An unexpected journey: A few lessons from sciences Po médialab's experience

Tommaso Venturini, Mathieu Jacomy, Axel Meunier, Bruno Latour

Original Research Article

BIG DATA & SOCIETY

Big Data & Society
July–December 2017: 1–11

DOI: 10.1177/2053951717720949
journals.sagepub.com/home/bds
SAGE

An unexpected journey: A few lessons from sciences Po médialab's experience

Tommaso Venturini[1], Mathieu Jacomy[2], Axel Meunier[3] and Bruno Latour[4]

Abstract

In this article, we present a few lessons we learnt in the establishment of the *Sciences Po médialab*. As an interdisciplinary laboratory associating social scientists, code developers and information designers, the médialab is not one of a kind. In the last years, several of such initiatives have been established around the world to harness the potential of digital technologies for the study of collective life. If we narrate this particular story, it is because, having lived it *from the inside*, we can provide an intimate account of the surprises and displacements of digital research. Founding the médialab in 2009, we knew that we were leaving the reassuring traditions of social sciences to venture in the unexplored territory of digital inscriptions. What we couldn't foresee was how much such encounter would change our research. Buying into gospel of Big Data, we imagined that the main novelty of digital research came from handling larger amounts of data. We soon realized that the interest of digital inscriptions comes instead from their proliferating diversity. Such diversity encouraged us to reshape our professional alliances, research practices and theoretical perspectives. It also led us to overcome several of the oppositions that used to characterize social sciences (qualitative/quantitative, situation/aggregation, micro/macro, local/global) and to move in the direction of a more continuous sociology.

Keywords

Actor-network theory, computational social sciences, datascape navigation, digital methods, digital sociology, quali-quantitative methods

A peek in our logbook

In the last couple of decades, the advent of digital media has created much ferment in the social sciences. The availability of a new source of information on collective phenomena has encouraged a growing number of academic institutions to create interdisciplinary initiatives at the crossroad of media studies, computational sciences and information design. Among these we can mention (citing only the groups with which we had direct contact): the *Digital Methods Initiative* of the University of Amsterdam (digitalmethods.net); the *Institutes for Complex Systems* of Rhône-Alpes (ixxi.fr) and Ile de France (iscpif.fr); the *Institut Francilien Recherche Innovation Société* (ifris.org); the *Centre for Interdisciplinary Methodologies* of the University of Warwick (warwick.ac.uk/fac/cross_fac/cim/); the *Density Design Lab* of the Politecnico if Milano (densitydesign.org); the *Laboratório de estudos sobre Imagem e Cibercultura* of the University of the University of Espírito Santo (labic.net); the *Medialab of the University of Rio de Janeiro* (medialabufrj.net); the V*irtual Observatory for the Study of Online Networks* of the Australian National University (vosonlab.net); the *Social Science Matrix* of the University of Berkeley. This list if far from exhaustive (a search in this very journal will easily return dozens of other groups), yet it gives an idea of how diffused and

[1]Univ. Lyon, Inria, ENS de Lyon, CNRS, UCB Lyon 1, 69342, Lyon, France
[2]Axel Meunier and Bruno Latour- Sciences Po Paris, médialab, Paris, France
[3]Independent Researcher
[4]Sciences Po Paris, médialab, Paris, France

Corresponding author:
Tommaso Venturini, King's College London, Room 207, 26-29 Drury Lane Strand Campus, London WC2B 5RL, UK.
Email: tommaso.venturini@kcl.ac.uk

diverse is this movement. The médialab of Sciences Po is only one of these initiatives. In this article, we will focus on its story not because it is unique, but because it will help us to narrate the peculiar way in which the encounter with 'the digital' has affected our research.

The tone of our text will be rather intimate. We are less interested here in proposing concepts or techniques that in conveying the peculiar atmosphere that characterizes these early years of digital methods (see Latour et al., 2012, for a more theoretical outlook). Richard Rogers, one of the protagonist of this movement, argued that digital methods are about 're-purposing' new media for social investigation (Rogers, 2013). While this is true, it is also interesting to appreciate how digital techniques have repurposed social research *in return*. This article narrates the story of this repurposing and leads the reader through the encounters that, in the first years of the médialab, changed our way of doing social research. Through its three sections, the article discusses how the digital transformed our relation to the data, the methods and the theory of social research.

While none of these transformations is entirely new, the way in which the médialab weaved them in a coherent narrative is, we believe, original. The *fil rouge* (the guiding thread) of this narrative is the way in which digital technologies displace and eventually dissolve the classic distinctions of our discipline: qualitative/quantitative in data, situation/aggregation in methods and micro/macro in theory. Encounter after encounter, our unexpected journey brought us beyond such distinctions, toward a more 'continuous sociology'.

A last note on style: in this article, we make an abundant use of metaphors and not by chance. The transformations that digital methods are encouraging are deep but still unaccomplished. Different approaches still co-exist and their outcomes can be varied. This openness is one of the things that we like the most about our field and we do not want to close it too soon or too much. Metaphors will allow us to provide a sense of where digital research is heading while leaving open the range of possible explorations.

Three displacements of digital traces (continuity in data)

When the Sciences Po médialab left the harbor in May 2009, we all had in mind stories of the *miraculous draught* of digital data. Having heard our fair share of Big Data tales, we cast our nets in the digital sea and waited for the miracle. Withdrawing our nets, however, we found them full to the brim, but not of data. What we got was instead a catch of *digital traces*, different from the Big Data in at least three respects. (1) The main novelty of digital technologies is their increased traceability, (2) which does not zero the cost of social data, (3) but augment their diversity. We will describe these three displacements as we experienced them at the médialab, but others scholars have provided similar accounts (see, for example, Bollier and Firestone, 2010; Boyd and Crawford, 2011; Marres, 2012; Uprichard, 2013).

(1) By 'digital traces', we intend loosely all the inscriptions produced by digital devices in their mediation of collective actions – for instance, a post published on a blog, a hyperlink connecting two websites or the log of an e-commerce transaction. Though the terms 'inscriptions', 'traces' and 'data' are often used as synonyms, it is here important to distinguish their different meanings.

'Inscription' is the more general terms and it refers to any piece of information that is materialized through a technical device (digital or not). Inscriptions are the foundation of any scientific enterprise for they allow to *imprint* knowledge on material supports, which can be stored, transformed and transmitted (Latour, 1983, 1990, 1979; Latour and Woolgar, 1979). For most natural and social researches, inscriptions are produced by devices specifically set up for this intent. Sometime, however, inscriptions are produced by other purposes and become object of scientific enquiry only later. This is distinctively the case for digital research, which is often performed as a form of secondary analysis of inscriptions produced for marketing, surveillance or technical purposes (cf. Ruppert et al., 2013).

The practice of using *second-hand* inscriptions for social investigation does not start with digital media. Since the invention of writing and increasingly with the introduction of the alphabet, the printing press, the broadcast media and the communication technologies in general, our collective life is organized through a web of material inscriptions. And, since their foundation, social sciences have exploited such inscriptions for their research. In this sense, there is nothing new about digital media (Boullier, 2015), *except for* the amplification that they impose to such process (in line with McLuhan, 1964, hypothesis of media acceleration). As social sciences are concerned, the main novelty of digital media is to have assembled an unprecedentedly extensive ensemble of inscribing infrastructure (Bowker et al., 2009).

Such crucial novelty was initially overlooked by social sciences. Opposing 'the virtual' to 'the material', early media scholars regarded to electronic networks as a new terrain for old methods. Notions such as 'cyberculture' (Negroponte, 1995), 'virtual communities' (Rheingold, 1993), 'online identities' (Turkle, 1995) were introduced to channel the novelty of new media within the tradition of social sciences (Rogers, 2009).[1995, 1993] Such approach, however, concealed

the most interesting feature of electronic media: the fact that, far from making social relations more virtual, they multiply their material inscriptions. Passing through routers, cables, databases, our associations leave traces that can be easily recorded, massively conserved and instantly retrieved:

> Once you can get information as bores, bytes, modem, sockets, cables and so on, you have actually a more material way of looking at what happens in Society. Virtual Society thus, is not a thing of the future, it's the materialisation, the traceability of Society. It renders visible because of the obsessive necessity of materialising information into cables. (Latour, 1998)

(2) As digital media increase the traceability of collective phenomena, our initial Big Data enthusiasm seemed justified. As we soon realized, however, the proliferation of digital inscriptions is not per se a gain for the social sciences. Here is where we need to distinguish between 'digital traces' and 'digital data'. With the term *traces*, we refer to inscriptions as *originally* produced by digital devices (e.g., a hyperlink as stored in a HTML code of a web page). With the term data, we refer to the same inscriptions having undergone the cleaning and refining necessary to make them useful knowledge objects (e.g., our hyperlink as translated in an edge of a website network).

While such distinction is somewhat artificial (there are no such things as 'raw traces' and all inscription processes entail adjustments and correction), it has the advantage of drawing attention to the multiple transformations that inscriptions must undergo before they can be exploited (see Bachelard, 1949, on 'evidence work' and Cardon, 2013, on 'digital evidence work'). In the case of digital methods, much of this process happens outside the academia, with bias that digital researchers are forced to inherit. The problem with digital data is not that they are biased (all data is). It is that the control that social scientists can exert on their bias is remarkably poor (Savage and Burrows, 2007). Digital data is still at an early stage in the process of sociological validation and should therefore be used with caution and never without investigating the conditions of their fabrication (Marres and Gerlitz, 2016).

To be exploited in research, digital data needs to be detached from their original contexts and refitted for research. Such 'repurposing' work (Rogers, 2013) requires significant efforts and specific expertise. This is why the Sciences Po médialab could not exist without its data experts. Digital media are undoubtedly a fertile ground, but they yield nothing if they are not skillfully labored. Engineers, computer scientists, developers, data geeks know how to nurture this kind of data in ways that social scientists still have to learn. Also, being more familiar with digital technologies, they are also more resistant to technological hype. Being acquainted to the hardships of 'data friction' (Edwards, 2010), they know that only a little fraction of digital inscriptions will transmute into usable data and only through a good deal of their work.

(3) Left as adventurers of the digital *El Dorado*, we settled as data farmers. We learned that information infrastructures pay the costs of the inscription of digital traces, but not those of their transformation into sociological data. 'More traces', therefore, does not mean 'more data', but more sources from which data can be extracted. At the beginning of our adventure, we neglected such crucial distinction and paid the consequences in our choice of sources. Like kids in a toy store, we went for the biggest. We coveted the huge reservoirs of Web 2.0 platforms, and forgot that the largest sources are not always the most valuable. Platforms such as Google, Facebook or Wikipedia do produce large quantities of traces (and do re-distribute some of them). Yet, such abundance of traces does not always convert into valuable data (Lazer et al., 2014) and mainstream social media are not always the most interesting sources. Google is ideal for studying the flow of online attention (Choi and Varian, 2009); Facebook is a decent proxy of Web sociability (Rieder, 2013); Wikipedia is great for studying knowledge debates (Borra et al., 2014; Niederer and Dijck, 2010). Still, none of them can answer all the questions of the social sciences. Nor should they: *there are more things in Internet and the Web (Horatio) than are indexed in your social media!* Seen from the outside, the jungle of digital inscription meets the eye for its extension. But from the inside, its most amazing feature is the stunning diversity of the species that it shelters.

An example comes from a research that the médialab carried out on the international negotiations on climate change. In two projects dedicated to the climate debate (www.emapsproject.com and projetmedea.hypotheses.org), we were looking for information on the United Nations Framework Convention on Climate Change (UNFCCC). Conventional data on such negotiations are available through the reports of ethnographers and through the UNFCCC documents archive (unfccc.int). Such data allows investigating respectively the detail of diplomatic practices and the overall structure of the climate regime, but not the connection between the two. Moving in controversy analysis tradition (Pestre, 2007; Pinch, 2015; Venturini, 2010, 2012), we wanted to follow the rise and fall of issues; the forming of alliances and oppositions; and the positioning of delegations on items of the climate agenda. We wanted to observe the dynamics of the 'diplomatic ballet' and the complex arrangements produced by such dynamics.

To follow these discursive actions, we first thought of exploiting Twitter and investigate climate talks through the messages exchanged by negotiators and policy-makers. Despite the existence of a remarkable project by UN Climate Change Secretariat (climatetalkslive.org), we found Twitter traces disappointing. Such traces were indeed massive (with thousands of messages exchanged every day), but they shed little light on the negotiations. Tweets are just too short to convey the subtleties of climate diplomacy and are generally addressed to the public more than to other negotiators. They allowed us to extract the general agendas pushed by different countries (and their reception by the media), but not to follow the negotiation dynamics.

So, we looked further and came across the Earth Negotiation Bulletin (www.iisd.ca/vol12), compiled by the International Institute for Sustainable Development (IISD). The ENB reports on all UN negotiations covering both the official proceedings and the discussions in the corridors. The ENB can hardly be called Big Data. Its traces are neither big (the volume dedicated to climate negotiations contains about 600 issues of a few pages each) nor 'natively digital' (the reports are first distributed *in paper* to the negotiators and then archived on the IISD website). The ENB inscriptions are not as rich as ethnographic observations (the reports iron out all nuances to propose a dry paraphrasing of the arguments exchanged) and not as extensive as the UNFCCC archive (the reports cover the positions taken in the meetings, but disregard the materials submitted by parties or produced by bodies of the Convention). Yet, the ENB traces have enough breadth to follow climate debate throughout the two decades of the UNFCCC negotiations *and* enough depth to identify single diplomatic moves (cf. Venturini et al., 2014a, and climatenegotiations.org).

Of course, this comes at a price. The ENB reports are not written for scholars, but for negotiators. Each issue summarizes one day of negotiations, compiling in one document different tracks and formats that researchers might want to separate (e.g., IPCC debates cannot be studied in the same way as UNFCCC's; discussions in the corridors demand a different treatment than talks in the plenaries). Also, the ENB reports include 'duplicates' in the form of summaries of the 'previous episodes' and analyses of the negotiations. Exploiting the ENB traces demanded therefore an extensive work of cleaning and refining that ended up occupying most of our project time.

The ENB corpus thus illustrates the three displacements characteristic of digital traces: (1) it is not larger or more detailed than traditional datasets; (2) it demands a thorough refinement before it can be used for research; (3) but it allows investigating climate negotiations *in breadth and depth*. This third displacement is, we believe, the most important. As we said, digital data is not necessary more abundant than their predecessors and they are certainly not cleaner or cheaper. They are, however, more diverse and more evenly distributed across the span of collective existence of which they therefore offer a more continuous appraisal.

Datascape navigation (continuity in methods)

The diversity of digital traces holds great potential, but it also demands to rethink our methods and, in particular, the classic opposition between qualitative and quantitative approaches. In our sociological training, we learned that knowledge can be obtained by two opposite strategies: either by focusing on particular events and collecting as much information as possible; or by selecting specific bits of information and harvesting them through time and space. Through qualitative methods, we learned to examine local interactions. Through quantitative methods, we learned to skim over global trends. Both strategies are undeniably productive in their sphere and most scholars are capable to alternate and mix them. Yet, the discontinuity between the two types of methods generated a blind spot in our vision. We can glance at the global picture and examine each piece of interaction, but we cannot observe how the puzzle of collective life is put together.

The arrival of digital traceability may bridge such discontinuity by generating information on an increasing variety of phenomena and, most importantly, at a variety of scales. By their proliferating diversity, digital traces diffract so much the perspectives of social research and dissolve traditional oppositions (Latour et al., 2012). In this sense, our approach may be called 'quali-quantitative' (Venturini et al., 2015a), but the label is inappropriate as digital methods do not lie *between* qualitative and quantitative methods. Nor do they get the best of both worlds: the quantity of digital data is not always larger than traditional statistics and their quality is often inferior to ethnographic observations. Yet, the more the digital media infiltrate collective life, the more varied become the traces that social scientists can exploit. The following figure tries to render this idea. It is far too simplistic, but has one quality: instead of representing the classic quali/quantitative *continuum*, it outlines a space in which inscriptions distribute unevenly but not binarily (Figure 1).

Looking at the figure below, readers will remark that data (i.e., inscriptions refined to serve as basis for research) tend to be more polarized than inscriptions and have a more skewed coverage of the social space. Not only traditional methods start from different types

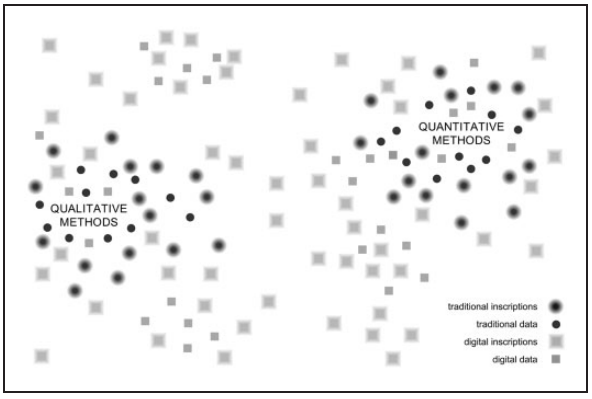

Figure 1. Distribution of traditional inscriptions (blurred black circles), traditional data (clear black circle), digital traces (blurred grey squares) and digital data (clear grey square).

of inscriptions, but they also refine them in opposite directions. Most traditional research protocols are unidirectional: they have been sharpened to produce more *situation* or more *aggregation*. Situated observation allows preserving the full richness of the situation, but entails renouncing to follow the influences that associate 'here and now' to other times and places. Aggregation allows extending the scope of sociological vision, but demands focusing on specific observables (e.g., not persons in their full complexity, but say their age or income) and concentrating on regularities rather than exceptions.

Part of the discontinuity between situation and aggregation derives from the difficulty of rewinding such procedures and to reverse the transformations that they operate. To be sure, good researchers have always provided their readers the means to reverse their protocols (Latour, 1995). Yet, before digitalization, the variety of research devices (notepads, audio and video recorders, questionnaires, data tables, typewritten reports, calculators, printed texts, etc.) made reversing objectively difficult. Disaggregating a statistical chart entailed, for instance, moving from the chart to the calculator, from the calculator to the data table, from the table to the research notebook, from the notebook to the sample and from the sample to the actual phenomenon. With the advent of digital technologies, an increasing number of research transformations take place in the same personal computers. This does not mean, of course, that such transformations have got any easier. It means, however, that they can be made reversible with reasonable effort.

Unfortunately, this opportunity is not always exploited and too often the technical expertise required by digital methods becomes an extra obstacle to the rewinding of research procedures. This is why, *particularly in digital research*, findings should always be accompanied by the data on which they are based, the traces from which this data is extracted and the scripts that allowed the extraction (cf. Ince et al., 2012; Willinsky, 2006). To this purpose, Rieder and Röhle (2012) encourage digital researchers to complement their articles with a 'companion web-site':

> Such a website could further document methodological choices, provide in-depth information on technical aspects, and, if possible, source code. It could also, as part of a data-sharing plan, provide access to the 'raw' data that went into the project, allowing other researchers to experiment with different approaches and arrive at different results. Also, a companion website seems like an ideal vehicle to present results more dynamically and interactively. Instead of providing the research results as a closed, finished product, Web-based interfaces could allow audiences to explore them both inductively and deductively, involving them in the process of knowledge production. (pp.80,81)

A more reversible digital publication not only improves transparency and replicability, but transforms the way we read, write and design research. A good digital research does more than catering a series of results to its readers, it provides them interfaces to navigate through the inscriptions.

This was never clearer to us than in an investigation that the médialab carried out on French parliamentary system (but see another example of our work in Leclercq, 2017). In such investigation, we were working with political scientists to assess how much French laws are substantially amended through parliamentary discussions. Both qualitative and quantitative methods seemed unfit for the project. We could have dissected the parliamentary journey of a few bills to examine their transformation, but how to know if findings could be generalized? And we could have devised some statistical measures of parliamentary transformation and compute them for all French laws, but how to know whether those metrics were not too simplistic (and capable to differentiate substantial from cosmetic modifications)?

Eventually, we created an online platform to explore the details some 300 laws proposals (www.lafabriquedelaloi.fr). The platform allows comparing how long different laws were discussed in different branch of the parliament and how many words were changed through such discussion (see Figure 2(a)); identifying how much each article of each law has been modified at each passage (see Figure 2(b)); considering all the amendments proposed by different political groups (see Figure 2(c)); and reading the transcription of each word spoken by each parliament member on each specific article at each stage of the discussion (see Figure 2(d)).

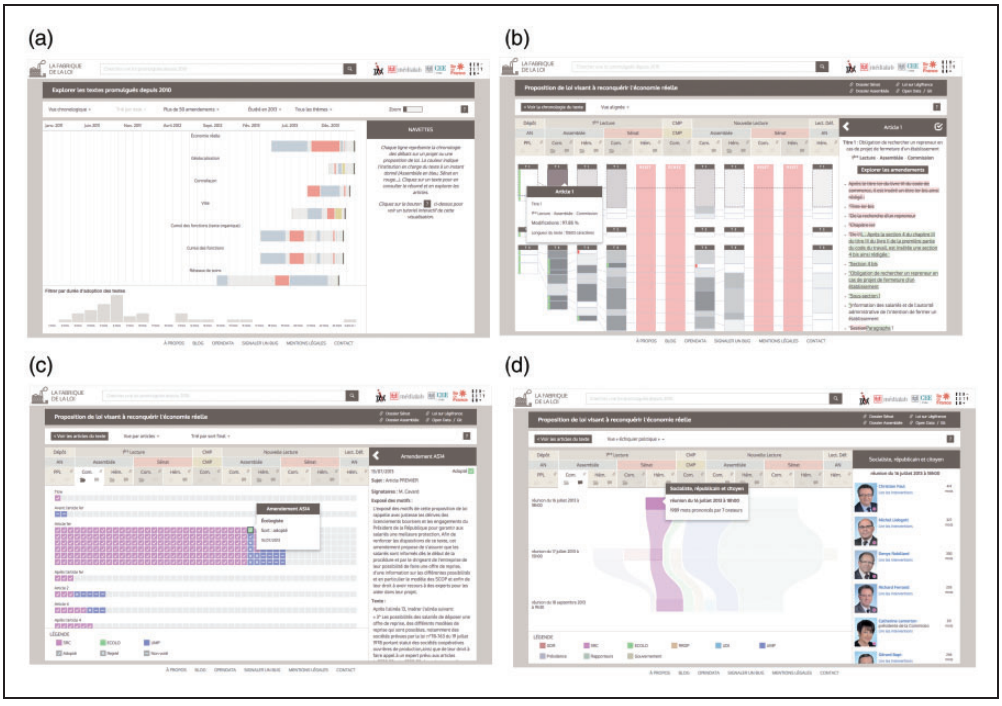

Figure 2. Four interfaces in La Fabrique de la Loi project (http://www.lafabriquedelaloi.fr/).

The development of such platform marked the médialab more than we expected (Venturini, 2017). Building ways to move from the most aggregated overviews to the debates' minutes (and back), we felt the discontinuity of social sciences fading away. Not only we encountered little opposition between the quality and the quantity of our data (parliamentary logs offer one and the other, provided that one is willing to invest the efforts to mine and refine them), but we also realized that situation and aggregation could happily coexist as functions of the same digital platform. Datascape navigation (as call this experience) recognizes the importance of both the examination of specificities and the generalization of trends, but instead of looking for a balance or compromise, it aims facilitating the movement between the two. And this is where the journey of the médialab met another type of expertise: besides social scientists and developers, our team came to include designers.

Unlike quantitative methods, the purpose of our research is not to deliver stable findings to its readers and, unlike qualitative methods, it is not to immerse them in specific situations. Its purpose is to offer its readers (but 'users' would be a more appropriate term) interfaces to browse large datasets and identify interesting data points. And this is where the help of interface design and information design (Bertin, 1967; Tufte, 1983) comes in handy. By multiplying the operations that users can perform on digital inscriptions, datascape design dissolves the discontinuity between situation and aggregation. Not only by finding ways of reconciling these movements, but also by showing they are only two among the many ways of navigating through data. In this sense, quali-quantitative methods are utterly different from 'mixed methods' (Johnson and Onwuegbuzie, 2004). Our objective is not to combine qualitative and quantitative techniques, but to overcome their opposition altogether, drawing on the increased diversity of data to promote a greater diversity of methods. We as are interested in alternating situation and aggregation, then in exploring other navigational operations such as highlighting, sorting, filtering, zooming, selecting, comparing, link surfing and annotating (cf. Heer and Shneiderman, 2012; Yi et al., 2007).

If all these operations existed before computers, their digital instantiation has greatly increased the ease with which they can be performed. In a sense, datascape navigation materializes the dream of the *Exploratory Data Analysis* (Behrens and Chong-Ho, 2003; Tukey, 1977) but also its nightmare – the risk that more exploration possibilities might end up producing greater disorientation. Over-complicated dashboards only add an 'interface overload' to the 'information overload' already experienced by digital researchers. This is why the contribution of designers is crucial to carefully select and combine different forms of navigation into meaningful forms of data exploration.

Beyond micro and macro (continuity in theory)

The dissolution of the quali/quantitative divide encouraged us to take one step further and rethink the theoretical opposition between individual actors and social collectives, micro-interactions and macro-structures. Central in social research at least from Durkheim (1897, 1912), such binary conception was convenient (even indispensable) when the separation between qualitative and quantitative methods pulled research in opposite directions. It provided a handy conceptualization of the two types of phenomena that situation and aggregation can best describe, respectively, local and ephemeral exchanges and global and enduring trends. However, as digital traceability and datascape navigation makes data and methods more continuous, the micro/macro distinction appears less significant.

> The best proof that those two levels do not correspond to any real ontological domains is that they begin to disappear, to be literally redistributed, every time one modifies or enhances the quality of access to the datasets, thereby allowing the observer to define any actor by its network and vice versa. This is exactly what the striking extension of digital tools is doing to the very notions of 'individual' and 'wholes'. The experience (more and more common nowadays) of navigating on a screen from elements to aggregates may lead researchers to grant less importance to those two provisional end points. Instead of having to choose and thus to jump from individuals to wholes, from micro to macro, you occupy all sorts of other positions, constantly rearranging the way profiles are interconnected and overlapping. (Latour et al., 2012: 595)

Digital research dissolves in practice a distinction that we still struggle to overcome in theory. Following the lead of its founder (Latour, 2005) and the lesson of Actor-Network Theory (Law and Hassard, 1999), the médialab has always tried to dispense with local/global distinctions. But attacking the micro/macro divide in theory, we were constantly trapped in using the same concepts that we tried to refuse (the very name 'actor-network' is a perfect example of such paradox). Talking of 'quali-quantitative methods', 'navigating through situation and aggregation' and 'overcoming the micro/macro divide', we cannot avoid giving the wrong impression that the adventure is located *between and not beyond* these traditional oppositions. Much clearer than any theoretical argument, the research practices embedded in our tools and methods demonstrate what a continuous sociology can deliver (tools.medialab.sciences-po.fr).

To understand the advantages of a continuous research, consider a playful diagram designed by Andy Greenwald (grantland.com/features/the-hbo-recycling-program) to highlight how the HBO television network 'recycle' the same performers in different series (see Figure 3(a)). Besides showing that most performers are connected to several lines (hence the 'recycling'), the diagram does not convey much. In particular, it is difficult observe which performers co-appear in the same series and which series share the same performers. This information is hidden by the separation between performers and series. Since the performers are aligned on the left of the diagram (the micro) and the series are aligned on the right (the macro), all other groupings disappear. It is enough, however, to remove this separation – and let performers and series position themselves *on the same plan* – to highlight other fractures and other groupings (notably the 'gangster' cluster at the bottom of the network) (see Figure 3(b)).

At the médialab, we encountered the same experience in a more academic case (Venturini et al., 2016). Using a network visualization (see Figure 4(b)), we tried to overcome the separation between the individuals and the institutions of the French dominant class, which a famous table by Luc Boltanski (1973) separated in rows and columns (see Figure 4(a)). Representing individuals and institutions on the same plan we could highlight not only how many connections had each entity (as in the table), but also how far in the social space did such connections extended.

Besides hiding other (often more relevant) fractures, the micro/macro distinction has another disadvantage: it conceals the work invested to build individual interactions and collective aggregates, conceiving the former given and the latter as 'emerging'. Introduced in biology, the notion of 'emergence' has been developed to express the idea that complex systems (e.g., living organisms) have properties (e.g., the fact of being alive) that are impossible to anticipate from their components. An emergent phenomenon is 'unlike its components insofar as these are incommensurable, and it cannot be reduced to their sum or their difference' (Lewes, 1875: 412, cfr. the 'Emergent Properties' entry in the *Stanford Encyclopedia of Philosophy*, 2012).

Transported in the social sciences, the notion of emergence risks however to *naturalize* (and *de-politicize*) collective systems. If institutions and norms are emergent, then none can be held responsible for them. They derive from the interactions among a multitude of actors, but not correspond to the intentions of any of them. The metaphor recurrently employed is that of ants' colonies. Like ants reacting to the pheromones that they encounter, but incapable to appreciate the larger architecture of their nests (Garnier et al., 2007), social actors would create global structures by reacting to their local environment.

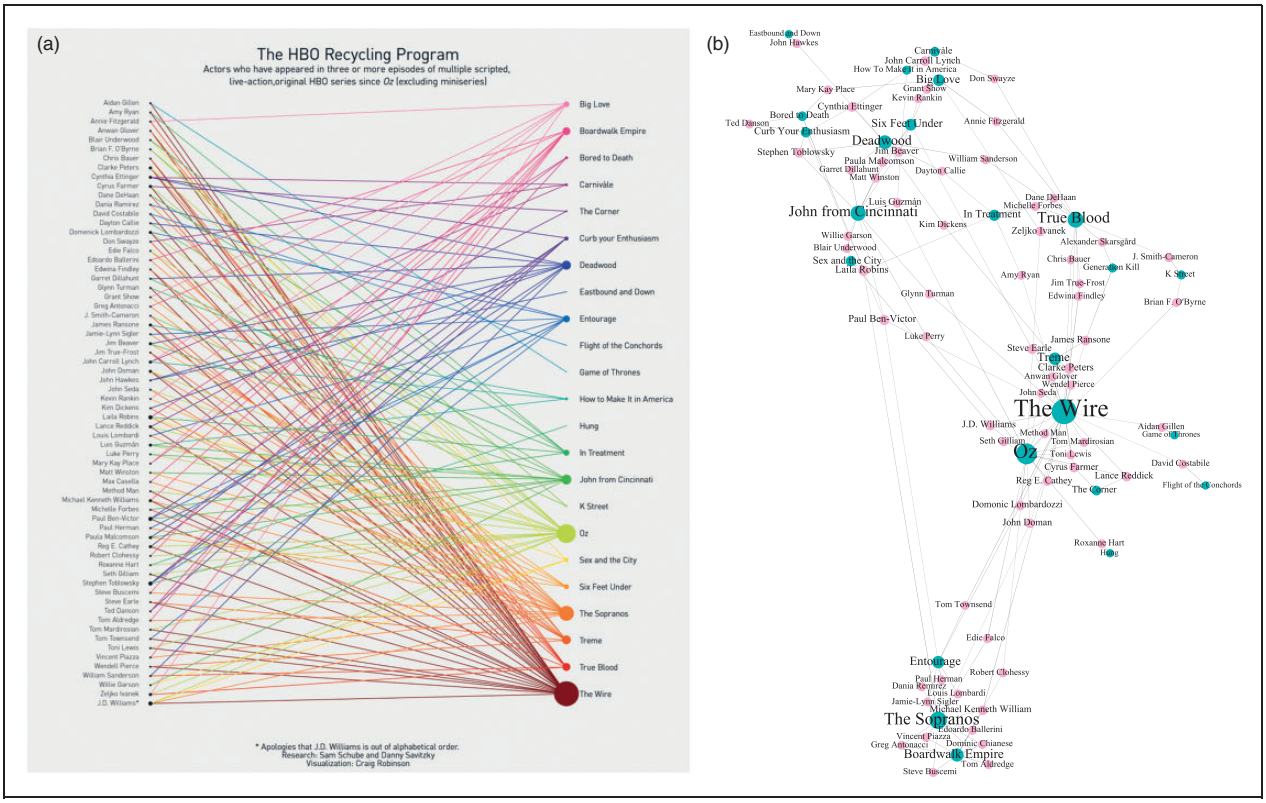

Figure 3. The network of HBO casting spatialized by (a) separating or (b) merging performers and series.

Remarkably, this 'emergent' approach characterized the first applications of digital computation to social sciences – particularly through agent-based models (see Castellano et al., 2009, and the special issues of *PNAS*, 99-3 and *American Journal of Sociology*, 110-4). Attributing global structures to mysterious emergent dynamics (that, by definition, cannot be studied) made sense when the opposition between qualitative and quantitative methods offered little information on anything else than situated interactions and aggregated trends. Yet, as digital traceability allows investigating collective actions at a plurality of levels, the opposition between 'micromotives and macrobehaviors' (Schelling, 1978) lose much of its interest. Even worse, such binary approach has become a tangible obstacle to understanding some of the collective crises faced by our societies (Venturini et al., 2015b).

At the médialab, we stumbled on such an obstacle while mapping the debate on climate change in the EMAPS project (Venturini et al., 2014b, http://climaps.eu). Such debate is blocked (among many other reasons) because of the tendency to present personal choices or worldwide agreements as the only possible dimensions to tackle the climate crisis (Aykut and Dahan, 2015, call this the 'reality schisms' of international climate regime). While both these levels are clearly important, they are by no means the only ones. Countless other collective actions are possible and necessary at other scales, including industrial policies, consumptions trends, ecological resiliencies, local communities, etc. While all these scales have been studied by social scientists, the discontinuity of our methods and theories have made it difficult to observe the 'scale-breaking' relations that connects, for example, the international funding decisions to the adaptation projects carried out in specific districts of Bangladesh (http://climaps.eu#!/narrative/climate-adaptation-in-bangladesh); or the discussions about natural and social resilience to the climate investments in different regions of Germany (http://climaps.eu#!/narrative/who-deserves-to-be-funded).

Inviting us to navigate through different scales of actions, digital methods have profound political implications. Alongside the transformation of digital inscriptions, another (more important) operation is taking place: the crafting of alternative narratives about social life. By dissolving the binary distinction between qualitative and quantitative methods, situation and aggregation, digital research makes space to a new set of collective stories. Such stories involve a multitude of actors other than individuals and structures – associations, communities, clusters, networks, interest groups, concerned publics and a variety of other characters. Though these characters have long been known

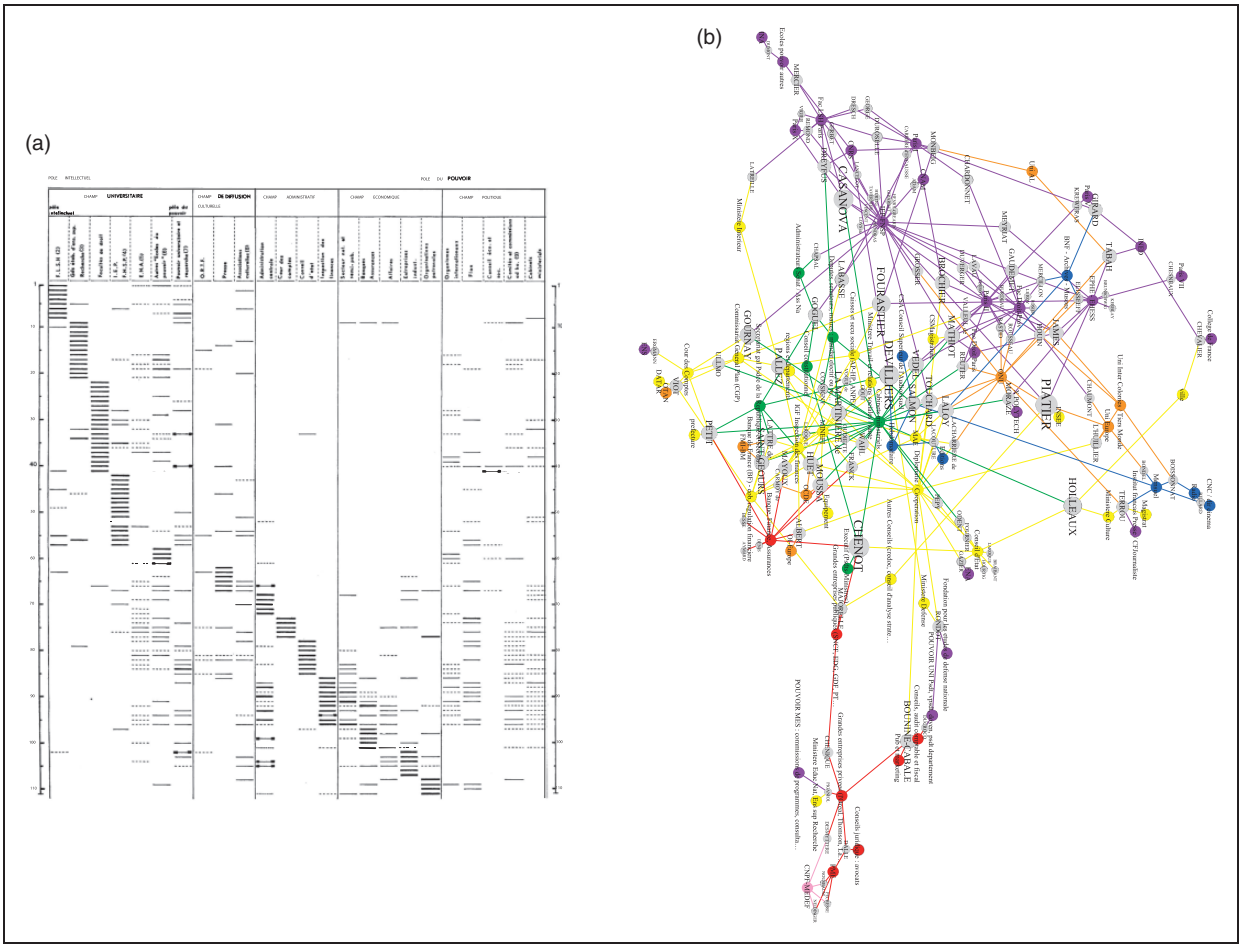

Figure 4. Luc Boltanski original table of the French dominant class and its redesign as a network.

by social sciences, the 'flatting' effect produced by the continuity of digital methods suddenly bring them together and force them to share the same stage.

Welcoming these actors and their interference is both crucial and crucially difficult. As we tried to convey in this article, there is nothing automatic about digital methods. The variety of digital traces comes with the responsibility to chose among different sources. Say we want to work on climate adaptation (as we did in EMAPS), where should we start from? from the report of the OECD Official Aid Report; from the UNDP database of adaptation projects; or from tweets about global warming? And once the sources chosen, we still have to decide how we will transform their inscriptions; how we will analyse the resulting data; and which forms of navigation we will encourage with our interfaces. While we know that each of these decisions will have a substantial influence on our research outcomes, we often lack the knowledge necessary to make them. Our methodological focus makes us experts on digital techniques, not on the issues we study through them. And this is where the dialogue with a variety of social actors becomes crucial. Provided that we find ways to organize the various (and sometime conflicting) contributions coming from, e.g., a spokesperson of the French Finance ministry, a development specialist working on an adaptation project in Tuvalu, or Bangladeshi negotiator at the UNFCCC.

Organizing the heterogeneity of these contributions and redistributing research decisions demand once more to renew our scientific practices. Different research formats have been experimented at the médialab, like workshops, data sprints (Munk et al., 2016; Venturini et al., 2016a), simulations and exhibitions. These different types of events are meant to engage different types of publics: workshops, for example, address fellow researchers; data sprints address experts; simulations address students; exhibitions address the general public. Yet, all these formats represent 'extreme' versions of the way in which we do research on a daily basis: we work collaboratively on data, we imagine uses and users for our datascape and we show our tools to actors, experts and

communities. Crafting these interactions has become an integral part of our approach to digital methods, a way of doing research in which face to face interactions ended up being every bit as important as digital computation.

The journey of the médialab is not finished and other unexpected encounters await in the future of social sciences. It is not too early, however, from taking stock of what we have learned. By our trials and by our errors, we experienced how different is the practice digital research from the dream of Big Data. The digital is not a land of abundance. It is not a place where information pour in freely or easily; not a place where computational tricks, powerful as they may be, can replace the hard work necessary to mine, nurture and refine inscriptions. Digital methods do not spare us from walking the walk, but they give us the chance to experiment new pathways. The diversity of digital traces and techniques challenges disciplinary orthodoxies. It compels us to consider new ways of collecting and redistributing information; it pushes us out of our comfort zone; forces us to ally with developers, designers and with a multiplicity of social actors. None of this is easy, to be sure, and nothing assures that the promise of digital methods will be kept. Pursuing a more continuous sociology is not the same thing as achieving the continuity of the social. Our journey is far from reaching its harbor, but 'the sails are set, the wind is east, the moorings fret'.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

American Journal of Sociology (2005) Special Issue on Agent-Based Modeling 110(4)..

Aykut SC and Dahan A (2015) *Gouverner Le Climat?* Paris: Presses de Sciences Po.

Bachelard G (1949) *Le rationalisme Appliqué*. Paris: PUF.

Behrens JT and Chong-Ho Y (2003) Exploratory data analysis. In: Weiner IB (ed.) *Handbook of Psychology*. London: Wiley, pp. 33–64.

Bertin J (1967) *Sémiologie Graphique*. Paris: La Haye, Mouton.

Borra E, Weltevrede E, Ciuccarelli P, et al. (2014) Contropedia – the analysis and visualization of controversies in Wikipedia articles. In: *OpenSym 2014 Proceedings*.

Boyd D and Crawford K (2011) Six provocations for big data. In: *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*. Available at: https://ssrn.com/abstract=1926431 or http://dx.doi.org/10.2139/ssrn.1926431 (accessed September 2011).

Bollier D and Firestone CM (2010) *The Promise and Peril of Big Data*. Washington: The Aspen Institute.

Boltanski L (1973) L'espace positionnel: multiplicité des positions institutionnelles et habitus de classe. *Revue Française de Sociologie* 14(1): 3–26.

Boullier D (2015) Pour des sciences sociales de troisième génération: des traces numériques aux répliques. In: Menger PM (ed.) *Big Data, Entreprises et Sciences Sociales*. Paris: Open Editions Press (Collège de France).

Bowker GC, Baker K, Millerand F, et al. (2009) Toward information infrastructure studies: Ways of knowing in a networked environment. In: Hunsinger J, Klastrup L, Allen MM, et al. (eds) *International Handbook of Internet Research*. Dordrecht: Springer, pp. 97–117. Available at: http://doi.org/10.1007/978-1-4020-9789-8_5.

Durkheim E (1897) *Le Suicide*. Paris: Presses Universitaires de France.

Durkheim E (1912) *Le Formes Elémentaires de la vie Religieuse*. Paris: Presses Universitaires de France.

Cardon D (2013) Travail et travailleurs de la donnée. *Internet Actu*. Available at: http://www.internetactu.net/2013/12/13/travail-et-travailleurs-de-la-donnee/ (accessed 5 July 2017).

Castellano C, Fortunato S and Loreto V (2009) Statistical physics of social dynamics. *Reviews of Modern Physics* 81: 591–646.

Choi H and Varian HR (2009) Predicting the present with Google trends. *Papers.ssrn.com* 15(8): 1327–1328.

Edwards PN (2010) *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge MA: The MIT Press.

Garnier S, Gautrais J and Theraulaz G (2007) The biological principles of swarm intelligence. *Swarm Intelligence* 1(1): 3–31.

Heer J and Shneiderman B (2012) Interactive dynamics for visual analysis. *Queue - Micoprocessors* 10(2): 30–55.

Ince DC, Hatton L and Graham-Cumming J (2012) The case for open computer programs. *Nature* 482: 485–488.

Johnson RB and Onwuegbuzie AJ (2004) Mixed methods research: A research paradigm whose time has come. *Educational Researcher* 33(14): 14–26. Available at: http://doi.org/10.3102/0013189X033007014.

Latour B and Woolgar S (1979) Laboratory life. *The Construction of Scientific Facts*. Los Angeles: Sage.

Latour B (1983) Give me a laboratory and I will raise the world. In: Knorr-Cetina K and Mulkay M (eds) *Science Observed*. London and Beverly Hills: Sage, pp.141–170.

Latour B (1990) Drawing things together. In: Lynch M and Woolgar S (eds) *Representation in Scientific Practice*. Cambridge, MA: MIT Press, pp. 19–68.

Latour B (1995) The ''Pédofil'' of Boa Vista: a photo-philosophical montage. *Common Knowledge* 4(1): 144–187.

Latour B (1998) Thought experiments in social science: from the social contract to virtual society. In: *1st Virtual Society? Annual Public Lecture*. London: Brunel University1 April.

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

Latour B and Woolgar S (1979) *Laboratory Life. The Construction of Scientific Facts*. Los Angeles: Sage.

Latour B, Jensen P, Venturini T, et al. (2012) 'The whole is always smaller than its parts': a digital test of Gabriel Tardes' monads. *The British Journal of Sociology* 63(4): 590–615.

Lazer D, Kennedy R, King G, et al. (2014) Big data. The parable of Google Flu: traps in big data analysis. *Science (New York, NY)* 343(6176): 1203–1205.

Law J and Hassard J (1999) *Actor Network Theory and After*. Oxford: Blackwell.

Leclercq C (2017) Le Projet E.A.T. Datascape ou ce que le design numérique peut faire à l'histoire sociale de L'art. In: *Colloque ''Des Pratiques Antiquaires Aux Humanités Numériques''*. Paris: Ecole du Louvre.

Lewes GH (1875) *Problems of Life and Mind*. London: Kegan Paul, Trench, Turbner, and Co.

Marres N (2012) The redistribution of methods: On intervention in digital social research, broadly conceived. *The Sociological Review* 60(June): 139–165.

Marres N and Gerlitz C (2016) Interface methods: Renegotiating relations between digital social research, STS and sociology. *The Sociological Review* 64(1): 21–46. Available at: http://doi.org/10.1111/1467-954X.12314.

McLuhan M (1964) *Understanding Media: The Extensions of Man*. New York: McGraw-Hill.

Munk AK, Meunier A and Venturini T (2016) Data sprints: A collaborative format in digital controversy mapping. In: Ribes D and Vertesi J (eds) *DigitalSTS: A Handbook and Fieldguide* (forthcoming).

Negroponte N (1995) *Being Digital*. New York: Alfred A. Knopf.

Niederer S and van Dijck J (2010) Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system. *New Media & Society* 12(8): 1368–1387. Available at: http://doi.org/10.1177/1461444810365297.

Pestre D (2007) L'analyse de Controverses Dans L'étude Des Sciences Depuis Trente Ans. Entre Outil Méthodologique, Garantie de Neutralité Axiologique et Politique. *Mil Neuf Cent. Revue D'histoire Intellectuelle* 1(25): 29–43.

Pinch T (2015) Scientific controversies. *International Encyclopedia of the Social & Behavioral Sciences*. Amsterdam: Elsevier.

PNAS (2002) Adaptive agents, intelligence, and emergent human organization: Capturing complexity through agent-based modeling (special issue), 99(suppl. 3).

Rheingold H (1993) *The Virtual Community: Homesteading on the Electronic Frontier*. Cambridge, MA: MIT Press.

Rieder B (2013) Studying Facebook via data extraction: The Netvizz application. In: *Proceedings of WebSci '13, the 5th annual ACM web science conference*, pp.346–355. Available at: http://doi.org/10.1145/2464464.2464475 (accessed 6 July 2017).

Rieder B and Röhle T (2012) Digital methods: Five challenges. In: Berry DM (ed.) *Understanding Digital Humanities*. Houndmills: Palgrave Macmillan, pp. 67–84.

Rogers R (2009) *The End of the Virtual: Digital Methods*. Amsterdam: Amsterdam University Press.

Rogers R (2013) *Digital Methods*. Cambridge, MA: MIT Press.

Ruppert E, Law J and Savage M (2013) Reassembling social science methods: The challenge of digital devices. *Theory, Culture & Society* 30(4): 22–46.

Savage M and Burrows R (2007) The coming crisis of empirical sociology. *Sociology* 41(5): 885–899. Available at: http://doi.org/10.1177/0038038507080443.

Schelling TC (1978) *Micromotives and Macrobehavior*. New York: W W Norton & Company.

Stanford Encyclopedia of Philosophy (2012) Emergent properties. Available at: https://plato.stanford.edu.

Tufte ER (1983) *The Visual Display of Quantitative Information*. Cheshire: Graphics Press.

Tukey JW (1977) *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Turkle S (1995) *Life on the Screen. Identity in the Age of the Internet*. New York: Simon and Schuster.

Uprichard E (2013) Big Data, Little Questions? *Discover Society*, no. 1. Available at: http://www.discoversociety.org/focus-big-data-little-questions/ (accessed 6 July 2017).

Venturini T (2010) Diving in Magma: How to explore controversies with actor-network theory. *Public Understanding of Science* 19(3): 258–273.

Venturini T (2012) Building on faults: How to represent controversies with digital methods. *Public Understanding of Science* 21(7): 796–812.

Venturini T (2017) The fish tank complex of social modelling. In: Nagatsu M and Ruzzene A (eds) *Frontiers of Social Science: A Philosophical Reflection*. New York: Bloomsbury Publishing.

Venturini T, Baya Laffite N, Cointet J-P, et al. (2014a) Three maps and three misunderstandings: A digital mapping of climate diplomacy. *Big Data & Society* 1(2): 1–19. Available at: http://doi.org/10.1177/2053951714543804.

Venturini T, Meunier A, Munk AK, et al. (2014b) Climaps by Emaps in 2 Pages (A summary for policy makers and busy people in general). *Social Science Research Network* (ID 2532946). Available at: http://papers.ssrn.com/abstract=2532946 (accessed 6 July 2017).

Venturini T, Cardon D and Cointet J-P (2015a) Méthodes digitales: Approches quali/quanti des données numériques - Présentation du numéro spécial. *Réseaux* 188: 9.

Venturini T, Jensen P and Latour B (2015b) Fill in the gap: A new alliance for social and natural sciences. *Journal of Artificial Societies and Social Simulation* 18(2): 11.

Venturini T, Jacomy M, Baneyx A, et al. (2016) Hors champs: La multipositionnalité par l'analyse des réseaux. *Reseaux* 199(5): 11–42. Available at: http://doi.org/10.3917/res.199.0011.

Willinsky J (2006) *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, MA: MIT Press.

Yi JS, Kang YA, Stasko J, et al. (2007) Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics* 13(6): 1224–1231.