DARIAH-EU Virtual Annual Event 2020

Towards a Global Study of Algorithms and Algorithmization

POIROUX Jérémie
Phd student in sociology, computational social science team, Centre Marc Bloch Berlin - poiroux@cmb.hu-berlin.de

Aims & scope. Over the last few years, a research object has been attracting the attention of quite a number of members of the scientific community of digital humanities and sociology: the study of algorithms, and more precisely of their impact on society. Examples include the study of the impact of work distribution algorithms (Uber), recidivism prediction algorithms (COMPAS) or recommendation algorithms (Youtube or Spotify). Both quantitative and qualitative methods are used for these studies. At the same time, we can find an increasing integration of the digital humanities in these studies, both to analyse the data used by the algorithms and their results.

However, an insufficient number of studies in this field have focused on the fabrication of these algorithms, either on the way engineers and more widely organizations design these programs that have the impacts that we now know a little better (their non-neutrality or discrimination to name only two).
On one hand, sociologists have gone to meet engineers to study their way of working: this is "humanities of the digital" [1]. On the other hand, computer scientists have conducted experiments to open the "black box" of algorithms, notably through reverse engineering techniques: this is "numerical humanities".

In order to overcome this dichotomy at the heart of the digital humanities, it appears important to intertwine these two mentioned aspects. This shall illustrate that fears about qualitative studies being supplanted by quantitative studies are not well-founded [2].

Between 2017 and 2018, we conducted a sociological study by interviewing thirty engineers [3] from about the same number of different companies on the manufacture of recommendation algorithms. For the purpose of this poster, we partially reproduce the chapter on "transparency and intelligibility of the algorithm".

Algorithmic transparency currently appears to be a major normative concern, whether in the academic literature, the media or civil society [4]. While there seems to be a consensus on its necessary and positive aspects, especially in the context of algorithms with high social impact, the details and scope of algorithmic transparency and the trade-offs it implies [5] remain a largely open question.

From the developers' perspective we interviewed, transparency appears as an overhanging "ideal", which mainly concerns the way the algorithm is designed. They frequently see it as a constraint, sometimes implicit, which does not allow to develop a recommendation algorithm as one develops a classical software. Its very meaning is questioned: engineers thus ask first, "for whom", and "how", because transparency does not imply that the code is open and accessible to all. The obstacles put forward are mainly of three kinds.
Out of reach for the user. Many engineers believe that users are either unwilling or unable to understand how the algorithm works. Many point out that an algorithm is primarily a sequence of mathematical operations that are difficult to make understandable and that users may not want to know.
Out of reach for the engineers themselves. For engineers, transparency is also a question of cost and time deployed, particularly because the algorithm is constantly evolving or its essence makes it difficult for their own designers to understand.
Too risky for intellectual property. Some engineers point out that being transparent is also synonymous with risk for the company, especially in terms of intellectual property in a competitive market.

In short, developers agree on the need for transparency: the value is strong but its application is subjective. There is a tension between an ideal of transparency and explicability and the practical and technical constraints linked to its implementation.

A framework for the study of algorithms. In order to deal simultaneously with the ideals, the practical and technical constraints and the impacts, we need to arrange the focus and to choose the right methods to grasp particularly the political and ethical tensions that are at the core of the study of algorithms.

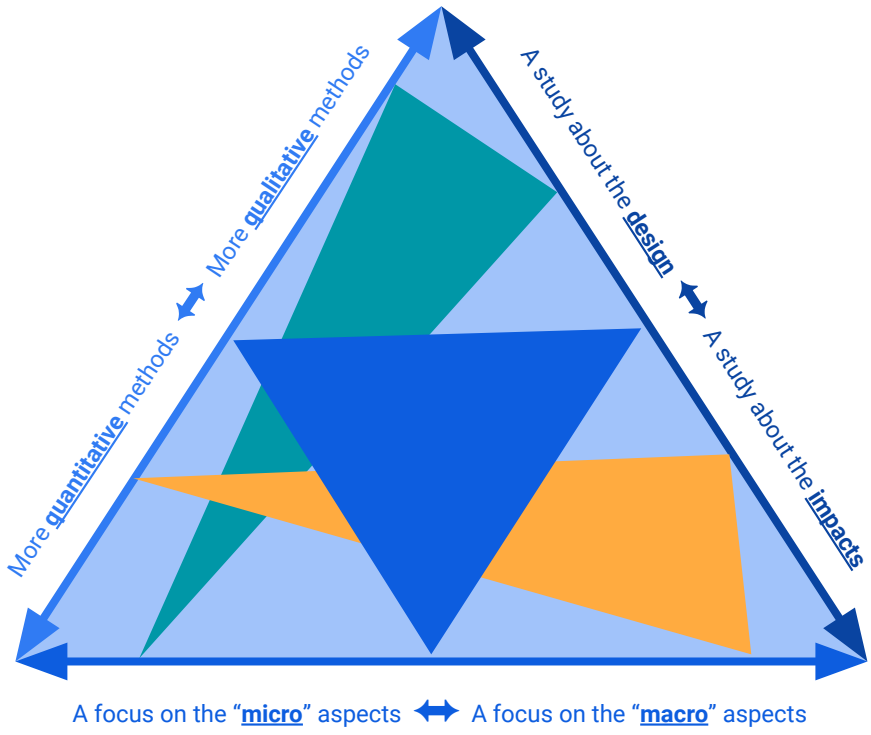
For this purpose, we suggest a simple framework that allows to situate one's research and to embrace the complexity of research on algorithms. Knowing that it is obviously impossible to cover everything with one study, this framework should encourage transdisciplinary initiatives to be set up around the three proposed axes, effectively responding to the seven "scholarly primitives" [6]:

- The **methodological** axis *(discovering, annotating, comparing)*: is the research rather quantitative or qualitative?
- The **focus** axis *(referring, sampling)*: is it rather on macro or micro dimensions?
- The **topic** axis *(illustrating, representing)*: is the study more concerned with the design of algorithms or their impacts?

With this poster, we call for "socio-technical" research to study those "socio-technical" objects. The two sides of the digital humanities shall intertwine, in a collaborative and transdisciplinary way, to answer the complexity of understanding algorithms and algorithmization.

Future work. It is likely that a change in approach will lead to a new way of working together and of presenting and publishing such research, in a hopefully more transparent and open way. This will be the aim of future meta-work.

A framework for the study of algorithms



"Socio" research
Qualitative studies about the "design" of algorithms by engineers ("micro" aspects).

"Technic" research
Quantitative studies about the "impacts" of algorithms on governance ("macro aspects").

"Socio-technical" research
Future studies exploiting all dimensions of digital humanities.
Ideally a quali-quantitative study on ties between the design and impacts of algorithms in a "meso" focus.

References
[1] Roth, Camille. 2019. "Digital, Digitized, and Numerical Humanities." *Digital Scholarship in the Humanities* 34(3):616–32.
[2] Bastin, Gilles, and Tubaro, Paula. 2018. "Le moment big data des sciences sociales." *Revue française de sociologie* 59(3):375–94.
[3] Poiroux, Jérémie and Roth, Camille. Forthcoming. "The manufacture of recommender systems".
[4] Burrell, Jenna. 2016. "How the Machine 'thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3(1):2053951715622512.
[5] Ananny, Mike, and Kate Crawford. 2018. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society* 20(3):973–89.
[6] Unsworth, John. 2000. "Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This." *In Symposium on Humanities Computing: Formal Methods, Experimental Practice*. King's College, London, 13:5–00.