



HAL
open science

Digital approaches and mobilities in the Big Data era

Jean-François Lucas

► **To cite this version:**

Jean-François Lucas. Digital approaches and mobilities in the Big Data era. Handbook of Urban Mobilities, 2020. hal-02939403

HAL Id: hal-02939403

<https://hal.science/hal-02939403>

Submitted on 15 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LUCAS Jean-François

Sociologue (PhD), expert ville numérique / Smart City chez Chronos, cabinet de conseil et de prospective.
Chercheur associé au Laboratoire de sociologie urbaine (LaSUR), École polytechnique fédérale de Lausanne (EPFL).

Digital approaches and mobilities in the Big Data era

2020

Attention, ce texte est un document de travail.

Il peut comporter des différences avec la version définitive du texte.

Merci de citer et de se référer à la version définitive :

Jean-François Lucas, « Digital approaches and mobilities in the Big Data era ». In: O.B. Jensen, C. Lassen, V. Kaufmann, M. Freudendal-Pedersen, I.S.G Lange, ed. 2020. *Handbook of Urban Mobilities*, London and New-York : Routledge International Handbooks, pp.77-86.

| |
|---|
| <p>Ce texte de travail est mis à disposition des internautes afin de faire connaître les travaux de LUCAS Jean-François. Les droits d'auteur de ce document demeurent aux auteurs des textes et/ou aux détenteurs des droits. Ce texte est destiné à une utilisation personnelle et non-marchande. Il ne peut être en aucun cas distribué et utilisé à des fins lucratives.</p> |
|---|

Chapter 9: “Digital Approaches and Mobilities in the Big Data era”

Jean-François LUCAS^{1 2}

<https://orcid.org/0000-0003-4284-321X>

1. Chronos, Paris, France
2. Urban Sociology Laboratory (LaSUR), École Polytechnique Fédérale de Lausanne, Suisse.

The data city

Data for optimisation

“In 2008, our global civilization reached three historical thresholds” (Townsend, 2013, 1). First, the world's urban population has become equivalent to the world's rural population of the world. Secondly, “the number of mobile cellular broadband subscribers surpassed the number of fixe DSL, cable, and fiber-optic lines”. Finally, during the same year we moved from the Internet of People to the Internet of Things (IoT) (*ibid.*).

The Internet of Things refers to a network of connected physical objects with their own digital identity and capable of communicating with each other. These objects can be mobile, such as our mobile phones, GPS, connected watches, transport cards with RFID chips. They can also be fixed, such as surveillance cameras or the various sensors that are increasingly hidden in each urban interstice. In 2010, the city of Santander, which has 180,000 inhabitants, deployed nearly 12,000 sensors to measure CO₂, NO₂, noise, light intensity, humidity, the amount of waste in the trash bins¹, etc. IT is therefore described as “*pervasive*” (Boullier, 2016, 48-49), because connected objects are deployed throughout our entire daily environment, to the point of constituting a “digital skin” (Rabari, Storper, 2014). Paradoxically, their multiplication favors their invisibility, because these technologies are integrated “into the fabric of everyday life until they are indistinguishable from it” (Wieser, 1991, 94).

¹. SmartSantander: <http://www.smartsantander.eu>

The data generated by these sensors lead to a multiplication of services and new uses related to the different functions of the city: moving around, living, working, lodging, etc. Many mobility applications allow knowing in real time public transport timetables, calculating alternative routes according to incidents taking place, finding a parking space near one's position, identifying points of interest, etc. At the city level, the analysis of mobility flows data allow to, for example, optimize travel times of city dwellers and to reduce pollution.

Technical and socio-political issues

Digital traces represent an unprecedented potential to understand and manage mobility issues, since they provide access to knowledge of spatial practices at the individual and the collective scales.

For users, services that use Big Data are decision support systems based on the promise of a personalized mobility experience that is simpler, more fluid, faster, cheaper and more social.

For companies, these data represent financial opportunities, be it through their monetization as a paid service or through the resale of the collected data to a third party.

For a municipality, Big Data offers the prospect of a better optimization and management of the services offered to users of transport networks, with a potential reduction in the effects of congestion and pollution. However, the plurality of actors and offers at the scale of a territory can sometimes complicate the management and regulation of mobility. This is for example what happens when Waze² suggests a user to take a road parallel to a main axis in order to avoid a traffic jam. This individual solution seems relevant, but becomes problematic as soon as users multiply and the application finally redirects an uninterrupted flow of vehicles onto secondary roads not intended to accommodate the flow of a main road (even if it is clogged). In addition to these structural problems, there are also security issues and the nuisances suffered by local residents. The issues and problems related to the use and interpretation of Big Data in the field of mobility are not only technical, they are fundamentally political.

In this chapter, we will discuss some of the possibilities offered by Big Data for capturing and understanding mobility practices. In the next part, we will present the ways in which the properties of these data can profoundly modify the analytical frameworks with which we understand mobility practices. Beyond the promises of Big Data, we will insist on the

². Mobile GPS navigation application acquired by Google in 2013.

importance of considering it, like any other technique, as human and therefore social constructions. Therefore, we will emphasize the importance of describing all the decisions, operations, constraints, etc. that shape these data and the algorithms they feed, in order to understand their limitations as well as to accept their part of subjectivity. Finally, we will discuss the challenges of Big Data for scientific research, from the point of view of the possibilities it offers to renew the understanding of mobility, but also the challenges it raises for the humanities and social sciences (HSS).

What Big Data changes

The 3V

A large quantity of various traces of our mobility practices are produced in real time thanks to fixed sensors (cameras, radars, telephone antennas, bicycle terminals, counting sensors...) and mobile sensors (GPS, smartphone...), or through our activities on the Web and social networks. These traces can be generated voluntarily or not, consciously or not. These data are called “Big Data” when they are varied, mass-produced and in real time. Although there are models with 7 or even 12 dimensions (Uprichard, 2013), “volume”, “variety” and “velocity” are the three most commonly cited dimensions. This is known as the “3V” of Big Data (Laney, 2001; Wilder-James, 2012; De Mauro, et al., 2016). In spite of the criticisms that highlight the weakness of these three dimensions to describe the diversity of what is commonly called “Big Data” (Kitchin, McArdle, 2016), it is interesting to understand how these data, because of the dimensions that characterize them, make it possible to rethink mobility.

Volume and correlation

The generation, collection, storage, processing and analysis of an ever-increasing volume of mobility data allows making spatial and therefore social practices visible, such as routes and traveling times. In addition to being more and more numerous, these data must be precise in terms of geolocation and time stamping to be relevant. These qualities can be specific to the data, or obtained by mixing different datasets, such as satellite data, road network or GSM data.

The analysis of a large number of accidentology or congestion-related data makes it possible, for example, to better understand the reasons (speed, dangerous passage, etc. in the case of accidents; urban planning, vehicle typology, meteorology, etc., for congestion), to predict or prevent occurrences in order to improve user safety and comfort. If certain variables such as

driving speed or blood alcohol content seem to be obvious accidentology factors, Big Data makes it possible to identify behavioural models (patterns) based on an ever-increasing number of variables, and therefore weak signals, that were difficult or even impossible to detect before.

The data are no longer used to test hypotheses, they allow the emergence of general laws based on correlations that will not “tell us precisely *why* something is happening, but will alert us to the *fact* that it is happening” (Mayer-Schonberger, Cukier, 2013, 14). In this sense, the inductive logic of the Big Data replaces the traditional deductive logic of the science.

These new possibilities have led some to speak of a new form of “data-driven” society (Pentland, 2012). Hypotheses, models, theories would then be obsolete in the face of data and algorithms that would find “patterns where science cannot” (Anderson, *ibid.*). All forms of subjectivity are thus evacuated, since, basically, “who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves” (Anderson, 2008).

This proposal elicited many comments, both on the possibilities offered by Big Data to discover new patterns and on the fact that this argument is part of a tradition that has seen this issue emerge each time a new form of “massive data” has appeared. Strasser thus recalls that “*perceptions of an "information overload" (or a "data deluge") have emerged repeatedly from the Renaissance through the early modern and modern periods and each time specific technologies were invented to deal with the perceived overload*” (Strasser, 2012, 85). Some consider that this does not translate a return to determinism but rather a “current mismatch between the characteristics of these massive data and social science’s methods and infrastructures” (Plantin, Russo, 2016).

Beyond the fuss that Anderson's “provocation” may have generated and the criticisms that have emerged in the academic sphere (Boyd, Crawford, 2012), the potential of the data and the correlations that can be made massively irrigate the industrial world and the direction that the development of new mobility services takes.

Variety and diversity

Variety illustrates the possibility of combining structured and unstructured data from multiple sources and different formats. It is thus possible to establish cross-references between descriptive data (ticket prices, theoretical timetables, geolocation of infrastructures, etc.), real-time data (geolocation of vehicles, occupancy rates, available fleet, video surveillance, incident alerts, etc.) or statistical data.

Many mobility services aggregate data from public and private actors relating to public transport, bicycle rental services, car sharing or carpooling, for example by compiling data on time (public transport timetables), stock (bicycles available at a terminal), cost (price of booking a car for a specific journey), etc., in order to offer a multimodal mobility service to users³.

Users can then compare offers and make choices according to several criteria: mode of transport, travel time, travel cost, type of lanes used, carbon footprint of the journey, or even more subjective criteria such as the type of landscape or odors they want to have on their route⁴.

Velocity and real time city

Velocity embodies the high frequency at which data are generated, collected, calculated and then published. While geolocated data are produced at a regular frequency, ranging from a few seconds to several minutes, their aggregation gives the impression that they are produced continuously.

Georeferenced and time-stamped digital traces make it possible to capture and make visible spatial practices over time (Beaude, 2015), i.e. frequencies, habits, rhythms... The visual representation of these flows on topographic maps may reveal the “pulse” of a city, which contributes to exacerbating the event dimension of contemporary urbanity (Picon, 2009).

These flows are grouped together with other data and other types of representation formats (figures, percentages, tables, videos, etc.) within dashboards, such as London’s⁵ or Dublin’s⁶,

³. "MaaS - Mobility as Services - consists of aggregating public and private transport offers into a single package of mobility services accessible to the user on a unified interface. It is certainly the concept that best embodies the service-switch that takes place in the world of mobility", in "Mobility as Networks", Le Lab OuiShare x Chronos, 2018: www.mobilityasnetworks.eu/publications-1

⁴. Goodcitylife.org : <http://goodcitylife.org>

⁵. <http://citydashboard.org/london>

⁶. <http://www.dublindashboard.ie/pages/index>

in order to facilitate the real-time management and regulation of the city and its flows (transport, energy, water, people, goods, etc.).

Rio de Janeiro's emblematic Operations Center undoubtedly embodies the pinnacle of the city "dashboard". Inaugurated in 2010, as a result of a partnership between IBM and the municipality of Rio, this project initially consisted in predicting and responding effectively to natural disasters. However, thanks to the interconnection of information from multiple sources (sensors, cameras, mobile phones, etc.), the intelligent operations center is able to assist the city in its daily operations and in many sectors (water and sanitation management, urban transport, traffic conditions, health system, civil security, law enforcement, waste collection, street lighting, housing, tourism, education system, etc.) as well as in emergency situations: heavy rainfall, road accidents, power cuts or landslides⁷. The operations room, made up of hundreds of screens, provides many of views of the city, its built environment, its flows, its inhabitants, etc. Like the eye of God in the sky overlooking the city, this project symbolizes the modern version of Bentham's Panopticon, the architectural device that "develops spatial units that allow to see continuously and to recognize immediately" (Foucault, 2016 [1975], 233).

Open the algorithm's hood

Data are "constructed data"

Data are a human production. They do not pre-exist to the many human operations that make it. Indeed, people are needed to create sensors and determine what they should pick up, how often, etc.

In fact, data are never "raw", they never speak for themselves (Bowker, 2005; Gitelman, Jackson, 2013). They are a production that makes us say things about the observed world, and produce it in turn. Whatever their origins, they are the result of multiple actions, techniques and ideologies of "data workers" (Bastard, et al., 2013) that shape them as soon as they are collected, stored, cleaned, processed, analyzed, shared, visualized... (Bowker, Star 1999; Ribes, Jackson, 2013). "What data are generated is the product of choices and constraints, shaped by a system of thought, technical know-how, public and political opinion, ethical considerations, the regulatory environment and funding and resourcing (...) Data then are situated, contingent, relational, and framed and used contextually to try and achieve certain aims and goals" (Kitchin,

⁷. "Smarter Business for a Sustainable Future ", IBM brochure available online.

Lauriault, McArdle, 2015, 21-23). To account for the goals and issues pursued, it is important to inform and know the choices made for each processing in order to make explicit (Boullier, 2016) how the data that are shared, visualized, consulted, etc., are the result of a co-construction process (Latour, 1999 ; Cames, Noyer, 2014).

Big is not all

While the prospects offered by Big Data for studying and understanding mobility are particularly stimulating, they should not hide limitations that are too often forgotten.

First of all, it must be considered that the amount of data obtained, however large, does not mean that it is exhaustive. There are always people who are not taken into account, either because they are not equipped with the appropriate technology or because they use a service different from that used, or because they do not want to, or because the data collection or the processing system have their own failures. Thus, with Big Data, “n=all” in which “all” is related to a particular system (Kitchin, McArdle, 2016, 8), itself “shaped by the technology and platform used, the data ontology employed and the regulatory environment, and it is subject to sampling bias” (Kitchin, 2014, 4).

Moreover, just because we access billions of data representing terabytes of digital information does not mean that we have more representative data. As large as these data may be, they are never global (Beaude, 2015). Moreover, linking digital data to identities, when it is not forbidden, is sometimes very complicated, as for social networks such as Facebook or Twitter where everyone is more or less free to create an alternative identity, with a pseudonym and an age that is not their own (Ravalet, Lucas, Lohou, 2017). Thus, “by insisting too often on the delicate representativeness of a population by a sample, we do not sufficiently understand what the data represent. Before representing a population, do the data accurately represent the individuals who make up the population?” (Beaude, 2015, 148).

The paradox of transparency

Few people probably know that the price of an Uber ride increases all the more as the battery level of the user's phone is low at the time of booking⁸. This example among many others

⁸. “This Is Your Brain On Uber”, National Public Radio, published on 17 May 2016:
www.npr.org/2016/05/17/478266839/this-is-your-brain-on-uber?t=1549134490084

illustrates the asymmetry that exists “between operators' demand for greater transparency on the part of users to be able to benefit from digital services, without them reciprocating. Uber knows everything about the identity of both drivers and customers on his platform, the routes they take, the rating they give each other. Conversely, the latter do not have all this information, either on the functioning of pricing algorithms or on those on users' rating. Trust offer is unilateral” (Basdevant, Mignard, 2018, 75).

The opacity surrounding the manufacture of algorithms and data processing complicates the contract of trust between a service provider and a user. However, the quality of the service provided is still the main criterion for membership, maybe because users are often unaware that their data are being sold to advertisers or insurance companies for current and future uses, the impact on personal life of which is still difficult to predict, both from a legal and ethical point of view.

At a time when figures are less and less representative of the social reality they are more and more shaping (Desrosières, 2008), it is becoming urgent to be able to analyze the “black boxes” (Pasquale, 2015) that are algorithms, since they construct forms and deploy statistical representations of society that organize the world in a certain way (Cardon, 2015).

In addition to the opacity of the algorithms, it is complicated to access the data that feeds them, when they are shared.

Researchers who have access to telephone operator data have often obtained a “pass” as part of a specific project. In this case, the rest of the scientific community does not have access to these data, which raises issues in terms of the very foundations of scientific practice, if only from the point of view of its validation or replicability. Academics who do not have an exclusive partnership with the big names in social networks must pay huge sums of money to access data relevant to their research purpose. For others, openly accessible data are often too fragmented to provide convincing results in a scientific study.

In addition, technical access to this type of data, whether through an API or a crawler, requires some specific skills. Finally, the issue of access to data also arises in terms of the balance of power that exists between private approaches within data-holding companies and public research.

Big data and mobility: challenges for research

Beyond travel, mobility

Beyond the perspectives mentioned in the field of mobility for public and territorial actors, companies and users, few academic actors recur to Big Data to question the concept of mobility, be it because social scientists are still not very interested in these data, because computer scientists who handle these data are not very interested in this concept, or because Big Data is mainly “movement” data.

However, mobility can be understood as the three-dimensional articulation of the realm of possibility to move around (available networks, performance, access conditions, territorial configurations, etc.), the potential for mobility that each individual acquires according to his or her life course, and actual movements in space (Kaufmann, 2008).

Big Data blur the lines between two currents of geography that study transport and ‘mobilities’, where ‘mobilities’ is considered as the central fact of modern and postmodern life (Urry, 2000; Hannam, et al., 2006; Sheller, Urry, 2006). Indeed, Big Data makes it possible to track flows and exchanges of people and goods, regardless of the micro, meso or macro scales. That allows us to capture some aspects of our contemporary, mobile and connected society. However, data collected from individuals or groups of individuals often only describe mappable and calculable movement.

So, while the realm of possibility or actual movements in space can easily be informed thanks to digital techniques, it seems more complex to obtain information relating to the life courses of individuals, defined as mobility capital or motility (Kaufmann, 2002), thanks to Big Data, since its paradigm is that of “the erasure of the subject, the subtle shift from the individual to the statistical individual and, more subtly, the dissolution of the subject in the network” (Beaude, 2015, 152).

From this point of view, qualitative methods “are the only way to access corners of social reality that are not yet digitized, or whose digital traces are privatized by large commercial companies that are reluctant to share them. Qualitative methods also remain of great interest as a complement to work using digital data: they provide a window into the feelings of individuals

[Kennedy, 2018], into their experiences and their interpretation of practices of which digital data are the trace, and which are often difficult to understand otherwise” (Bastin, Tubaro, 2018, 387).

Let me quickly point out that I have been referring since the beginning of this article to geographical mobility in an urban context, but that this would need to be extended to larger mobility, for example.

Big Data and mixed methods

Data from social networks can contribute to the analysis of mobility practices and representations, using a combination of qualitative and quantitative methods (Ravalet, Lucas, Lohou, 2017).

Nevertheless, while the combination of at least one qualitative element with at least one quantitative element is sufficient to characterize a mixed method (Bergman, 2008), several authors emphasize the need to go beyond the articulation of different data collection methods to include how qualitative and quantitative data analysis is conducted (Barbour, 1999; O’Cathain, Murphy, Nicholl, 2010). Two levels of articulation of a mixed method can thus be dissociated: the *primary level*, which corresponds to the types of combinations made between quantitative and qualitative methods to recover data, and the *synthesis level*, which refers to the level of analysis, depending on whether data processing is qualitative, quantitative or mixed; even though the latter option is not very used (Heyvaert, Maes, Onghena, 2013).

While some individuals have the skills and knowledge in *computer science*, HSS sufficient to conduct their research on their own, the use of Big Data to study mobility in the HSS field often involves collaborating with data analysts or computer scientists.

Mobility and Big Data at the crossroads of disciplinary fields

Collaboration between researchers in the humanities and social sciences and *computer scientists* is potentially fruitful from a scientific point of view because, “while computer scientists have produced powerful new tools for automated analyses of such “big data”, they lack the theoretical direction necessary to extract meaning from them. Meanwhile, cultural sociologists have produced sophisticated theories of the social origins of meaning, but lack the

methodological capacity to explore them beyond micro-levels of analysis” (Bail, 2014, 465). However, this type of collaboration implies negotiating as well as possible with the different “knowledge positions” (Darré, 1999; Morrissette, Desgagné, 2009) of the different actors, but also agreeing on terminology and a convergent vision (Star, Griesemer, 1989), which requires “articulation work” (Strauss, 1988), i.e. coordination operations to bring together different points of view, and time investments that can be costly. However, the purpose of these investments differs according to the actors involved, because researchers in the humanities and social sciences may seek to understand the meaning of the recovered data, while researchers in the computer sciences field may first seek to optimize algorithms (time gain, robustness and reliability of systems, etc.), without questioning the ontologies from which they condition data recovery.

Conclusion

Big Data is an opportunity to discover new patterns of behavior that had not been possible to observe at the individual and collective level until now. Nevertheless, if the discourses that surrounding them praise the many possibilities they open up for understanding such phenomena, particular attention must be paid to their properties, because "big" is not synonymous with exhaustiveness since it always has entities (individuals, objects, etc.) that are not taken into account, and to the algorithms that aggregate and calculate these data, because they are often "black boxes" that do not make it possible to understand how phenomena are analyzed, or even how they prescribed behavior, which raises many questions of equity, ethics and problems related to the general interest.

The data generated and collected in the field of mobility are generally related to travel by individuals or groups of individuals. However, if Big Data puts into question the role of quantitative methods, notably because of the rapid development of the machine learning (Bastin, Tubaro, 2018), qualitative methods still have a whole field of social issues that they are the only ones (for the time being?) to be able to explore.

In the coming years, we can only expect an increase in collaboration between researchers specializing in the processing of Big Data and researchers in the human sciences field specializing in mobility issues, as it is important to produce knowledge that goes beyond the analysis and understanding of movements, and to provide – particularly in the case of

sociologists - reflexivity to decision-makers in a political context in which “mobility” is a daily, social, economic, or even psychological issue for millions of people.

Bibliography

- Anderson, C., 2008. The end of theory. *Wired magazine*, [online] 23 June. Available at: <https://www.wired.com/2008/06/pb-theory> [Accessed 12 April 2019].
- Bail, C.A., 2014. The cultural environment: measuring culture with big data. *Theory and Society*, 43(3), pp.465-482.
- Barbour, R.S., 1999. The case of combining qualitative and quantitative approaches in health services research. *Health services research policy*, n°4, pp.39-43.
- Basdevant, A., Mignard, J.P, 2018. *L’empire des données. Essai sur la société, les algorithmes et la loi*. Paris: Don Quichotte éditions.
- Bastard, I., Cardon, D., Fouetillou, G., Prieur, C., Raux, S., 2013. Travail et travailleurs de la donnée. *Internetactu*, [online] 13 December. Available at: www.internetactu.net/2013/12/13/travail-et-travailleurs-de-la-donnee [Accessed 12 April 2019].
- Bastin, G., Tubaro, P., 2008. Le moment big data des sciences sociales. *Revue française de sociologie*, 59(3), pp.375-394.
- Beaude, B., 2015. Spatialités algorithmiques. In: M. Severo, A.Romele, ed. 2015. *Territoires et traces numériques*. Paris: Mines ParisTech, pp.135-162.
- Bergman, M.M. ed., 2008. *Advances in Mixed Methods Research*. London: Sage.
- Boullier, D., 2016. *Sociologie du numérique*. Paris: Armand Colin.
- Bowker, G.C., 2005. *Memory Practices in the Sciences*. Cambridge, MA: The MIT Press.
- Bowker, G.C, Star, S. L., 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: The MIT Press.
- Boyd, D., Crawford, K., 2012. Critical questions for big data. *Information, Communication & Society*, 15(5), pp.662–679
- Carmes, M., Noyer, J. M, 2014. L’irrésistible montée de l’algorithmique. *Les Cahiers du numérique*, 10(4), pp 63-102.
- Darré, J.P., 1999. *La production de connaissances pour l'action : arguments contre le racisme de l'intelligence*. Paris: Éditions de la Maison des sciences de l'homme et Institut National de la Recherche Agronomique.

- De Mauro, A., Greco, M., Grimaldi M., 2016. A Formal Definition of Big Data Based on its Essential Features. *Library Review*, 65(3), pp.122-135.
- Desrosières, A., 2008. *Gouverner par les nombres*. Paris: Presses de l'Ecole des mines.
- Foucault, M., 2016. *Surveiller et punir. Naissance de la prison*. Paris: Gallimard. First edition, 1975.
- Gitelman, L., Jackson, V., 2013. Introduction. In: L. Gitelman, ed. 2013. *“Raw data” is an oxymoron*. Cambridge, MA: The MIT Press, pp.1-14.
- Hannam K., Sheller M., Urry J., 2006. Mobilities, Immobilities and Moorings. *Mobilities*, 1(1), pp.1–22.
- Heyvaert, M., Maes, B., Onghena, P., 2013. Mixed methods research synthesis: Definition, framework, and potential. *Quality & Quantity*, 47, pp.659-676.
- Kaufmann, V., 2002. *Re-Thinking Mobility: Contemporary Sociology*. Farnham: Ashgate Publishing Company.
- Kaufmann, V., 2008. *Les paradoxes de la mobilité, bouger, s'enraciner*. Lausanne: Presses polytechniques et universitaires romandes.
- Kennedy, H., 2018. How People Feel about What Companies Do with their Data is Just as Important as What they Know about it. *LSE Impact Blog*, [online] 29 March. Available at: <<https://blogs.lse.ac.uk/impactofsocialsciences/2018/03/29/how-people-feel-about-what-companies-do-with-their-data-is-just-as-important-as-what-they-know-about-it/>> [Accessed 12 April 2019].
- Kitchin, R., 2014. Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), April–June 2014, pp.1-12.
- Kitchin, R., McArdle, G., 2016. The Diverse Nature of Big Data. *Big Data & Society*, 3, pp.1-10.
- Kitchin, R., Lauriault, T.P, McArdle, G., 2015. *Smart cities and the politics of urban data*. In: S. Marvin, A. Luque-Ayala, C. McFarlane, ed. 2015. *Smart Urbanism: Utopian Vision or False Dawn?* London: Routledge, pp.16-33.
- Laney, D., 2001. 3-D Data Management: Controlling Data Volume, Velocity and Variety. *META Group Research Note*, [online] 6 February. Available at: <<https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>> [Accessed 12 April 2019].
- Latour, B. 1999. *Pandora's Hope, Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.
- Mayer-Schonberger, V., Cukier, K., 2013. *Big Data: A Revolution That Will Transform How We Live Work and Think*. Boston: Houghton Mifflin Harcourt.

- Morrisette, J., Desgagne, S., 2009. Le jeu des positions de savoir en recherche collaborative: une analyse. *Recherches qualitatives*, 28(2), pp.118-144.
- O’Cathain, A., Murphy, E., Nicholl, J., 2010. Three techniques for integrating data in mixed methods studies. *British Medical Journal*, 341.
- Pasquale, F., 2015. *The Black Box Society, The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.
- Pentland, A., 2012. Reinventing Society in the Wake of Big Data. *Edge*, [online] 30 August. Available at: <https://www.edge.org/conversation/alex_sandy_pentland-reinventing-society-in-the-wake-of-big-data> [Accessed 12 April 2019].
- Wilder-James, E., 2012. What is big data? An introduction to the big data landscape. *O’Reilly Radar*, [online] 11 January. Available at: <<http://radar.oreilly.com/2012/01/what-is-big-data.html>> [Accessed 12 April 2019].
- Picon, A., 2009. Ville numérique, ville événement. *Flux*, 4(78), pp.17-23.
- Plantin, J.C., Russo F., 2016. D’abord les données, ensuite la méthode ? *Socio*, 6, pp.97-115. Available at: <<https://journals.openedition.org/socio/2328>> [Accessed 12 April 2019].
- Rabari, C., Storper, M., 2014. The digital skin of cities: Urban theory and research in the age of the sensed and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions Economy and Society*, 8(1), pp.27-42.
- Ravalet, E., Lucas, J.F., Lohou, A., 2017. Les retours d’une exploration méthodologique croisant données Twitter, recrutement via Facebook et questionnaires web. *Netcom*, 31(3/4), pp.309-334.
- Ribes, D., Jackson, S.J., 2013. Data bite man: The work of sustaining long-term study. In: L. Gitelman, ed. 2013. *“Raw Data” is an Oxymoron*. Cambridge, MA: The MIT Press, pp.147-166.
- Sheller, M., Urry, J., 2006. The new mobilities paradigm. *Environment and Planning A*, 38(2), pp.207-226.
- Star, S. L., Griesemer, J., 1989. Institutionnal ecology, ‘translations’, and boundary objects: Amateurs and professionals on Berkeley’s museum of vertebrate zoologie. *Social Studies of Science*, 19(3), pp.387-420.
- Strauss, A., 1988, The articulation of project work: an organizational process. *The Sociological Quarterly*, 29 (2), pp.163-178.
- Strasser, B., 2012. Data-driven sciences: From wonder cabinets to electronic databases. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(1), pp.85-87.
- Townsend, A., 2013. *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. New-York, London: W. W. Norton & Company.

Uprichard, E., 2013. Big data, little questions. *Discover Society*, [online] 1 October. Available at: <<http://discoversociety.org/2013/10/01/focus-big-data-little-questions>> [Accessed 12 April 2019].

Urry, J. 2000. *Sociology Beyond Societies: Mobilities for the Twenty-first Century*. London: Routledge.

Wieser, M., 1991. The Computer of the 21st Century. *Scientific American*, 265(3), pp.94-104.

Wilder-James, E., 2012. What is big data? An introduction to the big data landscape. *O'Reilly Radar*, [online] 11 January. Available at: <<http://radar.oreilly.com/2012/01/what-is-big-data.html>> [Accessed 12 April 2019].