Why Don't We Ever Talk About Education When We Talk About Computational Social Science?

Diliara Valeeva
University of Amsterdam, Nieuwe Achtergracht 166,
1018 WV, Amsterdam, The Netherlands
d.valeeva@uva.nl

Abstract

This paper discusses the challenges and opportunities of teaching computational social science (CSS) at the university level. With the growing demand for CSS expertise in industry, government, and non-profit sectors, there is a need to prepare a new generation of CSS practitioners. However, there is a lack of discussion in the field about how to teach CSS. The paper argues that CSS researchers need to start thinking about teaching and learning CSS as educators, not just as researchers. The paper also raises questions about what constitutes a CSS question and how to integrate open science principles and ethical considerations into CSS education. The paper concludes by suggesting that CSS need to be prepared to answer these questions and ally with other fields to provide comprehensive CSS education.

"Can you give us more examples of how computational social science can actually help us to learn something new about society?" I am teaching a class for first-year bachelor students in computational social science. It is November in Amsterdam, and it has not stopped raining since September, when the program was launched. A student in the back row raised their hand and asked this question in the middle of class.

Fair enough. As someone doing computational social science research, I have heard this question many times, and I have prepared a variety of answers. I usually start my response by mentioning a study on how personality attributes can be predicted from digital records (Kosinski, Stillwell, and Graepel 2013). Then I will continue by discussing what we can glean from analysing Twitter about how people protest (González-Bailón et al. 2011). Maybe I will talk about how online book purchases can predict one's political preference (Shi et al. 2017). And then I will finish up with some dating app research, just to engage everyone again (Dinh et al. 2022). Usually, these examples are enough to convince a computational social science skeptic that studies such as these can reveal something new about society. And that it would be much more difficult, even impossible, to achieve the same results by applying the classical social science toolbox. During the break, the student tells me they started a Sociology degree the previous year, but dropped it because it was *"lots of theory and not enough practice"*.

In this piece, I will discuss the following action point: we, as computational social science researchers, have to start thinking about computational social science education as soon as possible. I am convinced that, over time, we will find ourselves in classrooms such as this one more and more often, teaching bachelor-level and master-level courses to new generations of computational social science students. Only a fraction of them will become computational social science scholars, of course. Many of them will pursue careers in industry, government, or the non-profit sector. I argue for a shift towards discussing computational social science as educators, not only as researchers.

What unites the studies that I use to engage my students, is not necessarily the way they are conducted. It is not (only) about getting access to complex and large-scale datasets or applying cutting-edge methodologies. In my opinion, these studies are similar in the ways in which they state questions about society and how they seek answers to them.

And the real question here is, of course: what is a computational social science question? What kind of questions would be the *computational* social, and not just social science questions? Are we ready for raising a generation of computational social scientists that would predominantly ask *computational* social science questions?

The problem is that any empirical study that can be used as a good example of a computational social science research, often itself contains a number of questions and topics, interesting to various groups of stakeholders and applicable to a number of scientific fields. Salganik, when he talks about a research by Blumenstock, Cadamuro, and On (2015) on measuring wealth distribution in Rwanda using mobile phone data, formulates it nicely:

> What people see depends on their background. Many social scientists see a new measurement tool that can be used to test theories about economic development. Many data scientists see a cool new machine learning problem. Many business people see a powerful approach for unlocking value in the big data that they have already collected. Many privacy advocates see a scary reminder that we live in a time of mass surveillance. And finally, many policy makers see a way that new technology can help create a better world. In fact, this

> study is all of those things, and because it has this mix of characteristics, I see it as a window into the future of social research (Salganic, 2019: 1.1).

It does not make it easier that all of us, doing research on computational social science, have been trained as anyone and anything but computational social scientists. Obvious but true, these kinds of degrees just did not exist. We did sociology, computer science, physics, political science, mathematics, psychology, economics, biology. I could continue this list and probably name any existing discipline to describe all the microcosm of researchers doing computational social science work. At the same time, the field developed so rapidly that we can say with confidence that, by now, all the research infrastructure has been already established. There is a number of regular computational social science conferences, multiple workshops and summer schools, thousands of papers, the Journal of Computational Social Science, an increasing number of grants aimed at advancing computational social science research, and some of us identify as computational social science scholars, and proud of it.

The plurality of backgrounds within the field is not only a hindrance. It is a strength. A window into the future, indeed. But it does mean that anyone working in the field is holding together a variety of assumptions, perspectives, methodologies. I believe the computational social science classroom is an important space where we also get to define what the future of the field will look like. Not as top-down scholars, telling young people what to think or how to do their research, but figuring it out together with them, by continually asking the fundamental questions about our field. It is a space where we get to return, again and again, to the most essential questions about what we do, and why we do it.

Notably, there is very little discussion in our field about teaching computational social science. Education is only rarely discussed in a number of founding papers (to name just a few: Edelmann et al. 2020; Lazer et al. 2009, 2020). This is understandable, as some of these papers were written at the very naissance of the field. But the reality is that there are actually many university-level programs opening, mostly in Europe at the moment, that are ready to issue degrees specifically in computational social science. Most of them are master-level programs. Currently, there is only one bachelor program in computational social science: at the University of Amsterdam, which I am currently a part of.

The students in computational social science that I teach today are digital natives, they are highly privacy-aware, on top of identity politics, use encrypted tools to communicate, know everything about climate change, and are not afraid to act. They also study in highly international classrooms, and finished their high school degrees during the global pandemic. They live in extremely polarized societies, and see how artificial intelligence is advancing every day. Which of these disruptions should we embrace together, and which ones should we resist? Whom to ally with in learning how to conduct computational social science research? How to teach responsible and ethical ways of gathering, preparing, and analyzing data? How to integrate open science principles in learning computational social science? These are the questions we are grappling with in the classroom.

I believe we have to be ready to answer these fundamental questions when we enter classrooms with students who want to get a university degree specifically in computational social science. It is just a matter of time before more and more specialized computational social science educational programs will begin opening up at universities across the globe. And who is going to teach there? Us, computational social science researchers. We need coherent ways of thinking not only about our computational social science research, but also about computational social science education. Only if we start this discussion on the importance of education, teaching, and learning, we will be able to answer the most fundamental questions about the significance of computational social science–like the one that my student posed to me on a rainy November day.

Disclaimer

Opinions expressed in this piece are my own and do not express the views or opinions of my employer, my colleagues, or my students.

References

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264), 1073-1076.

Dinh, R., Gildersleve, P., Blex, C., & Yasseri, T. (2022). Computational courtship understanding the evolution of online dating through large-scale data analysis. *Journal of Computational Social Science*, 5(1), 401-426.

Edelmann, A., Wolff, T., Montagne, D., & Bail, C. A. (2020). Computational social science and sociology. *Annual Review of Sociology*, 46, 61-81.

González-Bailón, S., Borge-Holthoefer, J., Rivero, A., & Moreno, Y. (2011). The dynamics of protest recruitment through an online network. *Scientific Reports*, 1(1), 1-7.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802-5805.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... & Van Alstyne, M. (2009). Computational social science. *Science*, 323(5915), 721-723.

Lazer, D. M., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., ... & Wagner, C. (2020). Computational social science: Obstacles and opportunities. *Science*, 369(6507), 1060-1062.

Salganik, M. J. (2019). *Bit by bit: Social research in the digital age*. Princeton University Press.

Shi, F., Shi, Y., Dokshin, F. A., Evans, J. A., & Macy, M. W. (2017). Millions of online book co-purchases reveal partisan differences in the consumption of science. *Nature Human Behaviour*, 1(4), 0079.