Examining algorithms in the light of their ground truth datasets: Results, objections, and avenues of reflection

Link: https://doi.org/10.1007/s44206-025-00197-4

Final manuscript for Digital Society (Topical Collection: Politics of Machine Learning Evaluation)

Florian Jaton, EPFL, Switzerland

Abstract

Since at least the 1990s, social and media studies contributed to presenting algorithms as human, cultural constructs. More recently though, fine-grained ethnographic inquiries went further in the analysis and showed that most algorithms derive from benchmark referential datasets – often called "ground truth" – that gather input data and output targets, thereby establishing what can be approximated computationally and evaluated statistically. In this commentary, I briefly review the most recent results on this line of research one may call ground truth studies. I then consider some objections to this ground truth-centered conception of algorithms and point out avenues for future thinking.

Keywords

Algorithms, ethnography, machine learning, artificial intelligence, ground truths, referential datasets, benchmarks

Algorithms! While the term can be confusing – difficult indeed to base it on a rigorous definition[1] – it is nonetheless very much debated and disputed in Science and Technology Studies (STS).[2] But where does the magnetism for the social study of algorithms come from? What are the notable results of this line of research, as well as its limits and future challenges? In this commentary, I first briefly review the contemporary history of the social study of algorithms and its recent – and refreshing – ethnographic turn. I then discuss one of its significant results: most algorithms can be considered derivatives of referential datasets (often called "ground truths"). I then consider two objections to this ground truth-centered glance on algorithms and point out avenues for future thinking.

1. Social studies of algorithms: from critique, to drama, and then ethnography

The 1990s were particularly rich in utopian discourses on computer technologies.[3] The growing distribution of microcomputers in the professional and private spheres, the promising beginnings of internet technologies, Clintonian announcements of "information superhighways," and the imminent advent of a new millennium: the stars were aligned for the development of a promising – and often promotional – vision of then-called new information technologies.

In response to these discourses equating technological progress with social advances, many sociologists, anthropologists, and media scholars set out to measure the concrete, social, and political dynamics of this broad, and actual, computerization process (e.g., King, 1991; 1994; Brouillard & Lafrance, 1996; Vedel, 1996). In the US, the 1993 law stipulating that 5% of Human Genome Research Project funds be invested in research into the social, legal, and ethical consequences of genomics helped position social science practitioners as guardians of justice and fundamental freedoms with regard to the deployment of computer technologies (Poon, 2016; Reardon, 2016; Zwart & Nelis, 2009). The second edition of Rob Kling's anthology *Computerization and Controversy* (1996) – which covers professional, domestic, fictional, affective, and security issues associated with computerization processes – provides a rich overview of the topics that ran through the study of the social and political consequences of computer technologies during this pivotal period.

Throughout the 2000s, the social study of computer technologies became more frontally critical, especially in the US, which contributed to the emergence of the "algorithm" as an object of study. Researchers contested the claims of Web technology promoters regarding empowerment and accessibility, revealed the limitations of search engines, and examined disparities in data mining devices (e.g., Hoffman & Novak, 1998; Introna & Nissenbaum, 2000; Lawrence & Giles, 1999; Civille, 1996).

---

[1] The notion of algorithm has no rigorous, stabilized definition. In mathematics and logic, defining what is an algorithm is still an open question. On this topic, see Seiller (2024).

[2] As I say elsewhere (Jaton, 2021b: 1), building on Dear and Jasanoff (2010), STS can be considered "a subfield of social sciences that aims to document the co-construction of science, technology, and the collective world. What connects the practitioners of this heterogeneous research community is the conviction that science is not only the expression of a logical empiricism, that knowledge of the world does not preexist, and that scientific and technological truths are dependent on collective arrangements, instrumentations, and dynamics."

[3] Two arbitrary but illustrative examples of technological utopianism in the 1990s could be Weiland (1993) and Stewart (1993).

Progressively, as surveillance technologies grew post-9/11, studies of detection programs and *algorithms* – the term appears at that time in the critical literature[4] – emphasized their biases, culminating in increased focus on discrimination and invisibilization processes in the 2010s (e.g., Gandy, 2002; Zureik & Hindle, 2004; Introna & Wood, 2002; Kraemer et al., 2010; Gillepsie, 2013; Bucher, 2012; Bozdag, 2013).

The importance of these critical studies of the effects of algorithms should be emphasized and re-emphasized. In their own way, they acted as a firewall against the commercial and seductive rhetoric of computer technologies' promoters and over-admirers. More generally, and crucially, this line of research that Gillepsie and Seaver (2016) labeled "critical algorithms studies" managed to *problematize* the notion of algorithm and somewhat decouple it from its modernist trope of radically detached objectivity. However, the proliferation of these critical studies and, in a way, their redundancy also generated a series of aporias that Malte Ziewitz finely described as an algorithmic *drama* (2016).

Ziewitz's argument suggests that as social research increasingly – and rightly – focused on criticizing the negative effects of algorithms, the heteronomous sociotechnical arrangements that enable them to come into existence and operate were gradually overlooked, leading to algorithms being perceived as autonomous and abstract entities. Over time, the ties to algorithms' supporting networks were ignored, and algorithms became seen as mysterious, powerful, and complex. This shift led to an "algorithmic drama," where algorithms were viewed as powerful precisely because they were inscrutable, creating a circular perception of their influence and opacity. A looping and disempowering drama, in short, whose political effects are still tangible today (Jaton & Vinck, 2023).

To try to break out of this dramatic spiral, several STS-sensitive authors have recently embarked on ethnographic inquiries into the mundane and concrete construction of algorithms in applied fields like image processing (Jaton, 2017; 2021a; 2022; Vinck et al., 2018), computer audition (Kang, 2023; Seaver, 2022), video processing (Engdahl, 2024), text analysis (Bechmann & Bowker, 2019; Wu, 2024), or predictive medicine (Henriksen & Bechmann, 2020; Jaton, 2023; Muhr, 2023)[5]. In much the same way as the late 1970s' ethnographic work of Bruno Latour and Steve Woolgar (1986), Michael Lynch (1985), Karin Knorr-Cetina (1981), and others sought to move away from the abstraction of (critical) epistemology by providing a down-to-earth view of the production of scientific statements, the work of these new ethnographers of *computer* science sought to move away from the abstraction of (critical) digital epistemology to gradually build up a concrete and realistic – and therefore empowering – understanding of algorithms (Alexandre, 2024).

---

[4] Whereas the terms "software", "code" or "software-algorithm" were initially favored, the single term "algorithm" began to become increasingly common in Anglo-American surveillance studies literature from the 2000s onwards. It would be interesting to know more about the channels through which the term "algorithm" came to be confidently used in the Anglo-American critical literature.

[5] It is important to note here that another series of authors with a hybrid profile (often computer scientists with a social science education) have also greatly contributed to the progressive detachment from this algorithmic drama by critically examining deep learning code (e.g., Hua & Haley, 2024; Hua & Rhee) and by subjecting deep learning model outputs to extensive statistical analysis (e.g., Bender et al., 2021). Today, this line of research is often referred to as Critical AI Studies (Raley & Rhee, 2023).

2. Notable result: algorithms can be considered derivatives of referential ground truth datasets

The preliminary results of this still on-going ethnographic turn are very rich, particularly in terms of the professional coordination mechanisms and programming practices that enable the concrete day-to-day shaping of algorithms, as well as constant, laborious, and very material tinkering of computer scientists to achieve 'good enough' code and feature extraction (e.g., Thylstrup et al., 2022; Bialski, 2024). But the common observation that has been, for now, the object of most discussion is certainly that many algorithms are derived from, and thus refer to, data collected in more or less arbitrary ways and assembled in more or less structured *sets* (Le Ludec et al., 2023; Amoore et al., 2024). These referential datasets – known as "ground truth," "benchmark," or simply "dataset," depending on the applied field (Jaton, 2021b) – gather and connect at least two subsets: 1) an input-data subset, which contains the information the would-be algorithm is intended to process, and 2) an output-targets subset, which holds the results the would-be algorithm is expected to produce (Rettberg et al., 2024).

Far from being marginal, these referential ground truth datasets frame and define algorithm construction work, as shown by a series of still on-going inquiries one may call ground truth studies. Among them, one could mention the work of Grosman and Reigeluth (2019) who documented the development of surveillance systems within a European Research and Innovation project. In their case study, machine learning algorithms required a new ground truth dataset to detect so-called "suspicious behavior." But to create this dataset, computer scientists and engineers *acted out* suspicious and non-suspicious roles, and then annotated the recordings. Beyond the surprising theatrical aspect of the process, it appeared that the behaviors these engineers taught the algorithms were arbitrary and based upon their own projections. As a result, the algorithmic systems became indeed efficient at detecting abnormal behaviors, but this abnormality was based on the socio-cultural habits of thought of the engineers present during recordings and annotations.

Another exemplary ground truth study is the inquiry by Henriksen and Bechman (2020) into a Scandinavian software company developing "AI solutions" for healthcare. In this work, they emphasize the importance of constructing a reliable referential dataset for the company, as it was essential for evaluating the performances of newly developed algorithms and attracting potential clients. Interestingly, then, the company's main challenges were not computational *per se*, because powerful open-source deep learning tools were already available and adaptable without too many efforts. Rather, the main issues concerned the actual construction of reliable referential ground truth datasets that could support meaningful performance evaluations for healthcare customers.

More recently, Engdahl (2024) followed the construction of a referential dataset of daily activities captured "in the wild," aiming to support computer vision tasks in academia and the industry. But as he shows with great precision, the "naturalness" of this database project (Jaton & Vinck, 2016) was frustrated by its necessary research apparatus, as the construction of the dataset involved the careful alignment of standards and negotiations among data subjects. The activities recorded were thus not quite spontaneous but shaped by various factors, including privacy concerns and legal standards. Ultimately, and ironically (and perhaps also, necessarily), the ground truth dataset – and the algorithms derived from it – ended up reflecting framing and alignment processes rather than behaviors "in the wild" captured from a detached viewpoint.

In a similar vein, Edward Kang (2023) finely described the efforts required to arbitrarily ground truth the "vocability" of employee-fit for the case of an R&D project in a tech start-up. This discovery, and others (Kang, 2022), led him to develop the interesting notion of *ground truth tracing*, which refers to the potentially systematic examination of how issues are framed through the construction of ground truth datasets. This speculative proto-classification device is thought to assess forms of epistemological and ethical soundness, as it may help determine whether an issue can be accurately represented within a dataset and, consequently, whether it can be effectively addressed by an algorithmic system.

It is quite clear, then, that ground truth studies enrich the social analysis of algorithms by shifting the analytical gaze on what makes algorithms germinate, come into existence, and circulate (see also Heimstädt, 2023; Girard-Chanudet, 2023). It is a much welcome program of (quite) critical inquiry, that continues to produce fascinating results, as the present topical collection demonstrates. And more generally, it is safe to say that, beyond their differences, these studies show that referential datasets and their assemblage have a massive impact on algorithm construction and content. In fact, the massiveness of this impact even suggests what one may call a *socio-ethnographic definition of algorithms*, based on their dependence on referential datasets. According to this definition – which remains fragile and speculative – an algorithm could be considered *a computer-compatible approximation of an ideal function supposedly governing the relationship between the input-data and the output-targets of at least one referential dataset* (Rettberg et al., 2024: 85).[6]

3. First objection/avenue of reflection: what about un- and self-supervised learning?

Conceiving of algorithms as derivatives of referential datasets may seem a little iconoclast, to say the least, in this time of hype (as well as *criti*-hype, cf. Vinsel, 2021) around so-called 'artificial intelligence' (Jaton & Sormani, 2023). Hence, maybe, several more or less assignable objections to this socio-ethnographic conception of algorithms, which constitute, it seems to me, as many avenues for research and reflection.

The first objection, sometimes even hinted at within the ethnographic works mentioned above, is linked to what is commonly referred to as un- and self-supervised learning. It can be summarized as follows: referential datasets are central *only* for algorithms based on supervised learning, i.e., that

[6] Interestingly, Mackenzie (2017: 81) points out that some computer scientists reach a close definition, yet by other means. This is the case, for example, of Hastie et al. (2009) who, in their statistical learning textbook, state that the goal of algorithm design "is to find a useful approximation f'(x) to the function f(x) that underlies the predictive relationship between input and output" (28). More generally, another interesting feature of this ground truth-centered definition is that it allows one to make a welcome distinction between computer programs and algorithms. Indeed, if one accepts that the algorithms that interest us are programs (since they have to be computer-readable), not all programs are algorithms, as it becomes the act of referring to an external set to reproduce/approximate part of its constituent relationships that defines the algorithmic form.

requires annotated data. When it comes to unsupervised or self-supervised learning, ground truths and the practices underlying their construction lose their importance (the objection goes).

As far as strictly unsupervised learning is concerned (which, for the most part, refers to mere descriptive and exploratory algorithms such as k-means or principal component analysis), it is indeed hard to conceive, as things stand, the referential datasets from which such algorithms would be derived. However, as far as I know, there is currently no in-depth ethnographic description of the concrete shaping of this type of unsupervised algorithms, which tends to be carried out in theoretical computer science laboratories. In this sense, while the objection makes sense, it also suggests *an avenue of research*, which would either invalidate the proposition of firmly associating the very notion of algorithm with referential sets, or refine the phenomenon by documenting, for example, other – maybe looser – forms of referential sets.

As far as self-supervised learning is concerned, two things need to be emphasized, not to say clarified. The first is that self-supervised learning is fully part of supervised learning, once said that the labels reside within the collected data itself. Schematically, in the case for example of language models (popular products of self-supervised learning nowadays), the word/token that follows a given word/token constitutes a ground truth reference. And this supervision is obviously crucial, since it enables the calculation of an error, which is then back-propagated within the layers of the neural networks, more or less thick and deep depending on the chosen architecture. Between 2015 and 2020, there may have been some terminological confusion in data science and applied computing, with some eminent authors occasionally associating self-supervision with unsupervision. Today, however, these concepts seem to be more clearly separated in the vernacular literature.[7]

Secondly, self-supervised models remain fundamentally subordinated to referential datasets, since they must in any case be evaluated. As I explain elsewhere, this practical imperative comes from the fact that these algorithms "are not intended to remain theoretical: They are designed to be ultimately used and worked upon, which implies comparing them to benchmarked ground-truth datasets in order to show their relevance and efficiency" (Jaton, 2021b: 4). In that sense, then, referential datasets remain essential to make these self-supervised algorithms exist as valuable devices.

However, asserting this fundamental subordination of self-supervised algorithms to referential datasets does not mean that this recent type of machine learning and algorithmic construction process makes no difference. On the contrary, it seems to make the topic of ground truth dataset even more convoluted and interesting, not least because of the progressive dissociation between training sets and evaluation sets (Engdahl, 2024). This dissociation is not merely technical but also deeply political – as the present Topical Collection also indicates – as the criteria and benchmarks used to evaluate machine

[7] Laureate of the prestigious 2018 Turing Award, Yann Lecun notably contributed to this confusion in his famous "cake analogy," initially stipulating that the biggest slice of the machine learning cake (meaning, the most promising type of method) consisted of unsupervised learning (Lecun, 2016). Interestingly enough, Lecun went back on his analogy in 2019, stipulating that the biggest piece of cake was, in fact, self-supervised learning (and unsupervised learning even disappeared from the analogy).

learning systems shape whose knowledge, values, and perspectives are legitimized in machine learning/AI development. More generally, it seems to me that this is another research direction to take, which would seek to document precisely the data-activity regime required to implement, assess, and sustain large self-supervised algorithmic systems.

4. Second objection/avenue of reflection: what about diplomacy (or how to compose with cyber-modernism)?

The second objection is trickier, and refers to a less assignable phenomenon. To identify it, we need to go at a higher level and consider the *backlash* that critical algorithm studies seem to be experiencing for the last few years.

Beyond high-profile incidents like the abrupt firing of famous ethical AI researcher Timnit Gebru (Wong, 2020) – and, more broadly, opportunistic and cynical U-turns coming from Big Tech – they are indeed growing signs of normative tightening within computer science research and industry. This is evident in the downsizing of research units that focus on social sciences and humanities (Vynck, 2023), the removal of social science experts from the boards of major companies (Paul, 2024), and more or less direct criticisms aimed at the legitimacy and even the seriousness of scholars critically studying algorithms (Hern, 2024). It seems that computer science professionals are increasingly vocal about their apprehension and/or resistance to collaborating with social scientists, who are often perceived as obstacles engaged in abstract – and vain – intellectual exercises (Moats & Seaver, 2019).

This reactionary moment is likely to be linked to broader cynical undertakings to preserve inequalities (Mondon and Winter, 2020). However, we cannot completely rule out the possibility of a sincere response – at least by some honest computer science professionals – to a form of critical clumsiness. The fine-grained ethnographic descriptions of computer science in action mentioned above do not seem to be concerned here: on the contrary, they tend to be recognized as reliable, accurate, and important (see for example Dickson, 2022). The problem seems to come from broader, alternative theorizations of critical algorithms students (including me!) that may appear disappointing, not to say insulting, to some computer science practitioners. In this sense, similarly to the so-called "science wars" during the 1990s,[8] the current situation of relative backlash could also be, at least in part, linked to a lack of *tact* on the part of the critical faction when it comes to inferring more general propositions. And this would ultimately refer to an inability to identify and respect the *values* central to those whose work and achievements that we – ethnographers of algorithms in the making – try to make understandable to others.

[8] In very broad strokes, the "science wars" were a series of intellectual debates in the 1990s, sparked by natural scientists who criticized social science and literary analyses of science – particularly the sociology of scientific knowledge (SSK) and what will be later be called actor-network theory (ANT) – for allegedly undermining scientific objectivity by viewing scientific truths as socio-cultural constructs. Interestingly, when one reads the retrospective accounts on these debates by some of their most targeted social scientists (e.g., Collins et al., 2017; Latour, 2013: 12-14), one sees that they are leaning towards *mea culpa*. While they stand by what they described, they also recognize that they could have been more cautious with their more general theoretical propositions.

If we accept this (highly debatable) hypothesis, one of the future challenges of critical algorithm studies – and *a fortiori* also of ground truth studies – would be to attempt to make critical achievements compatible with the elements of the computer scientific *lore* to which practitioners are most intensely attached. An example might be the notion of intelligence, which today remains strongly ingrained and positively charged within computer science and industry, to the great dismay of most critical algorithm students (Jaton & Sormani, 2023). Despite our disaffection with the notion of intelligence, how could we incorporate it in our critical algorithm framework? Or rather, how could we disarm this notion – undoubtedly one of the most harmful, notably due to its colonial and patriarchal heritage (Cave, 2020) – while retaining it, in order finally to make it compatible with the findings of, for example, ground truth studies? It is a thorny issue of political philosophy, which would seek to cautiously adapt sticky cyber-modernist heritage to anthropological actualities.

A new line of socio-philosophical research seems to be opening up here, which would attempt to gather reasonable critics and admirers around the weaknesses and strengths of algorithmic agency (Jaton & Lenglet, 2025). Quite a challenging diplomatic – and therefore collective – undertaking (Janicka, 2023), that would attempt to somewhat prolong the cyber-modernist adventure while recentering it on empirical, not to say ecological, fundamentals.

5. Acknowledgements

I would like to thank Anna Schjøtt Hansen, Dieuwertje Luitse and Tobias Blanke for making this text possible, both by editing this topical collection and by organizing the wonderful workshop 'The Politics of Machine Learning Evaluation' held in Amsterdam in 2023. I would also like to thank the two anonymous reviewers for their valuable suggestions.

6. References

Alexandre, O. (2024). Numérique: En quête de terrains. *RESET. Recherches en sciences sociales sur Internet*, *13*, Article 13. https://doi.org/10.4000/12cn6

Bechmann, A., & Bowker, G. C. (2019). Unsupervised by any other name: Hidden layers of knowledge production in artificial intelligence on social media. *Big Data & Society*, *6*(1), 205395171881956. https://doi.org/10.1177/2053951718819569

Bender, E., Gebru, T., McMillan-Major, A., et al. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York, NY, USA, 1 March 2021, 610–623. Available at: https://dl.acm.org/doi/10.1145/3442188.3445922

Bialski, P. (2024). Middle Tech: Software Work and the Culture of Good Enough. Princeton University Press.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3), 209-227. https://doi.org/10.1007/s10676-013-9321-6

Brouillard, P., & Lafrance, J.-P. (1996). À la conquête électronique de la toison d'or ou la mise en place des contenus et des services sur les autoroutes de l'information. *Réseaux*, *78*(4), 53-83.

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 1461444812440159. https://doi.org/10.1177/1461444812440159

Cave, S. (2020). The Problem with Intelligence: Its Value-Laden History and the Future of AI. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 29-35. https://doi.org/10.1145/3375627.3375813

Civille, R. (1996). The Internet and the poor. In B. Kahin & J. Keller (Éds.), *Public access to the Internet* (p. 175-207). MIT Press.

Collins H, Evans R and Weinel M (2017) *STS as science or politics? Social Studies of Science* 47(4), 580–586.

Dickson, B. (2022, février 21). How algorithms come into being. *TechTalks*. https://bdtechtalks.com/2022/02/21/constitution-of-algorithms-florian-jaton/

Engdahl, I. (2024a). Agreements 'in the wild': Standards and alignment in machine learning benchmark dataset construction. *Big Data & Society*, *11*(2), 20539517241242457. https://doi.org/10.1177/20539517241242457

Engdahl, I. (2024b). Wuthering Weights—Localisation Trajectories of Machine Learning Models for Local Ends. In V. Fors, M. Berg, & M. Brodersen (Éds.), *The De Gruyter Handbook of Automated Futures* (p. 304-321). De Gruyter.

Gandy, O. H. (2002). Data Mining and Surveillance in the Post-9.11 Environment. In K. Ball & F. Webster (Éds.), *The Intensification of Surveillance Crime, Terrorism and Warfare in the Information Age* (p. 113-137). Pluto Press.

Gillepsie, T. (2013). The Relevance of Algorithms. In *Media Technologies* (MIT Press, p. 167-194). Tarleton Gillepsie, Pablo Boczkowski, and Kirsten Foot.

Gillepsie, T., & Seaver, N. (2015, novembre 5). Critical Algorithm Studies: A Reading List. *Social Media Collective*. https://socialmediacollective.org/reading-lists/critical-algorithm-studies/

Girard-Chanudet, C. (2023). "What if the algorithm mimics our mistakes?" Constraints and effects of data annotation for AI:Constraints and effects of annotating AI training data. *Réseaux*, *240*(4), 111-144. https://doi.org/10.3917/res.240.0111

Grosman, J. & Reigeluth, T. (2019). Perspectives on algorithmic normativities: engineers, objects, activities. Big Data & Society, 6(2), 2053951719858742. https://doi.org/10.1177/2053951719858742

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (2nd edition). Springer.

Heimstädt, C. (2023). Making plant pathology algorithmically recognizable. *Agriculture and Human Values*. https://doi.org/10.1007/s10460-023-10419-5

Henriksen, A., & Bechmann, A. (2020). Building truths in AI: Making predictive algorithms doable in healthcare. *Information, Communication & Society*, *23*(6), 802-816. https://doi.org/10.1080/1369118X.2020.1751866

Hern, A. (2024, août 6). Why AI's Tom Cruise problem means it is 'doomed to fail'. *The Guardian*. https://www.theguardian.com/technology/article/2024/aug/06/ai-llms

Hoffman, D. L., & Novak, T. P. (1998). Bridging the Racial Divide on the Internet. *Science*, *280*(5362), 390-391. https://doi.org/10.1126/science.280.5362.390

Hua, M. & Raley, R. (2023). How to Do Things with Deep Learning Code. *Digital Humanities Quarterly* 17(2), 1-19. https://doi.org/10.48550/arXiv.2304.09406

Introna, L. D., & Nissenbaum, H. (2000). Shaping the Web: Why the Politics of Search Engines Matters. *The Information Society*, *16*(3), 169-185. https://doi.org/10.1080/01972240050133634

Introna, L., & Wood, D. (2002). Picturing Algorithmic Surveillance : The Politics of Facial Recognition Systems. *Surveillance & Society*, *2*(2/3). http://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/3373

Janicka, I. (2023). Reinventing the Diplomat: Isabelle Stengers, Bruno Latour and Baptiste Morizot. *Theory, Culture & Society*, 40(3), 23-40. https://doi.org/10.1177/02632764221146717

Jaton, F. (2017). We get the algorithms of our ground truths: Designing referential databases in digital image processing. *Social Studies of Science,* 47(6), 811–840. https://doi.org/10.1177/0306312717730428

Jaton, F. (2021a). *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating*. MIT Press.

Jaton, F. (2021b). Assessing biases, relaxing moralism: On ground-truthing practices in machine learning design and application. *Big Data & Society*, 8(1), 20539517211013569. https://doi.org/10.1177/0306312717730428

Jaton, F. (2022). Éléments pour une sociologie de l'activité de programmation. *RESET. Recherches en sciences sociales sur Internet*, 11(1), 1-32. https://journals.openedition.org/reset/3829

Jaton, F. (2023). Groundwork for AI: Enforcing a benchmark for neoantigen prediction in personalized cancer immunotherapy. *Social Studies of Science*, 53(5), 787–810. https://doi.org/10.1177/03063127231192857

Jaton, F. & Lenglet, M. (2025). A critical praise of algorithmic agency: Phased wisdom, isomorphic trueness, and extrinsic intelligence. STS-CH Conference 2025, Zurich, Switzerland.

Jaton, F., & Sormani, P. (2023). Enabling 'AI'? The situated production of commensurabilities. *Social Studies of Science*, 53(5), 625–634. https://doi.org/10.1177/03063127231192857

Jaton, F., & Vinck, D. (2016). Unfolding frictions in database projects. *Revue d'anthropologie des connaissances*, 10(4), a-m. https://doi.org/10.3917/rac.033.a

Jaton, F. & Vinck, D. (2023). Politicizing Algorithms by Other Means: Toward Inquiries for Affective Dissensions. *Perspectives on Science*, 31(1), 84–118. https://doi.org/10.1162/posc_a_00582

Kang, E. B. (2022). Biometric imaginaries : Formatting voice, body, identity to data. *Social Studies of Science*, 52(4), 581-602. https://doi.org/10.1177/03063127221079599

Kang, E. B. (2023). Ground truth tracings (GTT) : On the epistemic limits of machine learning. *Big Data & Society*, 10(1), 20539517221146122. https://doi.org/10.1177/20539517221146122

Kling, R. (1996). Computerization and Controversy: Value Conflicts and Social Choices. Morgan Kaufmann.

Knorr-Cetina, K. (1981). The manufacture of knowledge: An essay on the constructivist and contextual nature of science. Pergamon Press.

Kraemer, F., Overveld, K., & Peterson, M. (2010). Is there an ethics of algorithms? *Ethics and Information Technology*, 13(3), 251-260. https://doi.org/10.1007/s10676-010-9233-7

Latour, B., & Woolgar, S. (1986). *Laboratory Life : The Construction of Scientific Facts*. Princeton University Press.

Latour, B. (2013). *An Inquiry Into Modes of Existence*. Cambridge, MA: Harvard University Press.

Lawrence, S., & Giles, C. L. (1999). Accessibility of information on the web. *Nature*, *400*(6740), 107. https://doi.org/10.1038/21987

Lynch, M. (1985). *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. Routledge Kegan & Paul.

Mackenzie, A. (2017). *Machine Learners: Archaeology of a Data Practice*. MIT Press.

Moats, D., & Seaver, N. (2019). "You Social Scientists Love Mind Games" : Experimenting in the "divide" between data science and critical algorithm studies. *Big Data & Society*, *6*(1), 2053951719833404. https://doi.org/10.1177/2053951719833404

Mondon, A., & Winter, A. (2020). *Reactionary Democracy: How Racism and the Populist Far Right Became Mainstream*. Verso Books.

Muhr, P. (2023). The Unobserved Anatomy: Negotiating the Plausibility of AI-Based Reconstructions of Missing Brain Structures in Clinical MRI Scans. In A. Flüchter, B. Förster, B. Hochkirchen, & S.

Schwandt (Éds.), *Plausibilisierung und Evidenz : Dynamiken und Praktiken von der Antike bis zur Gegenwart* (p. 169-192). Bielefeld University Press. https://doi.org/10.1515/9783839469781-008

Paul, K. (2024, mars 9). OpenAI reinstates CEO Sam Altman to board after firing and rehiring. *The Guardian*. https://www.theguardian.com/technology/2024/mar/08/openai-sam-altman-reinstated

Poon, M. (2016). Corporate Capitalism and the Growing Power of Big Data : Review Essay. *Science, Technology, & Human Values*, *41*(6), 1088-1108. https://doi.org/10.1177/0162243916650491

Raley, R. & Rhee, J. (2023). Critical AI: A Field in Formation. *American Literature* 95(2): 185–204. https://doi.org/10.1215/00029831-10575021

Reardon, J. (2013). On the Emergence of Science and Justice. *Science, Technology, & Human Values*, *38*(2), 176-200. https://doi.org/10.1177/0162243912473161

Rettberg W, Crawford K, Schultz J, et al. (2024) An AI society: Artificial intelligence is reshaping society, but human forces shape AI. *Issues in Science & Technology*, 40(2), 77-88. https://doi.org/10.58875/tjod6844

Seaver, N. (2022). *Computing Taste: Algorithms and the Makers of Music Recommendation*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/C/bo183892298.html

Seiller, T. (2024). *Mathematical Informatics* [thesis, Université Sorbonne Paris Nord]. https://theses.hal.science/tel-04616661

Stewart, T. (1993). Boom Time on the New Frontier. *Fortune*, *128*(7), 153-161.

Thylstrup N.B., Hansen, K.B., Flyverbom, M., et al. (2022). Politics of data reuse in machine learning systems: Theorizing reuse entanglements. *Big Data & Society* 9(2): 20539517221139785.

Vedel, T. (1996). Les politiques des autoroutes de l'information dans les pays industrialisés : Une analyse comparative. *Réseaux. Communication - Technologie - Société*, *14*(78), 11-28. https://doi.org/10.3406/reso.1996.3756

Vinck, D., Camus, A., Jaton, F., & Oberhauser, P.-N. (2018). Localités distribuées, globalités localisées: Actions, actants et médiations au service de l'ethnographie du numérique. *Symposium*, *22*(1), 41-60.

Vinsel, L. (2021, février 1). You're Doing It Wrong: Notes on Criticism and Technology Hype. *Medium*. https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5

Vynck, G. D., & Oremus, W. (2023, mars 30). As AI booms, tech firms are laying off their ethicists. *Washington Post*. https://www.washingtonpost.com/technology/2023/03/30/tech-companies-cut-ai-ethics/

Weiland, R. (1993). 2001: A Meetings Odyssey. *Successful Meetings*, *42*(13), 34-39.

Wong, J. C. (2020, décembre 4). More than 1,200 Google workers condemn firing of AI scientist Timnit Gebru. *The Guardian*. https://www.theguardian.com/technology/2020/dec/04/timnit-gebru-google-ai-fired-diversity-ethics

Wu, Y.-H. (2024). Capturing the unobservable in AI development: Proposal to account for AI developer practices with ethnographic audit trails (EATs). *AI and Ethics*. https://doi.org/10.1007/s43681-024-00535-1

Ziewitz, M. (2016). Governing Algorithms: Myth, Mess, and Methods. *Science, Technology & Human Values*, *41*(1), 3-16. https://doi.org/10.1177/0162243915608948

Zureik, E., & Hindle, K. (2004). Governance, Security and Technology: The Case of Biometrics. *Studies in Political Economy*, *73*(1), 113-137. https://doi.org/10.1080/19187033.2004.11675154

Zwart, H., & Nelis, A. (2009). What is ELSA genomics? Science & Society Series on Convergence Research. *EMBO Reports*, *10*(6), 540. https://doi.org/10.1038/embor.2009.115

7. Conflict of interest statement

The author states that there is no conflict of interest.