

Fill in the Gap. A New Alliance for Social and Natural Sciences

T. Venturini¹, P. Jensen², B. Latour¹

¹ médialab, Sciences Po, Paris.

² Institut Rhônealpin des Systèmes Complexes (IXXI) and Laboratoire de physique, UMR 5672 CNRS, ENS de Lyon

*Correspondence to tommaso.venturini@sciencespo.fr

In the last few years, electronic media brought a revolution in the traceability of social phenomena. As particles in a bubble chamber, social trajectories leave digital trails that can be analyzed to gain a deeper understanding of collective life. To make sense of these traces a renewed collaboration between social and natural scientists is needed. In this paper, we claim that current research strategies based on micro-macro models are unfit to unfold the complexity of collective existence and that the priority should instead be the development of new formal tools to exploit the richness of digital data.

In the last decade, the spread of digital technologies has flooded the study of social phenomena with more data than ever dreamed of. Like rural countries pushed to a sudden industrialization by the global economy, the social sciences entered their age of abundance abruptly and with little preparation. Seeking help to handle their sudden fortune, social scientists turned to their colleagues in the natural sciences. Alas, the persistence of old habits has so far prevented researchers from taking full advantage of this alliance.

So far, the collaboration between natural and social scientists has invested vast efforts on modeling the emergence of collective phenomena from individual interactions, particularly through agent-based models (Cho 2009; Castellano *et al.* 2009; see also the very popular web site <http://www2.econ.iastate.edu/tesfatsi/ace.htm> and the special issues of *PNAS*, 99[suppl. 3] and *American Journal of Sociology*, 110[4]). Applied to urban segregation (Schelling 1971), business locations (Krugman 1996), epidemics and cultural trends (Bouchaud, 2012) and many other collective phenomena (cfr. most articles published in this journal), these models may range from simple ‘toy simulations’ to sophisticated systems based on empirical data (for example Broeck 2011). Most of them, however, partake of the same conceptual approach in which individuals are taken as discrete and interchangeable ‘social atoms’ (Buchanan 2007) out of which social structures emerge as macroscopic characteristics (viscosity, solidity...) emerge from atomic interactions in statistical physics (Bandini *et al.* 2009).

In the past, this strategy found its rationale in the methodological discontinuity that characterized the social sciences. As long as the divide between qualitative and quantitative methods cast a blind spot between situated observations and aggregated indicators, little was the chance to *reconstruct empirically the continuity of collective existence*. In the shortage of data on the dynamics of folding and unfolding that determine both local interactions and global structures, simulations were a sensible way to investigate the so-called micro-macro link (Archer 1995; Giddens 1984). When the access to empirical data is too expensive, models allow at least to examine the logical consequences of the theoretical assumptions made at the local level. In the

artificial worlds created by micro-macro models, researchers can play with the actors' features or the interactions' rules to produce a variety of patterns observed at the global level. As one famous economist puts it, simulations "provide fully articulated, artificial economic systems that can serve as laboratories in which policies that would be prohibitively expensive to experiment with in actual economics can be tested out at much lower cost." (Lucas, 1981, p. 271).

Micro-macro models, however, have serious methodological and political problems. From a methodological viewpoint, most simulations work only at the price of simplifying the properties of micro-agents, the rules of interaction and the nature of macro-structures so that they conveniently fit each other. A little bit like Descartes' followers who explained the acidity of lemons by postulating the existence of 'lemon atoms' with tiny pricking needles (Hoddeson 1992). In the absence of empirical confirmation, social models tend to rely exclusively on internal coherence and are particularly vulnerable to the 'confirmatory bias' (Nickerson 1998 and Rabin & Schrag 1999). "We opt for deductive verification of our claims in order to achieve clarity, rigour and certainty. But to get it we have tied the results to very special circumstances; the problem is how to validate them outside" (Cartwright 1999, p. 229).

From a political viewpoint, micro-macro models assume by construction that agents at the local level are incapable to understand and control the phenomena at the global level. Only the modelers can observe collective phenomena. Most simulations assume that only "human beings external to those involved – scholars and public officials – are able to analyze the situation, ascertain why counterproductive outcomes are reached, and posit what changes in the rules-in-use will enable participants to improve outcomes. Then, external officials are expected to impose an optimal set of rules on those individuals involved. It is assumed that the momentum for change must come from outside the situation rather than from the self-reflection and creativity of those within a situation to restructure their own patterns of interaction" (Ostrom, 2010, p. 648). Ironically, a supposedly "bottom-up" approach (Epstein & Axtell, 1996) leads to "top-down" social politics!

The methodological and political difficulties of micro-macro models have been highlighted, for example, in the case of the so-called 'tragedy of the commons' (Hardin, 1968). In these situations, personal interest pushes actors to overuse a shared good (a common pasture for example) to the detriment of the community. Assuming the existence of atomic agents each with its individually defined interest (as required by game theory), most of this literature cannot but confirm the overexploitation of the common good. However, empirical work conducted by Elinor Ostrom (1990) has shown that human cooperation can often (but not always) find arrangements to successfully manage the commons. Social simulations fail to obtain these arrangements, because they disregard the subtle mechanisms that govern the establishment of trust needed for cooperation. As shown by Ostrom, common standards, family ties, reputation and even facial expressions are crucial to obtain social cooperation. Impossible to anticipate through conceptual models, these factors can only be revealed by empirical observation.

Ethnographic studies such as Ostrom's, however, can expose the failures of micro-macro models, but not replace them. Relying on direct observation, qualitative researches can examine a number of situated exchanges but cannot follow how thousands of such interactions fold in the fabric of collective trust. And this is where the digital deluge may turn into a blessing.

The most interesting feature of digital media is that everything that they mediate becomes potentially traceable and often actually traced (Rogers, 2013). Such traceability creates data that are as rich/thick as those collected by ethnographic techniques but covering much larger populations. Everyday new public and private archives are swallowed by computers memories, economic transactions migrate online, social networks root in the Web and the more this happens, the more traces become available on the collective dynamics that used to be hidden by the quali-quantitative divide (Latour et al. 2012). Since digital traces brought it to light, the continuum between local exchanges and global trends revealed much more interesting and rich than its extremes. Social existence does not jump from micro to macro and neither should social sciences. Structures do not pop up from interactions as rabbits from an illusionist's hat. They are constructed and maintained by the relentless work of connecting and folding that (sometimes) leads to stronger, wider and longer lasting 'associations' (Latour 2005), as exemplified by studies of memes spreading (Leskovec 2009, memetracker.org); fame in the blogosphere (Cardon et al. 2011); migrant communities (Social Science Information, special issue 51:4, 2012 and e-diasporas.fr); manga styles (Manovitch 2012); scientific paradigms (Chavalarias & Cointent 2009 and Börner 2010); open source collaboration (B. Heller & E. Marschner 2011), international negotiations (Venturini et al. 2014); lexical trends in history of literature (Michel 2011); law amendments (lafabriquedelaloi.fr); Wikipedia controversies (Borra et al. 2014 and contropedia.net).

An example of how digital data can renew modeling can be found in opinion dynamics, one of the most popular subjects of social simulations. Several hundreds of papers have been published on this topic and some of them are among the most cited JASSS references. Their proponents argue that, despite their simplicity, these models have managed to produce a surprising variety of patterns when changing the details of the models or their parameters. However, these simulations have not yet succeeded to connect in any significant way to real-world behaviors (Sobkowicz 2009). These models use "thin concepts" which are homonymous with everyday concepts ('opinion'), but "little of their behaviour from the real world is imported into the model" (Cartwright 1999). No surprise that "the impact of JASSS is higher in computer sciences, physics and ecology than it is in the social sciences, even though JASSS-indexed articles tend to be more concerned with social science-related topics" (Squazzoni and Casnici 2013).

Thanks to the growing traceability of online discussions, it is now possible to describe in more realistic ways how people change opinion. For example, instead of assuming that each node influences all of its neighbors with the same probability, one would learn that the likelihood to adopt an opinion is higher "when participants receive social reinforcement from multiple neighbors in the social network" (Centola 2010). This leads to a "farther and faster spread across clustered-lattice networks than across corresponding random networks", contrary to what most simulations suggested. In another online experiment tracing in detail the spread diet diaries, Centola (2011) showed that neighbors that are similar to oneself are much more influential, whereas most simulations assumed that each tie has equal weight in the diffusion process.

Simulating the emergence of macro structures from micro interactions has never been an optimal research strategy, neither methodologically nor politically. Its main justification – the possibility to bypass the supposed micro/macro divide – lost much of its interest since the advent of digital media. Empirical studies show that, contrarily to what most social simulations assume,

collective action does not originate at the micro level of individual atoms and does not end up in a macro level of stable structures. Instead, actions distribute in intricate and heterogeneous networks than fold and deploy creating differences but not discontinuities.

Digital traceability has transformed the context of the collaboration between social and natural scientists and its agenda should change accordingly. The problem is not anymore to simulate data that would be too expensive to collect. The problem is to handle an avalanche of traces whose magnitude and diversity is unprecedented for the social sciences. For the first time in their history, social scientists have continuous information about their objects and the last thing they need are models that break them in micro/macro oppositions.

This does not mean, of course, that the modeling tradition of natural sciences ceases to be relevant for the study of collective life. Quite the contrary! Such experience is crucial to develop the new methods necessary to handle larger and more diverse datasets. At the beginning of the 19th century natural and social scientists developed together a new discipline, “statistics”, that helped them to interpret the new data available at that time (Hacking, 1990; Desrosières 2002). Today, the advent of digital data poses a similar challenge and calls for a similar alliance. Micro-macro models have run their course. The time is now to develop the formal techniques necessary to unfold the origami of collective existence and this should be the aim of the renewed alliance between the social and natural sciences. For the next few years, at least, efforts should be shifted from simulating to mapping, from simple explanations to complex observations.

References:

AMERICAN JOURNAL OF SOCIOLOGY (2005), special issue on agent-based modeling 110(4).

ARCHER, M.S. (1995). *Realist Social Theory: The Morphogenetic Approach*. Cambridge: Cambridge University Press.

BANDINI, S., MANZONI, S. AND VIZZARI, G. (2009). Agent Based Modeling and Simulation: An Informatics Perspective. *Journal of Artificial Societies and Social Simulation* 12(4)4.

BÖRNER, K. (2010). *Atlas of Science*. Cambridge MA: MIT Press.

BORRA, E., WELTEVREDE, E., CIUCCARELLI, P., KALTENBRUNNER, A., LANIADO, D., MAGNI, G., MAURI, M., ROGERS, R., VENTURINI, T. (2014). Contropedia - the analysis and visualization of controversies in Wikipedia articles. *OpenSym '14: The International Symposium on Open Collaboration Proceedings*.

BOUCHAUD, J.P. (2012). Crises and collective socio-economic phenomena: simple models and challenges. *arXiv:1209.0453v1* (accessed Sep 6th, 2012).

BROECK, W.V. ET AL (2011). The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infectious Diseases*, 11(37).

BUCHANAN, M. (2007). *The Social Atom*. London: Bloomsbury.

CARDON, D. ET AL (2011). Two Paths of Glory. *Proceedings AAAI ICWSM'11 Conference Weblogs and Social Media*.

CARTWRIGHT, N. (1999) The Vanity of Rigour in Economics: Theoretical Models and

Galilean Experiments, available online at :

http://www2.lse.ac.uk/CPNSS/pdf/DP_withCoverPages/DP43/DP43F-99-C.pdf

CASTELLANO C. ET AL. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, 81, 591-646.

CENTOLA, D (2010). The Spread of Behavior in an Online Social Network Experiment, *Science*, 329(5996), 1194-1197.

CENTOLA, D (2011). An Experimental Study of Homophily in the Adoption of Health Behavior, *Science*, 334(6060), 1269-1272.

CHAVALARIAS, D., COINTET, J.-P. (2009). The Reconstruction of Science Phylogeny. *arXiv:0904.3154* (accessed Sep 2th, 2012).

CHO, A. (2009). Ourselves and Our Interactions: The Ultimate Physics Problem?. *Science* 325(406).

DESROSIÈRES, A. (2002) *The Politics of Large Numbers*, Harvard Univ Press.

EPSTEIN, J.M., AXTELL, R.L. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge MA: MIT Press.

GIDDENS, A. (1984). *Constitution of society: Outline of the Theory of Structuration*. Berkeley: University of California Press.

HACKING, I. (1990) *The Taming of Chance*, Cambridge University Press.

HARDIN, G. (1968). "The Tragedy of the Commons" *Science*, 162(3859), 1243-48.

HELLER, B., MARSCHNER, E. (2011). Visualizing Collaboration and Influence in the Open-Source Software Community". *Proceedings of the Conference on Mining Software Repositories*.

HODDESON, L. (Ed.) (1992), *Out of the Crystal Maze*. Oxford: Oxford University Press.

KRUGMAN, P. (1996). *The Self-Organizing Economy*. Cambridge MA: Blackwell.

LATOUR, B. (2005). *Reassembling the social*. Oxford: Oxford University Press.

LATOUR, B. ET AL. (2012). The Whole is Always Smaller Than Its Parts. A Digital Test of Gabriel Tarde's Monads. *British Journal of Sociology*, 63(4), 591-615.

LESKOVEC, J. ET AL (2009). Meme-tracking and the dynamics of the news cycle. *Proceedings 15th ACM SIGKDD Conf on Knowledge discovery and data mining*.

LUCAS, R. E., (1981). *Studies in Business-Cycle Theory*, Cambridge MA: The MIT Press.

MANOVITCH L. (2012). How to Compare One Million of Images. In Berry, D. (Ed.), *Understanding Digital Humanities*. Houndmills and New York: Palgrave.

MICHEL J.B. ET AL (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331(6014), 176-82.

NICKERSON, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175-220.

OSTROM, E. (2010). Beyond Markets and States: Polycentric Governance of Complex Economic Systems. *American Economic Review*, 100(3), 641-672.

OSTROM, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge; Cambridge University Press.

- PNAS (2002). Adaptive Agents, Intelligence, and Emergent Human Organization: Capturing Complexity through Agent-Based Modeling (special issue). *PNAS*, 99(suppl. 3).
- RABIN, M., & SCHRAG, J. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114(1), 37–82.
- ROGERS, R. (2013). *Digital Methods*. Cambridge MA: MIT Press.
- SHELLING, T.C. (1971). Dynamic Models of Segregation. *Journal of Mathematical Sociology*, 1(143).
- SOCIAL SCIENCE INFORMATION (2012). Diasporas on the Web (special issue). *Social Science Information*, 51(4).
- SOBKOWICZ, P. (2009). Modelling Opinion Formation with Physics Tools: Call for Closer Link with Reality. *Journal of Artificial Societies and Social Simulation* 12(1)11.
<http://jasss.soc.surrey.ac.uk/12/1/11.html>
- SQUAZZONI, F., CASNICI, N. (2013). Is Social Simulation a Social Science Outstation? A Bibliometric Analysis of the Impact of JASSS. *Journal of Artificial Societies and Social Simulation* 16(1), 10.
- VENTURINI, T., BAYA-LAFFITE, N., COINTET, J., GRAY, I., ZABBAN, V., & DE PRYCK, K. (2014). Three Maps and Three Misunderstandings : A Digital Mapping of Climate Diplomacy. *Big Data and Society*, 1(1).