent manifestes et injustifiables. Ces dernières années, de nombreux chercheurs ont effectivement dénoncé la mythologie sous-jacente à de telles approches, n'hésitant pas à rappeler les fondements de la démarche scientifique, et des journalistes s'autorisent légitimement à rappeler quelques principes élémentaires de la statistique [Harford, 2014]. La critique la plus aboutie de cette dérive est probablement celle de danah boyd et Kate Crawford [boyd & Crawford, 2011; boyd & Crawford, 2012]. Sans exception, les précautions qu'elles suggèrent s'appliquent parfaitement au cas plus particulier des traces numériques territoriales. De l'évolution de la pratique scientifique aux problématiques de vie privée ou d'accessibilité des données, en passant par les illusions de représentativité ou d'objectivité, danah boyd et Kate Crawford inscrivent les *big data* dans une filiation dont les exigences ne sauraient être si aisément remises en cause.

La propension des *big data* à changer la science est effectivement remarquable, mais en des termes différents de ceux suggérés par Chris Anderson, David Lazer ou Alex Pentland. En

¹³ Le modèle ne fut pas efficace dès 2009 et, malgré ses ajustements, il ne se révèle pas plus performant que de simples modèles s'appuyant sur une projection des données du CDC (Centers for Disease Control and Prevention) [Lazer, 2014]. Lors de la saison 2011/2012, l'erreur était de plus de 50%. Lors de la saison 2012/2013, l'erreur fut supérieure à 100%.

¹⁴ Bien que GFT ne soit généralement pas perçu comme territorial, il s'agit bien de traces numériques territoriales. GFT crée des analogies entre les recherches et leurs localisations, sans lesquelles l'épidémiologie ne serait pas envisageable.

¹⁵ Nous pouvons déjà émettre l'hypothèse selon laquelle GFT progressera avec l'introduction d'hypothèses et d'*a priori* plus explicites.

s'inscrivant plutôt dans une sociologie des sciences qui interroge la relation entre les faits et les énoncés scientifiques, par la médiation des instruments et des chercheurs, ce changement ne signifie pas l'émergence d'un paradigme ultime et l'accès définitif à l'objectivité du monde, mais plus simplement l'évolution des constructions subjectives, renouvelées par l'émergence de nouvelles données et de nouveaux outils pour les traiter.

Ce constat explique probablement pourquoi nous retrouvons avec les *big data* des engouements très semblables à ceux qui se manifestèrent avec l'émergence de la statistique, de l'informatique et des systèmes d'information géographique. À chacun de ces moments, l'amélioration des modalités pratiques de l'expérience empirique du monde se traduit par l'illusion d'une lisibilité ultime, dont la société serait changée, mais aussi considérablement améliorée.

Conjonctions de corrélations

L'objectivité supposée des traces numériques doit ainsi être questionnée, tant la confusion entre la quantification et l'objectivation est étonnamment récurrente. L'activité des chercheurs exige une réflexivité à l'égard de la quantification qui se traduit par une interprétation et non par une simple « lecture » de faits qui parleraient d'eux-mêmes. Cette critique du sens commun¹⁶ souligne que le vraisemblable n'est pas plus vrai que l'invraisemblable.

La corrélation n'est aucunement une relation, si ce n'est statistique, et moins encore une cause. S'il est de plus en plus simple de distinguer statistiquement une cause d'une conséquence [Mooij et al., 2014], il est en revanche toujours impossible d'avoir l'assurance qu'une corrélation soit une relation effective entre deux phénomènes. En bénéficiant d'une individualisation importante des données, les traces numériques donnent l'illusion de se soustraire non seulement à l'erreur écologique (biais d'agrégation), mais aussi à la faiblesse des approches corrélatives. Le réductionnisme à l'œuvre dans l'exploitation hâtive des traces numériques ne se résume effectivement pas à la réduction des faits sociaux, mais aussi à celle de leurs relations supposées. Lorsque l'exploitation de ces données ne se limite pas à des approches descriptives, elles s'appuient très largement sur la corrélation entre indicateurs, avec l'assurance que la quantité des données assure des corrélations plus significatives. Si cette considération est vraie d'un point de vue statistique, la portée de sa valeur scientifique est relativement faible. En sus du constat éprouvé selon lequel les corrélations ne permettent pas d'établir des relations effectives entre des phénomènes, il importe de considérer que plus le nombre de variables est important, plus les corrélations fallacieuses sont nombreuses [Ioannidis, 2005].

Dès le 18^e siècle, Hume a pourtant parfaitement souligné cette difficulté, lorsqu'il insista sur le fait que nous n'avons jamais l'assurance que des conjonctions puissent être des causes. Les traces numériques territoriales n'échappent pas à cette exigence. C'est pourquoi la

¹⁶ Le physicien Duncan Watts, devenu sociologue, souligne bien cet enjeu particulier du sens commun en sciences sociales, et la difficulté de s'y soustraire [Watts, 2011].

prolifération des données spatiales exige des analyses et des interprétations particulièrement délicates et requiert des précautions trop souvent négligées.

Représentativités représentatives

Aussi, au-delà de la quantification, les *big data* interrogent la quantité, suggérant que la qualité émergerait de la quantité lorsque celle-ci tendrait vers l'exhaustivité. Or, ce n'est pas parce que nous accédons à des milliards de données représentant des terabytes d'informations numériques que nous disposons de données plus représentatives. Aussi grandes que puissent être ces données, *n* n'est jamais la globalité.

Les biais inhérents à la constitution des traces numériques sont effectivement tout à fait remarquables. L'exemple de Twitter est incontestablement le plus édifiant. Il constitue en effet l'une des sources les plus utilisées, bien qu'il ne soit que très marginalement utilisé par la population. Les données relatives à la géolocalisation des téléphones cellulaires sont certes plus représentatives statistiquement, mais elles n'en demeurent pas moins d'une grande pauvreté sémantique.

Les analyses exploitant de telles sources n'auraient de sens que si elles étaient circonscrites aux réalités considérées, ce qui est rarement le cas¹⁷. Une approche plus rigoureuse exigerait d'exclure toutes problématiques que ces traces ne représenteraient que partiellement, ce qui en réduirait considérablement l'intérêt et la portée, au-delà d'une meilleure connaissance de la mobilité des personnes considérées.

Surtout, l'idée même de représentativité statistique est discutable. En insistant trop souvent sur la délicate représentativité d'une population par un échantillon, on ne s'interroge pas assez sur ce que représentent les données. Avant de représenter une population, les données représentent-elles bien les individus qui la constituent ? Sur Internet, et particulièrement sur Twitter, Flickr, Facebook ou Foursquare, la *mise en scène* est tellement travaillée, que les *coulisses* peuvent constituer l'essentiel de pratiques pourtant invisibles [Goffman, 1973]. Ainsi, une différence observée selon les genres témoignerait-elle d'une mise en scène différente ou de pratiques territoriales différentes [Cousin et al., 2014; Beecham & Wood, 2013] ? Les niveaux de privacité variant considérablement selon les personnes, leur âge ou leur genre [Li & Chen, 2010], quel sens donner aux spatialités émergeant de telles traces ?

Contextes hors contextes

C'est pourquoi le contexte de production de ces données est particulièrement important, bien qu'il soit la plupart du temps éludé lors de leur exploitation. L'agrégation de données produites dans des conditions très hétérogènes se traduit par une augmentation de leur quantité qui correspond conjointement à une réduction de leur qualification et par continuité de leur

¹⁷ L'exemple de l'application *Street Bump* (<http://www.streetbump.org/>) qui permet d'identifier les dégradations de la chaussée à Boston témoigne parfaitement de cela. Les problèmes relatés étaient essentiellement circonscrits aux quartiers rassemblant des jeunes relativement aisés, qui se révèlent être aussi les usagers de cette application [Harford, 2014].

qualité. La dissolution du contexte se confond ainsi avec celle du sens, au risque d'un réductionnisme particulièrement intense, qui ne se limite pas à la non représentation de l'ensemble des réalités sociales, mais aussi de ce qui les constitue singulièrement et subjectivement [Rouvroy & Berns, 2013]. La différence entre les plateformes Flickr et Geograph illustre parfaitement l'importance du contexte, l'une et l'autre présentant deux distributions spatiale et sociale tout à fait distinctes, qui correspondent à deux pratiques de publication très différentes [Purves & Edwardes, 2011].

La valorisation des traces numériques accessibles comme source innovante encourage de surcroît l'orientation des recherches quelles que soient leurs limites. Leur quantité, trop souvent, est effectivement valorisée indépendamment de leur qualité et de leur pertinence. Cela pose des problèmes évidents en urbanisme, par exemple, alors que les temporalités considérées ne correspondent généralement pas à celles des traces numériques disponibles [Batty, 2013]. Cette réduction du social à ce qui est quantifiable est particulièrement sensible lorsqu'elle s'applique à des problématiques spatiales complexes telles que la criminalité [Beaude, 2009]. Si nous dépendions d'une telle démarche, nous serions tels des ivrognes qui cherchent inlassablement leur clé à la lumière d'un lampadaire. Cette tentation quantitative est d'autant plus sensible qu'elle traverse l'ensemble des sociétés occidentales, alors que les chiffres représentent de moins en moins la réalité sociale et la façonnent de plus en plus¹⁸ [Desrosières, 2008].

New Deal on Data

Aussi, la quantité qui émerge de l'accumulation des traces résulte ultimement de l'association de spatialités singulières de millions d'individus, dont l'exposition et le traitement sont rarement consentis. C'est pourquoi il conviendrait de considérer plus systématiquement les problèmes éthiques relatifs à la traçabilité généralisée des pratiques individuelles. Car tels enjeux éthiques correspondants ne peuvent certainement pas être évités au nom d'un empirisme érigé en absolu.

Les *big data* et la *social physics* suppose en effet une assignation violente à la transparence, parfaitement incarnée par Google et Facebook, qui exigent l'accès au plus grand nombre de traces possible, sans lesquelles l'individualisation contextuelle ne serait pas possible. Or, le traitement automatisé de l'ensemble des données publiques d'un individu n'est pas nécessairement justifiable lorsqu'elles ont été élaborées dans un contexte de visibilité spécifique, soulignant la différence fondamentale entre être en public et être public [boyd & Marwick, 2011].

L'accès relativement opaque à des données privées sous couvert d'anonymisation, alors même que la robustesse de ces méthodes est largement critiquée, mériterait aussi d'être considéré avec beaucoup d'attention [Blumberg & Eckersley, 2009; de Montjoye et al., 2013;

¹⁸ La recherche ne fait d'ailleurs pas exception à cela. Alors qu'elle est de plus en plus réduite à ses traces, les citations, les chercheurs devraient être particulièrement sensibles de la réduction dont de telles approches sont porteuses.

de Montjoye et al., 2015]. Ce problème est particulièrement délicat, si l'on considère conjointement la multiplication des failles informatiques, l'augmentation des vols massifs de données, la généralisation de l'interconnexion de services différents et la fusion courante des bases de données lors d'acquisition. La concentration incontrôlée des traces expose ainsi à des atteintes inédites à la vie privée, hors de tout consentement et de toute négociation. Cela suppose une re-négociation « collective, conflictuelle et itérative » de la *privacy* [Casilli, 2013], et une réappropriation individuelle qui consisterait à *faire trace*, afin de limiter les effets de décontextualisation [Merzeau, 2013].

Cette stratégie inéluctable de contrôle engagé par les individus se traduit par une mise en scène croissante des traces numériques, qui limite plus encore les présupposés de représentativité, tout en soulignant la consubstantialité des données et de leur contexte de production.

La proposition d'un *New Deal on Data* [Pentland, 2014 :177-188] qui permettrait un contrôle beaucoup plus fin des données personnelles n'est en cela pas tout à fait satisfaisante. Elle répond à la demande légitime des individus de choisir plus précisément ce qu'ils souhaitent partager, mais elle accroît conjointement la disjonction entre ce qui est montré et ce qui ne l'est pas, mais aussi entre ceux qui montrent et ceux qui ne montrent pas.

Un tel changement exigerait aussi de repenser le modèle économique d'Internet, dont l'activité des principaux acteurs (Facebook, Google, Twitter, Netflix, Spotify...) repose très largement sur l'analyse des informations privées des utilisateurs. Un contrôle accru des données engage à se demander ce que serait une *data-driven society* qui reposerait sur les informations sélectives que les individus souhaiteraient partager, tout en éludant les autres, aussi importantes soient elles ? Quels modèles peuvent émerger d'une telle mise en scène de la vie des individus, en plus d'en être une réduction ?

Le pouvoir de l'opacité

L'opacité croissante des processus de collecte, de concentration et de traitement des traces numériques suscite un dernier enjeu particulièrement sensible. Alors que les traces numériques constituent une opportunité stimulante de renouveler la science, mais aussi la politique, les modalités contemporaines de leur production sont elles aussi parfaitement contradictoires avec les exigences de la science, mais aussi de la politique. Au-delà de la mise en scène des traces par les individus, les coulisses ne se résument pas aux pratiques invisibles en l'absence de traces. Elles recouvrent aussi l'ensemble des pratiques de collecte et de traitement des données, dont le contrôle et la visibilité sont particulièrement opaques.

Or, la science repose sur l'organisation de la réfutation des énoncés scientifiques, ce qui exige de pouvoir accéder aux données utilisées lors d'une expérience ou de pouvoir la reproduire. Pourtant, de plus en plus de recherches portent sur des traces numériques qui s'appuient sur des données privées, qui ne peuvent pas être communiquées à l'ensemble des chercheurs (coordonnées des téléphones cellulaires, usage des cartes bancaires, publications privées sur Facebook...). Cette « capitalisation de la connaissance » [Thrift, 2005] pose des problèmes

majeurs à l'exercice de la pratique scientifique. En plus de la sous-exploitation des données inhérente à cette situation, les résultats de telles recherches, malgré leur importance, ne peuvent pas être validés convenablement.

De surcroît, cette opacité concerne non seulement les données, mais aussi les algorithmes utilisés pour les exploiter, déployant subrepticement une double opacité particulièrement préjudiciable à l'exercice de la science. Cette double opacité est d'autant plus regrettable qu'elle se transpose avec la même intensité à la politique, lorsque les traces numériques sont exploitées pour informer des stratégies relatives à l'organisation des transports, de la lutte contre la criminalité et plus généralement des villes dans leur ensemble.

Avec Derrida, danah boyd et Kate Crawford nous rappellent que la démocratisation tient essentiellement en un principe : la participation et l'accès aux archives, à leur constitution et à leur interprétation [boyd & Crawford, 2012 :674]. Sans cela, la science, comme la politique, devient une boîte noire dont nous dépendrions sans en comprendre le processus [Pasquale, 2015], éliminant la possibilité même de critique et plus encore de projet. En faisant de l'individu une boîte noire, nous faisons inéluctablement de la société une boîte noire. Certains s'en réjouissent, y trouvant l'occasion d'en finir avec l'artifice des sociétés [Latour, 2006], d'autres y trouvent l'occasion de s'affranchir de l'inefficience supposée d'institutions politiques devenues obsolètes [Turner, 2006]. Néanmoins, que serait une politique sans société et sans institutions chargées de représenter les individus, non pas pour en réduire l'existence, mais pour transformer les présentes aspirations individuelles en coexistence à venir ?

L'effacement du sujet

Au-delà de l'opacité des processus, l'enjeu que recouvre le paradigme commun aux *big data* et à la *social physics* est paradoxalement celui de l'effacement du sujet, du glissement subtil de l'individu à l'individu statistique et, plus subtilement, de la dissolution du sujet dans le réseau [Rouvroy & Berns, 2013]. Alex Pentland, comme Bruno Latour, souligne l'importance du réseau, des relations et des flux pour comprendre la dynamique qui organise la coexistence. Ils désagrègent la société, trop fictionnelle, pour se concentrer sur le social, c'est-à-dire l'ensemble des relations qui constituent des associations plus ou moins stables. Les analogies organicistes et éthologiques sont ainsi récurrentes. Les références aux fourmis, mais aussi aux abeilles, sont d'ailleurs régulièrement mobilisées pour mieux souligner la puissance et la généralité du propos. L'individu n'est plus qu'une fiction, la distinction entre les humains et les non-humains un artefact, et les catégories, comme les structures, des constructions pré-scientifiques, qui appauvriraient le réel en réduisant les individus à des groupes fictifs, dont l'archétype serait la classe sociale.

On peut se demander en quoi la réduction d'un individu à ses traces numériques, aussi réticulaires soient-elles, serait moins préjudiciable que de le réduire à son métier, son âge, son sexe, son lieu de résidence et la multitude d'autres réductions du monde social ? L'effacement du sujet pose en cela une question politique majeure. Qu'est-ce qu'une société composée d'individus dont on ne considère pas spécifiquement les existences autrement que par des flux

et des relations qui les relient sans cesse à des collectifs innombrables ? Que signifie la coexistence de réalités pourvues d'intentions (projections dans un avenir collectif) et de réflexivité (construction des intentions), si la capacité des êtres humains à ne pas se conformer aux attentes que nous en avons et à transformer leurs intentions selon ces mêmes attentes était comparable à celles dont les coquilles Saint-Jacques seraient pourvues [Callon, 1986] ? En quoi dire que les humains sont comparables à des atomes, à des fourmis [Latour, 2010 :155,160] ou à des abeilles [Pentland, 2014 :97-98,190,209] permettrait de mieux comprendre le monde social, au-delà de sa seule dimension biophysique ?

Faudrait-il en conclure que l'humain n'est pas suffisamment singulier pour mériter des édifices théoriques plus sophistiqués que le modèle de gravité, la thermodynamique, la stigmergie ou la cybernétique ? Devrions-nous accepter l'avènement d'une gouvernamentalité algorithmique qui « évite les sujets humains réflexifs [et] se nourrit de données infra-individuelles insignifiantes en elles-mêmes, pour façonner des modèles de comportements ou profils supra-individuels sans jamais en appeler au sujet, sans jamais l'appeler à rendre compte par lui-même de ce qu'il est ni de ce qu'il pourrait devenir » [Rouvroy & Berns, 2013 :173-174] ? Dans un monde d'acteurs-réseaux qui serait régi par la *social physics*, il n'y aurait pas de place pour des sujets. Dans une *data-driven society*, ce seraient les données qui assureraient la conduite de nos actes.

La politique suppose précisément un monde qui ne se résume pas à son ordre actuel, et moins encore à son ordre passé. Alex Pentland parle d'ailleurs de *breadcrumbs*, et non de *footprints*. Des miettes de pain, tel le fil d'Ariane du Petit poucet, des marques plus que des traces, qui permettent certes de connaître le chemin parcouru, mais avec l'intention unique de le parcourir dans l'autre sens. Voici la limite des traces d'un chemin parcouru. Nous voyons comment cela permet de connaître le passé, de revenir éventuellement sur nos pas et, au mieux, de reproduire l'existant. Il est plus difficile de percevoir en quoi cela nous indique le chemin à venir. La politique exige certes de mieux connaître les réalités qui l'organisent, mais elle exige aussi de rendre commensurables les traces que nous laissons de nos actes passés et les intentions de nos actes à venir.

Changements dans les mouvements

L'approche des *big data* promue par Chris Anderson, Victor Mayer-Schonberger ou Alex Pentland semble non seulement confondre le mouvement et le changement, mais aussi le potentiel de la description partielle de l'ordre existant avec l'actualisation de l'ordre à venir. En s'appuyant sur des corrélations, aussi nombreuses soient-elles, de telles approches permettent de renouveler puissamment la lisibilité du présent, mais elles n'en demeurent pas moins insuffisantes pour déduire de cet ordre celui qui vient. Cette projection n'a de sens que dans un monde stable, dépourvu de réflexivité et expurgé de tout changement [Harford, 2014; Carr, 2014]. Il s'agit d'un monde sans réelle innovation. C'est pourquoi les références de Bruno Latour à Gabriel Tarde pour justifier son engouement pour les traces numériques est surprenant. Car en 1890, aux prémices de la sociologie, Gabriel Tarde soulignait déjà les limites de la statistique dans *Les lois de l'imitation* :

« À mesure que la statistique porte sur de plus grands nombres, on est quelquefois enclin à penser que, bien plus tard, si la marée montante de la population continue à accroître et les grands États à grandir, un moment viendra où tout, dans les phénomènes sociaux, sera réductible en formules mathématiques. D'où l'on induit abusivement que le statisticien pourra un jour prédire l'état social futur aussi sûrement que l'astronome la prochaine éclipse de Vénus. En sorte que la statistique serait destinée à plonger toujours plus avant dans l'avenir comme l'archéologie dans le passé. Mais nous savons par tout ce qui précède que la statistique est circonscrite dans le champ de l'imitation et que celui de l'invention lui est interdit » [Tarde, 2001 :196].

Cette opposition entre l'*imitation* et l'*invention* est au cœur de la pensée de Gabriel Tarde, qui considère que l'une et l'autre constituent les actes sociaux élémentaires, dont la dynamique serait animée par la *croissance* et le *désir* [Tarde, 2001 :203]. Gabriel Tarde émet alors l'hypothèse selon laquelle les « similitudes sociales » s'expliquent par l'importance de l'imitation, conçue comme « force sociale » dominante.

Alex Pentland procède selon le même raisonnement lorsqu'il construit la physique sociale autour de deux logiques élémentaires : les logiques d'engagement, qui relèvent du proche et du connu, et les logiques d'exploration, qui relèvent du lointain et de l'inconnu. Pour expliquer la régularité de l'ordre social, Alexis Pentland souligne que l'engagement constituerait l'essentiel des pratiques sociales, ce qui expliquerait la prévisibilité considérable de nos actes¹⁹ [Pentland, 2014 :239]. En revanche, bien qu'il conçoive que la *social physics* ne soit pas appropriée à l'appréhension des actes innovants, Alex Pentland semble considérer que la maîtrise des régularités suffit à organiser une société.

Ainsi, l'échec de Google Flu Trends, qui repose précisément sur cette incapacité à vraiment saisir le changement dans le mouvement, est sensiblement le même que celui des marchés financiers, dont les analyses s'appuyant sur le passé fonctionnent tant qu'il n'y a pas de changement. Leur conformité au passé laisse supposer leur conformité à venir, dans l'illusion la plus totale, aussi entretenue et dévastatrice soit elle. C'est pourquoi la réflexivité et plus précisément la capacité des individus à transformer significativement leur comportement ne doivent pas être négligées. Sans cette précaution, l'usage systématique des traces numériques se résumera à la description d'invariants relativement banaux, mais aussi à maintenir l'ordre social existant [Carr, 2014], en négligeant la composante la plus essentielle de l'humanité : l'invention. Comme le soulignait déjà Gabriel Tarde, l'imitation est certes la force sociale dominante, mais l'innovation, à la force des désirs, est celle qui motive le changement.

Dès 1999, Lawrence Lessig a souligné l'importance du code et plus précisément des algorithmes dans la gouvernance des pratiques [Lessig, 1999]. Aussi, Martin Dodge et Rob Kitchin ont parfaitement décrit en quoi l'intrication du code et de l'espace permet de repenser

¹⁹ Alex Pentland justifie la régularité des logiques d'engagement en les associant à la *pensée rapide*, telle que proposée par Daniel Kahneman [Kahneman, 2012]. La *pensée rapide*, constituée par les habitudes, procéderait par associations automatiques et inconscientes. Elle s'opposerait à la *pensée lente* qui procéderait quant à elle par raisonnements conscients, qui permettraient l'innovation [Pentland, 2014 :235].

la perspective ontogénétique du changement proposé par Gilbert Simondon [Dodge & Kitchin, 2005; Kitchin & Dodge, 2011]. Les traces numériques s'inscrivent effectivement dans une spatialité algorithmique complexe, qui relève d'une *transduction spatiale* particulièrement puissante. Cette *transduction* se caractérise par un processus continu de renouvellement de l'espace par l'actualisation de spatialités à venir qui mobilisent les traces de spatialités passées.

La spatialité algorithmique suppose ainsi de considérer précisément les codes qui rendent possibles la production et la circulation des représentations de l'espace. De la codification des traces aux traitements qui les associent pour leur donner des sens particuliers, il est urgent de rendre visibles les processus de production de la visibilité. Car les codes sont porteurs de valeurs [Lessig, 1999; Cardon, 2013], et ceux qui en ont la maîtrise, et plus encore le contrôle, ont le pouvoir d'organiser les spatialités algorithmiques contemporaines.

L'avenir à inventer

Il est « facile de prévoir le passé » ironisait Steven Salzberg en 2014 en réaction à l'échec récurrent de Google Flu Trends, rappelant l'abondance des modèles qui reposent sur des analyses statistiques qui, sans la moindre dimension explicative, sans hypothèses véritables, parviennent toujours à se conformer au passé [Salzberg, 2014]. Pourtant, l'enjeu de la politique n'est-il pas précisément de définir l'avenir qui convient plutôt que de le prévoir ? La prédiction algorithmique de l'avenir sur la base des actes passés ne serait-elle pas la négation même de la politique, à savoir le choix d'un avenir parmi des options non déterminées ? La politique n'est-elle pas essentiellement l'organisation de la coexistence, au-delà de son simple constat ? La politique, dans les démocraties modernes, ne consiste-t-elle pas à organiser, ensemble, le monde que l'on souhaite ?

Les logiques de marché elles-mêmes répondent bien à cela, alors qu'Apple est devenu la plus importante capitalisation mondiale en réfutant la pertinence des études de marché, qui projettent l'avenir à partir du présent, sans projet véritable²⁰. Steve Jobs aimait rappeler la célèbre phrase d'Henri Ford : « Si j'avais demandé à mes clients ce qu'ils voulaient, ils auraient répondu “un cheval plus rapide”, et pas une voiture ». C'est bien là la limite des algorithmes qui reposent sur les traces numériques : être incapable de prévoir l'avenir que nous leur confions, car dans le meilleur des cas, ils ne connaissent que notre passé recomposé ! Le futur est plein de propositions inattendues, qui répondent à une demande qui n'existe pas encore.

Les traces constituent en cela une opportunité de mieux connaître le monde, et les algorithmes un moyen de mieux en saisir la richesse, à la seule condition d'en comprendre la portée et les limites. Nous devons sans aucun doute saisir cette opportunité de renouveler les conditions pratiques de l'empirie et de multiplier les oligoptiques. Il convient à présent de mobiliser

²⁰ Cela ne signifie aucunement qu'Apple ne sorte pas de produits qui ne rencontrent pas de marché, mais que ces échecs ne sont pas significativement plus importants que chez d'autres constructeurs qui organisent leur avenir autour d'une idée présente du futur.

activement le numérique en sciences sociales [Rogers, 2013], d'en saisir la pluralité des approches [Plantin & Monnoyer-Smith, 2014], d'associer les traces numériques aux données conventionnelles [Lazer et al., 2014] et d'embrasser pleinement cette complexité plutôt que d'exclure le potentiel dont les traces numériques sont porteuses [Ruppert, 2013; Gorman, 2013; Severo & Romele, 2014].

Une telle démarche a comme préalable de rappeler que ces traces n'ont pas vocation à dicter notre conduite. Elles ne dispensent certainement pas d'interprétations, d'analyses, et moins encore de projets ! Encore faut-il le rappeler, les spatialités algorithmiques à venir sont largement à inventer et non à attendre !

Bibliographie

- [Ahas et al., 2010] Ahas, Rein et al. 2010. «Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: Case study with mobile positioning data». *Transportation Research Part C: Emerging Technologies*, 18, 1, pp 45–54.
- [Anderson, 2008] Anderson, Chris. 2008. «The end of theory». *Wired magazine*, 23 juin 2008.
- [Barnes, 2013] Barnes, Trevor J. 2013. «Big data, little history». *Dialogues in Human Geography*, 3, 3, pp 297–302.
- [Batty, 2013] Batty, Michael. 2013. «Big data, smart cities and city planning». *Dialogues in Human Geography*, 3, 3, pp 274–279.
- [Bayir et al., 2009] Bayir, M.A.; Demirbas, M. & Eagle, N. 2009. «Discovering spatiotemporal mobility profiles of cellphone users», *World of Wireless, Mobile and Multimedia Networks & Workshops*, IEEE, pp 1-9.
- [Beaude, 2009] Beaude, Boris. 2009. «*Crime Mapping*, ou le réductionnisme bien intentionné». *EspacesTemps.net*, mis en ligne le 30 juillet 2013, consulté le 19 mars 2015. URL : <http://www.espacestemp.net/document7733.html>.
- [Beaude, 2010] Beaude, Boris. 2010. «Espace de la carte, espace de la ville. Des analogies à la coexistence» in Zreik, Khaldoun (dir.). *Nouvelles cartographies, nouvelles villes. HyperUrbain2*, Paris: Europa Productions, pp 15–39.
- [Beaude, 2012] Beaude, Boris. 2012. *Internet. Changer l'espace, changer la société*. Limoges: FYP.
- [Beaude, 2013] Beaude, Boris. 2013. «Synchronisation» in Lussault, Michel & Lévy, Jacques (dir.). *Dictionnaire de la géographie et de l'espace des sociétés*, Paris: Belin, pp 972.
- [Beaude, 2014] Beaude, Boris. 2014. *Les fins d'Internet*. Limoges: Fyp.
- [Beecham & Wood, 2013] Beecham, Roger & Wood, Jo. 2013. «Exploring gendered cycling behaviours within a large-scale behavioural data-set». *Transportation Planning and Technology*, 37, 1, pp 83–97.
- [Benkler, 2006] Benkler, Yochai. 2006. *The Wealth of Networks: How Social Production Transforms Markets And Freedom*. London: Yale University Press.
- [Blondel et al., 2010] Blondel, Vincent; Krings, Gautier & Thomas, Isabelle. 2010. *Regions and borders of mobile telephony in Belgium and in the Brussels metropolitan zone. Brussels Studies*.
- [Borgnat et al., 2009] Borgnat, Pierre et al. 2009. «Spatial analysis of dynamic movements of Vélo'v, Lyon's shared bicycle program», *European Conference on Complex Systems 2009 - ECCS'09*, Warwick : United Kingdom.

- [boyd & Crawford, 2012] boyd, danah & Crawford, Kate. 2012. «Critical questions for big data». *Information, Communication & Society*, 15, 5, pp 662–679.
- [boyd & Crawford, 2011] boyd, danah & Crawford, Kate. 2011. «Six Provocations for Big Data». *SSRN*.
- [boyd & Marwick, 2011] boyd, danah & Marwick, Alice E. 2011. «Social Privacy in Networked Publics: Teens' Attitudes, Practices, and Strategies».
- [Butler, 2013] Butler, Declan. 2013. «When Google got flu wrong.». *Nature*.
- [Calabrese et al., 2010] Calabrese, Francesco et al. 2010. «The Geography of Taste: Analyzing Cell-Phone Mobility and Social Events» in Floréen, Patrik et al. (dir.). *The Geography of Taste: Analyzing Cell-Phone Mobility and Social Events*, Lecture Notes in Computer Science. Berlin: Springer, 6030, pp 22–37.
- [Callon, 1986] Callon, Michel. 1986. «Éléments pour une sociologie de la traduction». *L'Année sociologique*, 36, pp 169–208.
- [Cardon, 2013] Cardon, Dominique. 2013. «Dans l'esprit du PageRank». *Réseaux*, 177, 1, pp 63–33.
- [Carmes & Noyer, 2014] Carmes, Maryse & Noyer, Jean-Max. 2014. «L'«irrésistible montée de l'algorithmique». *Les Cahiers du numérique*, 10, 4, pp 63–102.
- [Carr, 2014] Carr, Nicholas. 2014. «The Limits of Social Engineering». *MIT Technology Review*.
- [Casilli, 2013] Casilli, Antonio A. 2013. «Contre l'hypothèse de la « fin de la vie privée ». La négociation de la privacy dans les médias sociaux». *Revue française des sciences de l'information et de la communication*, 3, mis en ligne le 30 juillet 2013, consulté le 19 mars 2015. URL : <http://rfsic.revues.org/630>.
- [Castells, 2001] Castells, Manuel. 2001. *The Internet galaxy: reflections on the Internet, business, and society*. Oxford: Oxford University Press.
- [Comte, 1995 (1842)] Comte, Auguste. 1995 (1842). *Discours sur l'esprit positif*. Paris: Librairie philosophique J. Vrin.
- [Cortright, 2009] Cortright, Joe. 2009. «Walking the Walk: How Walkability Raises Home Values in U.S. Cities», CEOs for Cities.
- [Cousin et al., 2014] Cousin, Saskia et al. 2014. «Étudier TripAdvisor. Ou comment Trip-patouiller les cartes de nos vacances.». *EspacesTemps.net*, mis en ligne le 29 août 2014, consulté le 19 mars 2015. URL : <http://www.espacestemp.net/articles/etudier-tripadvisor-ou-comment-trip-patouiller-les-cartes-de-nos-vacances/>.
- [Crandall et al., 2009] Crandall, David J et al. 2009. «Mapping the world's photos». *WWW '09 : Proceedings of the 18th international conference on World wide web*, pp 761–770.
- [Cranshaw et al., 2012] Cranshaw, Justin et al. 2012. «The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City.». *ICWSM*.
- [de Montjoye et al., 2015] de Montjoye, Yves-Alexandre et al. 2015. «Unique in the shopping mall: On the reidentifiability of credit card metadata». *Science*, 347, 6221, pp 536–539.
- [de Montjoye et al., 2013] de Montjoye, Yves-Alexandre et al. 2013. «Unique in the Crowd: The privacy bounds of human mobility». *Scientific Reports*.
- [Deleuze, 1968] Deleuze, Gilles. 1968. *Différence et répétition*. Paris: Presses Universitaires de France - PUF.
- [Deleuze, 1990] Deleuze, Gilles. 1990. *Pourparlers*. Paris: Les Éditions de Minuit.
- [Desrosières, 2008] Desrosières, Alain. 2008. *Gouverner par les nombres*. Paris: Presses de l'École des mines.
- [Deville, 2014] Deville, Pierre et al. 2014. «Dynamic population mapping using mobile phone data.». *Proceedings of the National Academy of Sciences of the United States of America*, 111, 45, pp 15888–15893.

- [Dodge & Kitchin, 2005] Dodge, Martin & Kitchin, Rob. 2005. «Code and the Transduction of Space». *Annals of the Association of American Geographers*, 95, 1, pp 162–180.
- [Eagle & Pentland, 2006] Eagle, Nathan & Pentland, Alex. 2006. «Reality mining: sensing complex social systems». *Personal and Ubiquitous Computing*, 10, 4, pp 255–268.
- [Eagle & Pentland, 2009] Eagle, Nathan & Pentland, Alex. 2009. «Eigenbehaviors: identifying structure in routine». *Behavioral Ecology and Sociobiology*, 63, 7, pp 1057–1066.
- [Blumberg & Eckersley, 2009] Blumberg, Andrew J & Eckersley, Peter. 2009. *On Locational Privacy, and How to Avoid Losing it Forever*, mis en ligne le 3 août 2009, consulté le 19 mars 2015. URL : <https://www.eff.org/fr/wp/locational-privacy>.
- [Foucault, 1975] Foucault, Michel, 1975, *Surveiller et punir*, Paris: Gallimard.
- [Froehlich et al., 2006] Froehlich, Jon et al. 2006. «Voting with Your Feet : An Investigative Study of the Relationship Between Place Visit Behavior and Preference», Lecture Notes in Computer Science, UbiComp'06, 4206, pp 333–350.
- [Froehlich, 2008] Froehlich, Jon; Neumann, Joachim & Oliver, Nuria. 2008. «Measuring the Pulse of the City through Shared Bicycle Programs», Proceedings of UrbanSense08 - International Workshop on Urban, Community, and Social Applications of Networked Sensing Systems, Raleigh, North Carolina, USA, pp 16–20.
- [Ganascia, 2009] Ganascia, Jean-Gabriel. 2009. *Voir et pouvoir : qui nous surveille ?*, Paris: Édition le Pommier.
- [Ginsberg, 2009] Ginsberg, Jeremy et al. 2009. «Detecting influenza epidemics using search engine query data». *Nature*, 457, 7232, pp 1012–1014.
- [Girardin, 2009] Girardin, Fabien. 2009. «Aspects of implicit and explicit human interactions with ubiquitous geographic information». *Phd Thesis*, Barcelona: Universitat Pompeu Fabra. Departament de Tecnologies de la Informació i les Comunicacions.
- [Girardin, 2008] Girardin, Fabien; Calabrese, Francesco, Fiore, Filippo Dal et al. 2008. «Digital Footprinting: Uncovering Tourists with User-Generated Content». *IEEE Pervasive Computing*, 7, 4, pp 36–43.
- [Goffman, 1973] Goffman, Erving. 1973. *La mise en scène de la vie quotidienne*. Paris: Les Éditions de Minuit.
- [Gonzalez-Bailon, 2013] Gonzalez-Bailon, Sandra. 2013. «Big data and the fabric of human geography». *Dialogues in Human Geography*, 3, 3, pp 292–296.
- [González et al., 2008] González, Marta C; Hidalgo, César A & Barabási, Albert-László. 2008. «Understanding individual human mobility patterns». *Nature*, 453, 7196, pp 779–782.
- [Gordon et al., 2011] Gordon, Eric & Silva, Adriana de Souza e. 2011. *Net Locality*. Hoboken, New Jersey: John Wiley & Sons.
- [Gorman, 2013] Gorman, Sean P. 2013. «The danger of a big data episteme and the need to evolve geographic information systems». *Dialogues in Human Geography*, 3, 3, pp 285–291.
- [Graham & Shelton, 2013] Graham, Mark & Shelton, Taylor. 2013. «Geography and the future of big data, big data and the future of geography». *Dialogues in Human Geography*, 3, 3, pp 255–261.
- [Graham, 2005] Graham, Stephen. 2005. «Software-sorted geographies». *Progress in Human Geography*, 29, 5, pp 562–580.
- [Harford, 2014] Harford, Tim. 2014. «Big data: are we making a big mistake?». *Financial Times*, mis en ligne le 28 mars 2014, consulté le 19 mars 2015. URL : <http://www.ft.com/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html>.

- [Hollenstein & Purves, 2011] Hollenstein, Livia & Purves, Ross. 2011. «Exploring place through user-generated content: Using Flickr tags to describe city cores». *Journal of Spatial Information Science*, 0, 1, pp 21–48.
- [Ioannidis, 2005] Ioannidis, John P A. 2005. «Why Most Published Research Findings Are False». *PLoS Medicine*, 2, 8, e124.
- [Kahneman, 2012] Kahneman, Daniel. 2012. *Thinking, Fast and Slow*. Penguin.
- [Kitchin, 2013] Kitchin, Rob. 2013. «Big data and human geography: Opportunities, challenges and risks». *Dialogues in Human Geography*, 3, 3, pp 262–267.
- [Kitchin & Dodge, 2011] Kitchin, Rob & Dodge, Martin. 2011. *Code/Space: Software and Everyday Life*. Cambridge, Massachusetts: The MIT Press.
- [Klauser, 2010] Klauser, Francisco, 2010. «Splintering spheres of security: Peter Sloterdijk and the contemporary fortress city». *Environment and Planning D: Society and Space*, 28, 2, pp 326–340.
- [Krings et al., 2009] Krings, Gautier et al. 2009. «Urban gravity: a model for inter-city telecommunication flows». *Journal of Statistical Mechanics: Theory and Experiment* 2009, 07, L07003–9.
- [Lambiotte et al., 2008] Lambiotte, Renaud et al. 2008. «Geographical dispersal of mobile communication networks». *Physica A: Statistical Mechanics and its Applications* 387 [21], pp 5317–5325.
- [Latour, 1999] Latour, Bruno. 1999. *Pandora's Hope*. Harvard University Press.
- [Latour, 2006] Latour, Bruno. 2006. *Changer de société, refaire de la sociologie*. Paris: La Découverte.
- [Latour, 2007] Latour, Bruno. 2007. «Paris, ville invisible : le plasma» in Macel, Christine et al. (dir.). *Airs de Paris, 30 ans du Centre Pompidou*, Paris:ADGP, pp 260–263.
- [Latour, 2010] Latour, Bruno. 2010. *Cogitamus : Six lettres sur les humanités scientifiques*. Paris: La Découverte.
- [Latour & Hermant, 1998] Latour, Bruno & Hermant, Emilie. 1998. *Paris ville invisible*. Paris: La Découverte.
- [Latour et al., 2012] Latour, Bruno et al. 2012. «'The whole is always smaller than its parts«—a digital test of Gabriel Tarde's monads». *The British journal of Sociology*, 63,4, pp 590–615.
- [Lazer et al., 2009] Lazer, David et al. 2009. «Computational Social Science». *Science*, 323, 5915, pp 721–723.
- [Lazer et al., 2014] Lazer, David et al. 2014. «Big data. The parable of Google Flu: traps in big data analysis.». *Science*, 343, 6176, pp 1203–1205.
- [Lessig, 1999] Lessig, Lawrence. 1999. *Code and other laws of cyberspace*. New York: Basic Books.
- [Lévy, 1994] Lévy, Jacques. 1994. *L'espace légitime*. Paris: Presses de la Fondation nationale des sciences politiques.
- [Lévy, 1999] Lévy, Jacques. 1999. *Le tournant géographique*. Paris: Belin.
- [Li & Chen, 2010] Li, Nan & Chen, Guanling. 2010. «Sharing location in online social networks». *Network, IEEE*, 24, 5, pp 20–25.
- [Lussault, 2003] Lussault, Michel. 2003. «Spatialité» in Lussault, Michel & Lévy, Jacques (dir.). *Dictionnaire de la géographie et de l'espace des sociétés*, Paris: Belin, pp 866–868.
- [Lussault, 2009] Lussault, Michel. 2009. *De la lutte des classes à la lutte des places*. Paris: Grasset.
- [Lussault, 2013] Lussault, Michel. 2013. «Hyperspatialité» in Lussault, Michel & Lévy, Jacques (dir.). *Dictionnaire de la géographie et de l'espace des sociétés*, Paris: Belin, 521–522.

- [Lussault & Stock, 2010] Lussault, Michel & Stock, Mathis. 2010. «Doing with space: towards a pragmatics of space». *Social Geography*, 5, 1, pp 11–19.
- [Mayer-Schonberger & Cukier, 2013] Mayer-Schonberger, Victor & Cukier, Kenneth. 2013. *Big Data : A Revolution That Will Transform How We Live Work and Think*. Boston: Houghton Mifflin Harcourt.
- [McLuhan, 1969] McLuhan, Marshall. 1969. «An Interview with Marshall McLuhan». *Playboy*, mars 1969.
- [Merzeau, 2013] Merzeau, Louise. 2013. «L'intelligence des traces». *Intellectica*, 1, 59, p.115–135.
- [Mitchell, 1996] Mitchell, William J. 1996. *City of bits: space, place, and the infobahn*. Cambridge, Massachusetts: The MIT Press.
- [Mooij et al., 2014] Mooij, Joris M et al. 2014. «Distinguishing cause from effect using observational data: methods and benchmarks».
- [Musso, 1997] Musso, Pierre. 1997. *Télécommunications et philosophie des réseaux: la postérité paradoxale de Saint-Simon. La politique éclatée*. Paris: Presses universitaires de France.
- [Neuhaus, 2010] Neuhaus, Fabian. 2010. «UrbanDiary - A Tracking Project Capturing the beat and rhythm of the city: Using GPS devices to visualise individual and collective routines within Central London.». *The Journal of Space Syntax*, 1, 2, pp 315-336.
- [Nova, 2009] Nova, Nicolas. 2009. *Les médias géolocalisés: comprendre les nouveaux espaces numériques*. Limoges: FYP.
- [Olteanu et al., 2011] Olteanu, Ana-Maria; Couronné, Thomas & Fen-Chong, Julie. 2011. SAGEO '11 - International Conference on Spatial Analysis and GEOmatics - Paris, «Modélisation des trajectoires spatio-temporelles issues des traces numériques de téléphones mobiles».
- [O'Neill, 2006] O'Neill, Eamonn et al. 2006. «Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape» in Dourish, Paul & Friday, Adrian (dir.), *Ubicomp 2006, LNCS, 4206*, pp 315–332.
- [Pasquale, 2015] Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
- [Pentland, 2014] Pentland, Alex. 2014. *Social Physics: How Good Ideas Spread-The Lessons from a New Science*. New York: The Penguin Press.
- [Phithakkitnukoon & Olivier, 2011] Phithakkitnukoon, Santi & Olivier, Patrick. 2011. «Sensing Urban Social Geography Using Online Social Networking Data», *The Social Mobile Web*.
- [Plantin & Monnoyer-Smith, 2014] Plantin, Jean-Christophe & Monnoyer-Smith, Laurence. 2014. «Ouvrir la boîte à outils de la recherche numérique. Trois cas de redistribution de méthodes». *tic&société*, 7, 2.
- [Purves et al., 2011] Purves, Ross S; Edwardes, Alistair J & Wood, Jo. 2011. «Describing place through user generated content». *First Monday*, 16, 9.
- [Rattenbury et al., 2007] Rattenbury, Tye; Good, Nathaniel & Naaman, Mor. 2007. «Towards automatic extraction of event and place semantics from flickr tags». *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pp 103–110.
- [Ratti et al., 2006] Ratti, Carlo et al. 2006. «Mobile Landscapes: using location data from cell phones for urban analysis». *Environment and Planning B: Planning and Design*, 33, 5, pp 727–748.
- [Ratti et al., 2010] Ratti, Carlo et al. 2010. «Redrawing the Map of Great Britain from a Network of Human Interactions». Sporns, Olaf (dir.). *PLoS ONE*, 5, 12, e14248–6..

- [Reades et al., 2007] Reades, Jonathan et al. 2007. «Cellular Census: Explorations in Urban Data Collection». *IEEE Pervasive Computing*, 6, 3, pp 30–38.
- [Reades et al. 2009] Reades, Jonathan; Calabrese, Francesco & Ratti, Carlo. 2009. «Eigenplaces: analysing cities using the space- time structure of the mobile phone network». *Environment and Planning B: Planning and Design*, 36, pp 824–836.
- [Robinet, 1957] Robinet, André; Leibniz, Gottfried Wilhelm & Clarke, Samuel. 1957. *Correspondance Leibniz-Clarke*, Paris: Presses universitaires de France.
- [Rogers, 2013] Rogers, Richard. 2013. *Digital Methods*. MIT Press.
- [Rouvroy & Berns, 2013] Rouvroy, Antoinette & Berns, Thomas. 2013. «Gouvernementalité algorithmique et perspectives d'émancipation». *Réseaux*, 177, 1, pp 163–196.
- [Ruppert, 2013] Ruppert, Evelyn. 2013. «Rethinking empirical social sciences». *Dialogues in Human Geography*, 3, 3, pp 268–273.
- [Salzberg, 2014] Salzberg, Steven. 2014. «Why Google Flu Is A Failure». *Forbes*, mis en ligne le 23 mars 2014, consulté le 19 mars 2015.
URL : www.forbes.com/sites/stevensalzberg/2014/03/23/why-google-flu-is-a-failure/.
- [Severo & Romele, 2014] Severo, Marta & Romele, Alberto. 2014. «Une approche philosophique de la ville numérique : méthodes numériques et géolocalisation» in Carmes, Maryse & Noyer, Jean-Max (dir.). *Devenirs urbains*, Paris: Mines ParisTech, pp 205–226..
- [Shoval & Isaacson, 2007] Shoval, Noam & Isaacson, Michal. 2007. «Tracking tourists in the digital age». *Annals of Tourism Research*, 34, 1, pp 141–159.
- [Song et al., 2010] Song, Chaoming et al. 2010. «Limits of Predictability in Human Mobility». *Science*, 327, 5968, pp 1018–1021.
- [Tarde, 2001 (1890)] Tarde, Gabriel. 2001 (1890). *Les lois de l'imitation*. Paris: Les empêcheurs de penser en rond.
- [Thrift, 2005] Thrift, Nigel. 2005. *Knowing Capitalism*. Thousand Oaks: SAGE.
- [Thrift & French] Thrift, Nigel & French, Shaun. 2002. «The automatic production of space». *Transactions of the Institute of British Geographers*, 27, 3, pp 309–335.
- [Tobler, 1970] Tobler, Waldo R. 1970. «A Computer Movie Simulating Urban Growth in the Detroit Region». *Economic Geography*, 46, pp 234-240.
- [Turner, 2006] Turner, Fred. 2006. *From Counterculture to Cyberculture*. Chicago: The University of Chicago Press.
- [Vaccari, 2009] Vaccari, Andrea et al. 2009. «Towards the SocioScope: an information system for the study of social dynamics through digital traces», ACM GIS '09 - Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, New York, NY, USA, pp 52–61.
- [Venturini & Latour, 2010] Venturini, Tommaso & Latour, Bruno. 2010. «Le tissu social : traces numériques et méthodes quali-quantitatives» in Chardonnet, Ewen (dir.). *Actes Futur en Seine 2009*, Paris:Cap Digital, pp 89–106.
- [Watts, 2011] Watts, Duncan J. 2011. *Everything Is Obvious: *Once You Know the Answer*. New York: Crown Business.
- [Wiener, 1950] Wiener, Norbert. 1950. *The human use of human beings: cybernetics and society*. Cambridge: Da Capo Press.
- [Yoshimura, 2014] Yoshimura, Yuji et al. 2014. «An analysis of visitors' behavior in The Louvre Museum: a study using Bluetooth data». *Environment and Planning B: Planning and Design*, 41, 6, pp 1113–1131.