Interpretive flexibility in data science and artificial intelligence

Tiphaine Viard, Mélanie Gornet
i[3], SES,
Télécom Paris
Institut Polytechnique de Paris
tiphaine.viard@telecom-paris.fr

Simon Delarue
LTCI,
Télécom Paris
Institut Polytechnique de Paris
simon.delarue@telecom-paris.fr

December 18, 2023

1 Context

Interpretive flexibility – the capacity that relevant social groups appropriate technology to create and sustain divergent opinions, uses and meanings – is a concept widely used in science and technology studies (STS), and in particular the EPOR [1] and Social Construction of Technology (SCOT) method [2]. In the SCOT method, innovation follows a *swirling* process, rather than a linear one, where different relevant social groups offer different uses and meanings of their relation to a technological artifact, thereby modifying the artifact through operations of *variation* and *selection*. When this process stabilises, the technological artifact reaches *closure*, *i.e.* the stabilisation around a set of uses and meanings.

This has led to contextualising multiple forms of interpretive flexibility [3], expanding the scopes adopted in EPOR and SCOT to encompass nuance: EPOR is framed as describing a *regress of truth* (typically, in a scientific controversy), while SCOT describes a *regress of usefulness*; the authors in [3] finally introduce the notion of *regress of relevance* as an additional lens, in their case to study the controversy around neural networks [4]. The work of Humphreys extends this further, reinforcing the dialogic relationship between technology and culture, and enriches the notion of *relevant social groups*, by separating different stake holders: bystanders, users, advocates and producers [4]. Notice however that this fails to encode the relations and interactions between individuals and social groups.

Furthermore, in all these cases, interpretive flexibility has been used as a concept to describe and understand how different relevant social groups shape the use of a technological artifact [5]. To the best of the authors' knowledge, interpretive flexibility has only been used descriptively and has not permeated computer science research. We would like to explore the converse perspective, *i.e.* how the technological artifact can shape human uses.

In data science and machine learning in particular, it is increasingly common that algorithms designed in research labs become available to the public, or used on a wider audience. In that sense, not considering interpretive flexibility at conception time makes it increasingly likely that closure will happen along already established power divides, reinforcing them rather than shifting them: one may think, without exhaustivity, of the COMPAS affair [6]; facial recognition's lack of representative datasets, especially when it comes to BIPOC[1] and gender minorities [7]; the quantitative and qualitative lack of datasets in languages other than a few selected ones, such as english [8], the exploitation of labour along colonial North-South lines [9]; the use of machine learning in fraud detection leading to unfair discrimination in the Netherlands [10]; among many others [11, 12, 13].

We argue that the concept of interpretive flexibility can be used in a prescriptive way, rather than a purely descriptive one. We observe a lack of research in this aspect, in exact and social sciences, and we identify and discuss three beliefs that contribute to this situation:

1. AI is a technical object, outside of the realm of social questions;
2. TProviding flexibility is about implementation with respect to end users, and does not belong to the scientific realm of study;

[1] Black, Indigenous, People of Color

3. Introducing interpretive flexibility in mathematical models means an increase in complexity, both conceptual and computational.

2 Addressing the beliefs

The first belief, namely seeing artificial intelligence and machine learning as a purely technical object, is the most studied already. Indeed, there is a rich body of works defining and studying sociotechnical systems, their limits and the interplay of the technical and the social [14]. Though it is not a new topic [15], the recent years have seen the development of multiple works contextualising data science and artificial intelligence in its broad sociotechnical context, under the "AI ethics" umbrella. However, this thinking is still not commonplace, as exemplified by a recent press release by the French governement. [2] It is still customary, in the scientific, institutional and public discourse, to see machine learning and data science discussed only through its technical aspects.

This makes it tempting to relegate the problem to a policy implementation one: researchers can devise models, formalisms and algorithms, and let end users and institutions decide of their deployment. This leads to devising a plethora of methods to make models explainable, fair, transparent, accountable and so on, which sound as hollow if they are researched without addressing the structures of power in which data science and artificial intelligence are rooted. In other words, to paraphrase Pratyusha Kalluri, *we should ask how AI shifts power, rather than if it is good or fair* [16]. We argue that being able to model, maintain and understand both consensus and dissensus is a desirable feature, that shapes the horizon of possibilities and thereby the real-world implementations of data science and machine lerning models and algorithms.

One may argue that interpretive flexibility comes as a fuzzy concept, difficult to quantify and make operational. In our fields, methods and algorithms that classify and work with partitions are easier to conceive, typically more tractable, which are desirable properties. Introducing flexibility at this level can be seen as dramatically raising the conceptual and computational complexity of such models. Yet, the idea of reformulating problems or changing the object of focus is a staple of mathematics and computer science: for example, optimisation methods offer primal and dual problems; multiple problems can be computationally resolved to the 3-SAT problem; and so on. Our argument is that interpretive flexibility is another goal that can be reformulated around. For example, on graphs (*i.e.* set of nodes V connected by edges $E \subseteq V \times V$), node clustering and classification is a commonplace task, with numerous efficient solutions [17, 18]. Though many of them output a partition of the nodeset (*i.e.* each node belongs to one and only one cluster), partitioning the edges instead gives overlapping information about the nodes; there are many examples of community detection on graphs that partition the edges in order to obtain overlapping node communities [19, 20]. In other words, shifting the ontological level of study can provide flexibility on lower levels.

3 Proposal and discussion

In conclusion, we argue that interpretive flexibility as a concept would benefit from going from a descriptive tool to a prescriptive one, in order to root data science and machine learning in its sociotechnical nature, and shift power structures rather than reinforce them. We also argue that the prism of relationships between relevant social groups is key to achieving this goal. While it appears initially both as a more complex and downstream problem, we argue that both assumptions should be challenged. Indeed, we argue that interpretive flexibility, like many problems, can be reformulated and incorporated into deterministic problems; and that embedding it early on contributes to rooting data science and machine learning in transformative practices.

4 About the authors

Tiphaine Viard is a *maîtresse de conférences* (roughly equivalent to tenured assistant professor) of AI and society in the social sciences laboratory at Télécom Paris/Institut Polytechnique de Paris. her work and expertise is at the crossroads of graph modelling for real-world phenomenon and expliciting the social world of artificial intelligence.

Mélanie Gornet is a PhD student at Télécom Paris/Institut Polytechnique de Paris, working on the interplay between technical models and constraints, and regulatory and standardisation questions.

Simon Delarue is a PhD student at Télécom Paris/Institut Polytechnique de Paris, working on leveraging graphs for interpretable data science and machine learning.

[2] `https://www.numerique.gouv.fr/uploads/CP_Stanislas_GUERINI_experimente_l-IA_generative_dans_les_services_publics.pdf`

References

[1] Harry M Collins. Stages in the empirical programme of relativism, 1981.

[2] Trevor Pinch. The social construction of technology: A review. *Technological change*, pages 17–35, 2012.

[3] Uli Meyer and Ingo Schulz-Schaeffer. Three forms of interpretative flexibility. 2006.

[4] Lee Humphreys. Reframing social groups, closure, and stabilization in the social construction of technology. *Social epistemology*, 19(2-3):231–253, 2005.

[5] Neil F Doherty, Crispin R Coombs, and John Loan-Clarke. A re-conceptualization of the interpretive flexibility of information technologies: redressing the balance between the social and the technical. *European Journal of Information Systems*, 15:569–582, 2006.

[6] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6):1–35, 2021.

[7] Inioluwa Deborah Raji, Timnit Gebru, Margaret Mitchell, Joy Buolamwini, Joonseok Lee, and Emily Denton. Saving face: Investigating the ethical concerns of facial recognition auditing. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 145–151, 2020.

[8] Iroro Orife, Julia Kreutzer, Blessing Sibanda, Daniel Whitenack, Kathleen Siminyu, Laura Martinus, Jamiil Toure Ali, Jade Abbott, Vukosi Marivate, Salomon Kabongo, et al. Masakhane–machine translation for africa. *arXiv preprint arXiv:2003.11529*, 2020.

[9] Paola Tubaro, Clément Le Ludec, and Antonio A Casilli. Counting 'micro-workers': societal and methodological challenges around new forms of labour. *Work Organisation, Labour & Globalisation*, 14(1):67–82, 2020.

[10] Marvin Van Bekkum and Frederik Zuiderveen Borgesius. Digital welfare fraud detection and the dutch syri judgment. *European Journal of Social Security*, 23(4):323–340, 2021.

[11] Trishan Panch, Heather Mattie, and Rifat Atun. Artificial intelligence and algorithmic bias: implications for health systems. *Journal of global health*, 9(2), 2019.

[12] Joanne Hinds, Emma J Williams, and Adam N Joinson. "it wouldn't happen to me": Privacy concerns and perspectives following the cambridge analytica scandal. *International Journal of Human-Computer Studies*, 143:102498, 2020.

[13] Frederik Zuiderveen Borgesius et al. Discrimination, artificial intelligence, and algorithmic decision-making. *línea], Council of Europe*, 2018.

[14] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 59–68, 2019.

[15] Steve Woolgar. Why not a sociology of machines? the case of sociology and artificial intelligence. *Sociology*, 19(4):557–572, 1985.

[16] Pratyusha Kalluri et al. Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature*, 583(7815):169–169, 2020.

[17] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[18] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17:395–416, 2007.

[19] Santo Fortunato. Community detection in graphs. *Physics reports*, 486(3-5):75–174, 2010.

[20] Noé Gaumont, François Queyroi, Clémence Magnien, and Matthieu Latapy. Expected nodes: a quality function for the detection of link communities. In *Complex Networks VI: Proceedings of the 6th Workshop on Complex Networks CompleNet 2015*, pages 57–64. Springer, 2015.