Workshop
The Social Complexity of Big Data: On the Limits of Naturalism in a Digital World
Ca' Foscari, Venice, Italy. January 29-30, 2018.

Over the course of this two-day workshop, we aim to facilitate an intellectually open and curious debate on the implications of Big Data for thinking about complexity in the social world. Big Data seems to have generated a pressing need for meta-theory among those with method – and for method for those with meta-theory. We therefore invite a diverse group of prominent scholars to revisit old questions of social ontology in the light of a new reality. What are the possibilities of social scientific knowledge and the limits of naturalism in our new digital world?

Workshop Theme
The emergence of Big Data has given social scientists unprecedented access to previously unimaginable data – traces of the lives, dreams, and feelings of hundreds of millions of people. This brings immense promises for social scientific work, as the "data deluge … is leading us to an ever-greater understanding of life on Earth and the Universe beyond … [it may] transform the process of scientific discovery. The more data there is the more discoveries can be made" [1]. Some have even pointed to an approaching "end of theory" [2], or a "fourth paradigm" for science, as new algorithmic, computational, and analytical tools produce "gold" from this data resource [3,4].

But for the traditional approaches to the social science, whether qualitative or quantitative, these hopes have often run into a wall of methodological and epistemological problems, resulting in the competing view that "the glittering promise of online data abundance too often proves to be fool's gold" [5]. Because Big Data is not only different in quantity, but also in quality, and it seems that the new shapes of data do not fit into the holes of our old social theory [6]. This has resulted in bringing fundamental questions to the fore: the questioning of epistemological assumptions, debates about the validity of disciplinary divides, critique of methodological monism, and the rejection of long-trusted simplifications, as the data deluge stirs long-submerged epistemological questions to float to the surface [7]. These epistemological challenges have been described as Big Data clashing with traditional understandings of quantitative social science, by unveiling some of the true relational complexity of the social world.

One way of understanding these tensions is the view that traditional social scientific data consist of static, representative samples of a population, rich in detailed attribute data, but poor in relational information; it is not a coincidence that such data lends itself so well to aggregation and variable-based analysis – it was specifically constructed to do so. Big Data, on the other hand, consists of traces of social processes rather than data constructed for scientific consumption, and does not fit as neatly into the epistemological box of the variable-based paradigm. Big Data does not hide the intricate relational complexity of the social, nor the mass-interaction that play part in social pattern [8]. It tends to be in many ways the opposite of old data: dynamic, often unrepresentative of anything but itself, rich on relations, but poor on attribute information. This makes the

static variable- and aggregation-based approaches chafe uncomfortably against the interactional frenzy of Big Data.

While the traditional social scientific approaches have stumbled, disciplines with their roots in the natural sciences have stepped forth to meet the tide of digital data. Methods and theories developed under the umbrella of "complexity science" [9,10,11], with their methodological toolkit filled from decades of focus precisely on relationality and mass-interaction, have been particularly successful in formulating a response. Research on *emergence* and *self-organization* and the large strand of theories and ideas associated thereto, has been deeply influential in the formation of disciplines like *Computational Social Science*.

This renegotiation of roles between the natural and social sciences seems to be leading to a renewed *naturalism* [12], exemplified by Lev Manovich's [13] view that "Digital is what gave culture the scale of physics, chemistry or neuroscience. Now we have enough data and fast enough computers to actually study the 'physics' of culture". Society is increasingly being subsumed under the metatheory of particles and flows, and understood under analogues such as "avalanches and granular flows, flocks of birds and fish, networks of interaction in neurology, cell biology and technology" [14].

But at the same time, the amorphous shape of Big Data does not seem to fit neatly into this theoretical hole either. Big Data appears, paradoxically, to emphasize the distinguishing ontological features of society, what Dahrendorf called its "vexatious fact" [15]: aspects of society that the complexity-inspired analogues tend to leave out. The human capacity of reflexivity of its own conditions, creating causal thickets [16] that cut through the timescale separation that lies as a foundational cornerstone assumption of many scientific methods – what Herbert Simon [17] called "near-decomposability" and e.g. Roy Bhaskar [12] referred to through the "closed"/"open" dichotomy. People shape society, and society shapes us, as we, individually and collectively, go about changing it or maintaining it. The dematerialization associated to Big Data means that it also constitutes an important step in late modernity's constant march toward increasing liquidity and acceleration [18], what Karpf [5] terms "web time" essentially means that the causal thickets of the social world grow thicker. With Big Data, they may indeed have transitioned from the realm of abstract theory to hard-to-ignore practical problems for research, expressed through larger time-scale overlaps or increasing "ontological uncertainty" [19]. Through this specific lens, culture is not becoming more like physics; it is becoming *less* like physics.

This opens the question of how to study a social world whose conditions for study are simultaneously - in terms of relational complexity and data availability - *more similar* to the ontological assumptions of the natural sciences than ever before, while at the same time - in terms of the "epistemological and ontological conditions of the social sciences" [12] – it has never been *more different*. It seems that the methodological and epistemological demands of Big Data are somehow in contradiction with its very ontological implications.

Big Data has therefore stirred up both sides of the entrenched conflict between, as David Byrne [20] puts it, "the gang [who] can count but don't know what they are counting [and the gang who] can't count, won't count, and assert that counting is a vile and perverse activity which ought not to be allowed". For the former, the challenges come in the form of an increasingly frail connection between the counting and reality of the social world. For the latter, an era in which quantitative approaches are less and less reliant on studying causality through variable-variance approaches, and in which social experiments have become ubiquitous, seems to prod a revisiting of the conclusions of authors like Roy Bhaskar regarding the possibilities of accessing causality in the social world.

Briefly put, the need for meta-theory among those with method, and method for those with meta-theory, seem more acute than ever.

This workshop aims at bringing together scholars from different backgrounds in a constructive and mutually curious setting, with the aim of allowing complexity/computer science engage seriously with philosophy of social systems, and vice versa. What are the possibilities of social scientific knowledge and the limits of naturalism in a digital world? Could an integration between an explicit social ontology with methods and theories from complexity science provide the foundation for "the development of a situated, reflexive and contextually nuanced epistemology" [7] capable of panning for gold in the muddy waters of the digital data deluge?

Preliminary Activities and Timeline

Day 1
09:00 – 09:10 – Welcome to Ca'Foscari
09:10 – 09:30 – Introduction to the workshop
09:30 – 10:15 – Speaker 1
10:15 – 10:45 – COFFEE
10:45 – 11:30 – Speaker 2
11:30 – 12:15 – Speaker 3
12:15 – 14:00 – Lunch
14:00 – 14:45 – Speaker 4
14:45 – 15:30 – Speaker 5
15:30 – 16:00 – Coffee
16:00 – 16:45 – Speaker 6

Day 2
09:00 – 09:45 – Speaker 7
09:45 – 10:30 – Speaker 8
10:30 – 11:00 – Coffee
11:00 – 13:00 – Panel discussion
13:00 – 14:00 – Lunch
14:00 – 16:00 – Panel discussion

References
1. Rosling, H. (2010). The joy of stats. Retrieved from http://www.gapminder.org/videos/the-joy-of-stats/
2. Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. Wired magazine, 16(7), 16-07.
3. Tansley, S., & Tolle, K. M. (2009). The fourth paradigm: data-intensive scientific discovery (Vol. 1). T. Hey (Ed.). Redmond, WA: Microsoft research.
4. Bell, G., Hey, T., & Szalay, A. (2009). Beyond the data deluge. Science, 323(5919), 1297-1298.
5. Karpf, D. (2012). Social science research methods in Internet time. Information, Communication & Society, 15(5), 639-661.
6. boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Information, communication & society, 15(5), 662-679. 7. Kitchin 2014
8. Sawyer, R. K. (2005). Social emergence: Societies as complex systems. Cambridge University Press.
9. Urry, J. (2003). Global complexity (p. 107). Cambridge: Polity.
10. Walby, S. (2007). Complexity theory, systems theory, and multiple intersecting social inequalities. Philosophy of the social sciences, 37(4), 449-470.
11. Mitchell, M. (2009). Complexity: A guided tour. Oxford University Press.
12. Bhaskar, R. (1978). On the possibility of social scientific knowledge and the limits of naturalism. Journal for the Theory of Social Behaviour, 8(1), 1-28.13. Lev Manovich 2016
14. Ball, P., 2012. Why society is a complex matter: Meeting twenty-first century
challenges with a new kind of science. Springer Science & Business Media.
15. Archer, M. S. (2013). Morphogenic society: Self-government and self-organization as misleading metaphors. In Social morphogenesis (pp. 145-164). Springer Netherlands.
16. Wimsatt, W. C. (1994). The ontology of complex systems: levels of organization, perspectives, and causal thickets. Canadian Journal of Philosophy, 24(sup1), 207-274.
17. Simon, H. A. (2002). Near decomposability and the speed of evolution. Industrial and corporate change, 11(3), 587-599.
18. Hayles, N. K. (2008). How we became posthuman: Virtual bodies in cybernetics, literature, and informatics. University of Chicago Press.
19. Lane, D. A., & Maxfield, R. R. (2005). Ontological uncertainty and innovation. Journal of evolutionary economics, 15(1), 3-50.
20. Byrne, D. (2002). Interpreting quantitative data. Sage.