

The Social Fabric: Digital Traces and Quali-quantitative Methods

By
Tommaso Venturini and Bruno Latour
médialab — Sciences Po Paris

Introduction

According to Livy in 493 B.C. the Roman plebeians withdrew to the Sacred Mount to protest against the Senate. The patricians sent Menenius Agrippa to negotiate with the rebels. Agrippa convinced the rebels to come back to the city by haranguing the mob with a br fable, which William Shakespeare later rewrote (Coriolanus, Act 1, Scene 1):

*There was a time when all the body's members
Rebell'd against the belly, thus accused it:
That only like a gulf it did remain
I' the midst o' the body, idle and unactive,
Still cupboarding the viand, never bearing
Like labour with the rest, where the other instruments
Did see and hear, devise, instruct, walk, feel,
And, mutually participate, did minister
Unto the appetite and affection common
Of the whole body. The belly answer'd*

...

*'True is it, my incorporate friends,' quoth he,
'That I receive the general food at first,
Which you do live upon; and fit it is,
Because I am the store-house and the shop
Of the whole body: but, if you do remember,
I send it through the rivers of your blood,
Even to the court, the heart, to the seat o' the brain;
And, through the cranks and offices of man,
The strongest nerves and small inferior veins
From me receive that natural competency
Whereby they live: and though that all at once,
You, my good friends,'--this says the belly, mark me,--*

...

*'Though all at once cannot
See what I do deliver out to each,
Yet I can make my audit up, that all*

*From me do back receive the flour of all,
And leave me but the bran.'*¹

When Menenius Agrippa recited his fable, medical arts were still far from developing the technique of scintigraphy. Such diagnostic technique makes it possible to follow the movements and accumulation of radioactive isotopes administered to a patient. With the help of scintigraphy, the body's members personified in the fable would have been able to know not only how the stomach distributed the nutrients, by which channels and in what proportions, but also how each member cooperates with the others in the human body. If they had known all this, Agrippa's story might have ended quite differently...

It is often said that digital technologies are revolutionizing the social sciences, much as they revolutionized the natural sciences a few decades ago. This is not the case. The situation in the social sciences is more like that of the natural sciences in the 15th century, just after the introduction of the printing press. In *The Printing Press as an Agent of Change*, Elizabeth Eisenstein (1979) thoroughly describes how the introduction of the press triggered the Scientific Revolution of the XVI century. Yet, as Eisenstein makes very clear, the revolution did not broke overnight. Several decades after Gutenberg invention, naturalists were still printing the same mistakes they used to hand-copy. It was only after press attained a certain level of maturity and circulation that it started to have a real impact on the natural sciences.

Social sciences are today in a similar position. Rather than drawing on digital data to revitalize their practices, they are still trying to pass them off as new terrains for old methods. *Cyberculture, virtual communities, online identities, computer-mediated communication*: all these notions have been developed to quarantine the novelty of electronic media. And yet, the very speed at which digital technologies infiltrate modernity makes such resistance more and more untenable. It is no longer possible to circumscribe digital interactions to a specific sector, as their threads are now ubiquitously woven into the fabric of collective existence. Digital media offer much more than just another field to apply existing methods: they offer the possibility of restructuring the study of social existence (Rogers, 2004).

Up to now, access to collective phenomena has always been both incomplete and expensive. Compared to their colleagues in the natural sciences, social scientists have always been relatively poorly equipped. While physicists could follow billions of atoms in their accelerators and biologists could grow millions of microbes under their microscopes, social scientists could not simultaneously maintain breadth and depth in their observations. Their methods offered them a bipartite view of social existence, as they could either focus on specific interactions or skim the surface of global

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

structures.

Among other things, this explains why in social sciences the notion of emergence has taken a different meaning than in the natural sciences. In physics, biology, and medicine, the notion of emergence is used to describe phenomena of self-organization due to the spontaneous coordination of many agents. In the social sciences, emergent phenomena become structures, classification tables hosting and influencing interactions. This structuralist vision is due to a great extent to the fact that the social sciences have never had methods to reconnect micro and macro and show how global phenomena are built by the assemblage of local interactions. Digital technology promises to revolutionize this situation, providing the social sciences the possibility of following each thread of interaction and showing how social life is woven together by their assemblage.

The Problem of Social Life

The problem of social life is that every situation involving more than two people can quickly become unbelievably complicated. Take the complexity of a human being and multiply it by the number of people involved in the interaction and then by the number of interactions needed to organize the simplest of social phenomena. Add to that the fact that interactions can affect and be affected by any kind of non-human agent (technologies, elements of nature, pieces of art...) and that these agents can carry the consequences of other interactions stretching far into time and space. If we are able to make anything of this complexity, it is because our societies can count on a large number of simplifying devices such as institutions, norms, standards, procedures, etc. Despite what we might think, the so-called *emergent* social structures are in reality simpler than the interactions of which they are made.

Emergence is a very useful notion, an idea that benefits both social actors and social scientists: it allows actors to assume the stability of collective phenomena and allows researchers to presume that these phenomena make sense (Johnson, 2001). However, the notion of emergence conceals entirely different meanings in these two cases. For social actors, there is no other way to access emergent properties than constructing them. Everyone who has tried to gather a rock band, organize a chess tournament, or start a knitting circle knows this all too well: social phenomena do not exist by themselves. Collective phenomena emerge through the work of participants, extend as far as the actors carry them, and last as long as they keep them up.

To be sure, social facts can be materialized: decisions can be transcribed, borders reinforced by barriers, relationships sealed with symbols. Social structures can be stabilized by material infrastructures (Latour, 2005), but it is only through actors' coordinated work that collective phenomena can emerge and last (Weisman, 2007). This labor, however, is only rarely recognized by the social sciences that, since their earliest days, never had the resources necessary to explore the construction of social structures. They can follow a few specific interactions, but breaking down social facts into the thousands of constituent interactions has always been beyond their reach. And

so, they preferred to avoid the question.

In the 19th century, a young sociology found the workaround it needed in statistics. By combining the two separate mathematical traditions of probability calculus and error estimation, statistics provided the social sciences with a convenient shortcut (Desrosières, 1998). Instead of tracing the way emergent structures were built by simplifying thousands of local exchanges, the researchers could gain access to social phenomena through a *mathematical leap*. Where social actors *assembled*, the researchers could *aggregate*. For example, instead of studying how criminal organizations selected and socialized their recruits, the researchers could calculate the profile of the “average criminal.” Instead of studying how crime rings spread their investments through the coordination of a multitude of individual criminals, the researchers could estimate the *average incidence* of each crime. The gain in efficacy was tremendous. Not only could the social sciences finally portray social structures, they could even contribute to their solidification.

Of course, the adoption of statistics came at a cost. In bypassing the work of collective construction via the statistical shortcut, researchers introduced a fictive distinction between micro-interactions and macro-structures. Though such distinction has never been anything more than an optical illusion reflecting the leap of statistical aggregation, social scientists always found it very convenient. Through the micro/macro distinction, researchers who were interested in local interactions could observe them without having to mess with the tangle of influences that surround each social event. And, those who wanted to observe the global picture of social life were could do it without bothering about single brush-strokes. By separating micro and macro, the social sciences simultaneously validated qualitative *and* quantitative methods.

Still, the dualistic approach based on the micro/macro distinction has a major disadvantage: it only works *after* the actors have finished their work of simplification and *until* that they manage to maintain it. One cannot cross-sample, if there are no clear boundaries among classes, groups or genders or. One cannot rely on averages or normal curves, if actors do not conform to norms or normality. One cannot estimate deviance, if deviant behaviors are deterred or sanctioned. Statistical analyses can approximate collective facts only to the extent that they have been defined, normalized and stabilized by social actors. The dualistic approach can describe *emerged* phenomena rather well, but it is useless with *emerging* phenomena.

Learning from Controversies

The weakness of the dualistic approach becomes particularly clear if, instead of concentrating on areas of consensus, we turn to the study of controversies, observing social life where agreement is no longer (or not yet) reached. In the study of disputes, we can depend neither on quantitative methods (because the local positions are so varied that there is nothing to be gleaned from aggregating them into a general average) nor on qualitative methods (because even the most specific controversy is always part of a global network of influences). Where disputes rage, where actors cannot come to agreement, where social ties come undone or, quite simply, where change unfolds so quickly that old institutions dissolve before new ones can be created, quantitative and qualitative methods reach their limits.

This is why the social sciences have always been uncomfortable with controversies: their methods

demand that they reduce them to residual phenomena or, at best, to transitional stages between periods of stability. This is particularly evident in the sociology of science and technology. For ages, arguments among scholars were considered to be insignificant events on the road to scientific advancement. Even authors like Thomas Kuhn (1962) who took scientific controversies very seriously ended up defining them as passing phases between two periods of “normal science.”

Much of the interest of Science and Technology Studies, as developed by the Edinburgh (Bloor, 1976) and Bath (Collins and Pinch, 1993) Schools, lies in having shown that, far from being residual or transitional phenomena, controversies are the true engine of science. In science and technology (Bijker and Law, 1992), disagreements and conflicts are the norm and consensus can only be reached through tremendous efforts of coordination and stabilization (Callon, Lascoumes, and Barthe, 2009). Scientific paradigms are not coherent systems that channel scholars' work. Instead, they are the fragile product of this work. They require endless negotiations for their upkeep and their existence is constantly threatened by the risk that these negotiations might fail (Latour, 1988).

Working with students at the Paris Institut d'Études Politiques in a course entitled *Cartography of Controversies*,² we came to understand that, far from being limited to the sciences, this situation is common to all collective phenomena. Controversies have taught us that nothing in our social lives endures without changes. Communal life resembles less a long, quiet river than a restless sea of thousands currents that no social contract could ever definitively calm. Thus, social institutions are not structures imposing themselves on individuals but a sort of “Law of War” always subject to negotiation. Sure these institutions exist, and sure they have effects, but these effects are the result of interactions rather than being their premise.

The combination of stability and change that characterizes controversies cannot be untangled using a dualistic approach (Venturini, 2010a). As they keep us from breaking down social disputes into the network of multiple oppositions that characterize them, quantitative methods can only address the few elements that create consensus among the actors (and which are typically the least interesting part of a dispute). As they keep us from showing how every quarrel is tied to a broader network of opposition, qualitative methods inhibit any understanding of how equilibrium points can be temporarily negotiated while leaving the controversy open. Studying controversies requires a new set of “quali-quantitative” methods that allow us to trace social phenomena throughout the processes of construction, deconstruction, and reconstruction that constitute them.

The Promise of Computerized Methods

Quali-quantitative methods are not merely the juxtaposition of statistical analysis with ethnographic observation. Erasing the micro/macro border implies being prepared to collect and process an unprecedented amount of data. Being interested in the construction of social phenomena implies tracking each of the actors involved and each of the interactions between them. This was an impossible goal just a dozen years ago, but it starts to become more and more realistic as digital technologies spreads.

² <http://controversies.sciences-po.fr>

The interest of electronic media lies in the fact that every interaction that passes through them leaves traces that can be easily recorded, massively stored and inexpensively retrieved. Each day, researchers discover new pools of digital data: public and private archives are sucked into computer memory, economic transactions migrate online, social networks take root on the web. Digital mediation spreads out like a giant roll of carbon paper, offering the social sciences more data than they ever dreamt of.

Thanks to digital traceability, researchers no longer need to choose between precision and scope in their observations: it is now possible to follow a multitude of interactions and, simultaneously, to distinguish the specific contribution that each one makes to the construction of social phenomena. Born in an era of scarcity, the social sciences are entering an age of abundance. In the face of the richness of these new data, nothing justifies keeping old distinctions. Endowed with a quantity of data comparable to the natural sciences, the social sciences can finally correct their lazy eyes and simultaneously maintain the focus and scope of their observations.

The potential of this quali-quantitative approach has been illuminated by the MACOSPOL (*Mapping Controversies on Science for POLitics*) project. Funded by the European Union, this project coordinated eight European labs in an effort to inventory and test the digital tools available for controversy mapping (Venturini, 2010b).

While the quantity of the identified resources was encouraging, project results also revealed the relative immaturity of digital methods. Today, more and more tools are available to extract and analyze electronic data and many researchers have begun using them (Lazer et al., 2009). Still, the success of these initiatives remains limited by their intermittent nature. In most cases, the scholars restricted themselves to digitizing one piece or another of their research without reconsidering the whole research process. The results of this approach are certainly interesting and have the advantage of being easily comparable with those of classical methods. Yet, a true transition to quali-quantitative methods will not be possible until the social science research chain is entirely digitized. For the new methods to realize their innovative potential, it is necessary that each step in the research chain be rethought in a coherent manner:

1. The identification of data sets should privilege digital archives, the web, online media and networks, digital documentation and literature, and, more generally, all sorts of digital traces.
2. The extraction of data should be based on assisted browsing or on different techniques of data collection (crawling, querying, scraping, parsing...).
3. Data from diverse sources should be integrated in a common relational database (*data mashing*).
4. The analysis and modeling of data should lean on web, digital, and, where possible, *open source* tools.
5. Research results should be published on the web, preferably in open archives and in a standard format in order to make their reuse easier.

It is easy to see that the process we describe is in fact a loop. Available on the web, the results of a digital investigation can become the data for a new analysis. Even more important: the digital data as well as their extraction and analysis tools can be published alongside the results, making accessible to the public not only the conclusions of a research project, but all the path that have led

to them. This is the true appeal of digital methods: instead of just describing the experimental procedures it is now possible to give direct access to them. In a digital research chain, no path is blocked, no transformation is irreversible, no result is given. Every element in the chain is made available to the public and is open to scrutiny. In a digital research project, navigating the data and navigating the chain of data processing are ultimately the same thing.

The Quali-quantitative Oligopticon

The promise of digital methods lies in the possibility of opening the research chain of social sciences. To keep this promise, however, the whole research chain must be digitized. As long as they will be confined to one part of the research process, digital methods will not deploy their innovative potential and will risk generating all sorts of incompatibility effects.

For example, the question of representativeness perfectly illustrates the difficulty of mixing digital and traditional methods. A criticism that has often been addressed at the analysis of digital traces is that these traces regard a non-representative sample of society. From the point of view of the traditional social sciences, this is certainly true. It is well known that digital literacy is not uniformly diffused in society (men, young people, and those with high levels of education are generally overrepresented in online samples). Still, this disproportion is a problem only as long as we insist on treating digital data as if they were survey data. The advantage of the new methods is that they allow tracing the assemblage of collective phenomena instead of obtaining them through statistical aggregation. The question of representativeness is thus posed in an entirely different way.

Digital data are representative and interesting only if their processing chain (identification, extraction, integration, analysis, publication) remains close to the work of social actors. To be sure, we are not saying that quali-quantitative methods will allow us to smooth out all the complexity of collective life. Quite the contrary, the advantage of these methods is that they are flexible enough to follow *some* social phenomena along each of their folds.

No research method offers a panoptic vision of collective existence and quali-quantitative methods are no exception. Digital methods can only offer an *oligoptic* vision of society (Latour and Hermant, 1988), exactly as traditional methods. However, for the first time in the history of the social sciences, this vision will at least be continuous spanning from the tiniest micro-interaction to the largest macro-structure.

This is why we opened this chapter with the story of *Menenius Agrippa* and the technology of *scintigraphy*. The advantage of scintigraphy is not that it allows observing any biological phenomenon, but that it can follow some of them in a continuous fashion, without jumps or breaks. Social existence is not divided in two levels, as traditional methods led us to believe. Micro-interactions and macro-structures are only two different ways of looking at the same collective canvas, like the warp and weft of the social fabric. There - in the unity generated by the multiplication of differences, in the stability produced by the accumulation of mutations, in the harmony hatching from controversies, in the equilibrium relying on thousands on fractures - lie the marvel of communal existence. Qualitative and quantitative methods have too long hid this spectacle from us. Digital methods will open our eyes.

Bibliography

- Bijker, W. E., and Law, J. (eds.) (1992), *Shaping Technology, Building Society: Studies in Sociotechnical Change*. Cambridge, Mass.: MIT Press.
- Bloor, D. (1976), *Knowledge and Social Imagery*. Chicago: University of Chicago Press.
- Callon, M., Lascoumes, P., and Barthe, Y. (2009), *Acting in an Uncertain World: An Essay on Technical Democracy*. Cambridge, Mass.: MIT Press.
- Collins, H., and Pinch, T. (1993), *The Golem: What Everyone Should Know about Science*. Cambridge, U.K.: Cambridge University Press.
- Desrosières, A. (1998), *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, Mass.: Harvard University Press.
- Eisenstein, E. (1979), *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early Modern Europe*. Cambridge, U.K.: Cambridge University Press.
- Johnson, S. (2001), *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*. New York: Touchstone.
- Kuhn, T. (1962), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Latour, B. (1988), *Science in Action: How To Follow Scientists and Engineers through Society*. Boston: Harvard University Press.
- Latour, B., and Hermant, E. (1988), *Paris ville invisible*. Paris: La Découverte.
- Latour, B. (2005), *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford: Oxford University Press.
- Lazer, D. et al. (2009), "Computational Social Science," in *Science*, Vol. 323: 721-723.
- Rogers, R. (2004), *Information Politics on the Web*. Cambridge, Mass.: MIT Press.
- Venturini, T. (2010a), "Diving in Magma: How To Explore Controversies with Actor-Network Theory," in *Public Understanding of Science*, Vol. 19, No. 3, 258-273.
- Venturini, T. (2010b), "Building on Faults: How To Represent Controversies with Digital Methods," in *Public Understanding of Science* (under review).
- Weisman, A. (2007), *The World Without Us*. New York: St. Martin's Press.