



# Editorial: At the Crossroads: Lessons and Challenges in Computational Social Science

Javier Borge-Holthoefer<sup>1,2,3\*</sup>, Yamir Moreno<sup>3,4,5</sup> and Taha Yasseri<sup>6,7</sup>

<sup>1</sup> Complex Systems Group (CoSIN3), Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Spain, <sup>2</sup> Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar, <sup>3</sup> Institute of Biocomputation and Physics of Complex Systems, Universidad de Zaragoza, Zaragoza, Spain, <sup>4</sup> Department of Theoretical Physics, Faculty of Sciences, Universidad de Zaragoza, Zaragoza, Spain, <sup>5</sup> Institute for Scientific Interchange, Torino, Italy, <sup>6</sup> Oxford Internet Institute, University of Oxford, Oxford, UK, <sup>7</sup> Alan Turing Institute, London, UK

**Keywords:** computational social science, simulation, models, big data, complex systems

## The Editorial on the Research Topic

### At the Crossroads: Lessons and Challenges in Computational Social Science

The interest of physicists in economic and social questions is not new: during the last decades, we have witnessed the emergence of what is formally called nowadays sociophysics [1] and econophysics [2] that can be grouped into the common term “Interdisciplinary Physics” along with biophysics, medical physics, agrophysics, etc. With tools borrowed from statistical physics and complexity science, among others, these areas of study have already made important contributions to our understanding of how humans organize and interact in our modern society. Large scale data analyses, agent-based modeling and numerical simulations, and finally mathematical modeling, have led to the discovery of new (universal) patterns and their quantitative description in socio-economic systems.

At the turn of the century, however, it was clear that huge challenges—and new opportunities—lied ahead: the digital communication technologies, and their associated data deluge, began to nurture those models with empirical significance. Only a decade later, the advent of the Web 2.0, the Internet of Things and a general adoption of mobile technologies have convinced researchers that theories can be mapped to real scenarios and put into empirical test, closing in this way the experiment-theory cycle in the best tradition of physics.

We are nowadays at a crossroads, at which different approaches converge. We name such crossroads computational social science (CSS) : a new discipline that can offer abstracted (simplified, idealized) models and methods (mainly from statistical physics), large storage, algorithms and computational power (computer and data science), and a set of social hypotheses together with a conceptual framework for the results to be interpreted (Social Science) [3–5]. Despite its youth, the field is developing rapidly in terms of contents (articles, books, etc.), but also institutionally—either under the form of labs, institutes, and academic programs; or as consolidated events and scientific gatherings.

This “work-in-progress” spirit is reflected as well in this volume: the call was launched in late 2014 and 10 articles were eventually accepted and published, including reviews—a look behind—, one methods paper, and six original contributions—a look ahead—introducing a broad range of research, from models with a strong analytical flavor to data-driven problems.

As mentioned above, each new research line in CSS starts with analysing a sizable dataset, containing transactional data or user generated content on social web (also called Big Data), The availability of data however poses new methodological challenges. Among them are statistical

## OPEN ACCESS

### Edited and reviewed by:

Alex Hansen,  
Norwegian University of Science and  
Technology, Norway

### \*Correspondence:

Javier Borge-Holthoefer  
jborgeh@uoc.edu

### Specialty section:

This article was submitted to  
Interdisciplinary Physics,  
a section of the journal  
Frontiers in Physics

**Received:** 29 July 2016

**Accepted:** 16 August 2016

**Published:** 29 August 2016

### Citation:

Borge-Holthoefer J, Moreno Y and  
Yasseri T (2016) Editorial: At the  
Crossroads: Lessons and Challenges  
in Computational Social Science.  
*Front. Phys.* 4:37.  
doi: 10.3389/fphy.2016.00037

analysis and how the methods and conventions that have been used in social sciences to analyse small datasets collected via questionnaires and interviews, can be now used to analyse data generated by millions of people. Vidgen and Yasserli address this challenge and in particular discuss the confusion and (mis)uses around the highly-popular  $p$ -values. They call for a more careful use of statistical tests and show few directions for improvement. Apart from the methodological challenges, as Holme and Liljeros note in their review on mechanistic models in computational social science, “Quantitative social science is not only about regression analysis or, in general, data inference.” Computer simulations, whose history is reviewed by Holme and Liljeros, are one the main connecting bridges between empirical observations and theoretical models.

Adopted from physics, the next step in the “scientific method” of CSS is experimentation. Experiments might not only be used to validate the generated theories, but also to yield new observations that might eventually lead to new lines of research themselves. However, experimentation in social content and with human subjects is nothing similar to physics experiments. Recruitment, representativeness, privacy issues, and ethical challenges are very central to social experimentation. Sagarra et al., focus their attention on citizen science and offer a methodological guideline for experimentation outside the laboratory.

Spreading phenomena make one of the key topics in CSS. With application in innovation [6], political change [7], epidemiology [8], etc., it is important to understand how “things” navigate through social networks. To further develop the extensive literature of social contagion, O’Sullivan et al. extend the modeling of complex contagion in the context of clustered networks—of upmost relevance in the social context. Solé-Ribalta et al., on their side, focus on communities on multilayer networks [9] and develop new mathematics for information transfer. These two articles make an excellent case of how the insights from complex systems and statistical physics can and do play a role in CSS, offering solid foundations to the methods and insights developed within.

One of the main advantages of CSS over classic social sciences, is the possibility to perform temporal analysis and come up with dynamical models. Most of the datasets under study in CSS contain timestamps, allowing for fine-grained analysis of interactions over time. Sanli and Lambiotte provide an original approach to online communication based on complex time series—rather than on network structure— that emerge from user dynamics on social media. The analysis is performed on a set

of collected messages which correspond to an exceptional event, which is common practice in the field to study collective behavior [10]. Also Omodei et al. take this approach, analyzing a broad range of events (policy, culture, science) which they characterize as multiplex networks [9, 11]. And finally Aledavood et al. use mobile phone records to study diurnal patterns of human communication and provide a cohesive picture of regularities in communication patterns both at individual and society levels.

Large scale analysis of socially generated data is not limited to transactional records: huge amount of digital content is being produced on daily basis. In a novel work, Tanase et al. apply linguistic analysis to user reviews that they collected from the web and study social influence and its interplay with network topology, and how it affects users’ opinions.

Generalizing on opinion dynamics, there is no doubt that the Internet in general and social media in particular have changed the political environment and the way people engage in political activities [12]. At the same time, the digital footprint of online political activities provides a great opportunity to conduct political science studies at scale and in close-to real time, leading to the emergence of some sort of data-driven political science within CSS [13]. However, these opportunities come with challenges and shortcomings. Cihon and Yasserli take a critical point of view toward such studies and in particular discuss the “biases” in Twitter-based research on political collective action in a short “biased review.”

Computational Social Science emerges as a wide set of scientific opportunities, to tackle the fundamental features of social complexity—multidirectional connections, layer interdependences and interferences, accelerated diffusion, and so on [14]. The complex systems approach that underlies CSS is a key feature toward creating a truly interdisciplinary, non-compartmental science.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## ACKNOWLEDGMENTS

We thank Matjaž Perc for the great editorial help and all the reviewers who assisted us in preparing the manuscripts of this Research Topic.

## REFERENCES

- Sen P, Chakrabarti BK. *Sociophysics: An Introduction*. Oxford, UK: Oxford University Press (2013).
- Mantegna RN, Stanley HE. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge, UK: Cambridge University Press (1999). doi: 10.1017/CBO9780511755767
- Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, et al. Life in the network: the coming age of computational social science. *Science* (2009) 323:721. doi: 10.1126/science.1167742
- Watts DJ. Computational social science: exciting progress and future directions. *Bridge Front. Eng.* (2013) 43:5–10. Available online at: <https://www.nae.edu/Publications/Bridge/106112/106118.aspx>
- Mann A. Core concepts: computational social science. *Proc Natl Acad Sci USA* (2016) 113:468–70. doi: 10.1073/pnas.1524881113
- Pentland A. *Social Physics: How Good Ideas Spread—the Lessons from a New Science*. New York, NY: Penguin (2014).
- Galam S. *Sociophysics: A Physicist’s Modeling of Psycho-Political Phenomena*. New York, NY: Springer Science & Business Media (2012).
- Newman ME. Spread of epidemic disease on networks. *Phys Rev E* (2002) 66:016128. doi: 10.1103/PhysRevE.66.016128

9. Kivela M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. *J Comp Netw.* (2014) 2:203–71. doi: 10.1093/comnet/cnu016
10. Borge-Holthoefer J, Perra N, Gonçalves B, González-Bailón S, Arenas A, Moreno Y, et al. The dynamics of information-driven coordination phenomena: a transfer entropy analysis. *Sci Adv.* (2016) 2:e1501158. doi: 10.1126/sciadv.1501158
11. De Domenico M, Solé-Ribalta A, Cozzo E, Kivela M, Moreno Y, Porter MA, et al. Mathematical formulation of multilayer networks. *Phys Rev X* (2013) 3:041022. doi: 10.1103/PhysRevX.3.041022
12. Margetts H, John P, Hale S, Yasseri T. *Political Turbulence: How Social Media Shape Collective Action*. Princeton, NJ: Princeton University Press (2015).
13. Weber I, Popescu AM, Pennacchiotti M. Data-driven political science. In: *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*. New York, NY: ACM (2013). p. 777–778. doi: 10.1145/2433396.2433498
14. Conte R, Gilbert N, Bonelli G, Cioffi-Revilla C, Deffuant G, Kertesz J, et al. Manifesto of computational social science. *Eur Phys J Spec Top.* (2012) 214:325–46. doi: 10.1140/epjst/e2012-01697-8

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Borge-Holthoefer, Moreno and Yasseri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.