



Algorithms Don't Have A Past: Beyond Gadamer's Alterity of the Text and Stader's Reflected Prejudiced Use

Matthew S. Lindia¹

Received: 20 February 2024 / Accepted: 23 February 2024 / Published online: 4 March 2024
© The Author(s) 2024

Abstract

This commentary on Daniel Stader's recent article, "Algorithms Don't Have a Future: On the Relation of Judgement and Calculation" develops and complicates his argument by suggesting that algorithms ossify multiple kinds of prejudices, namely, the structural prejudices of the programmer and the exemplary prejudices of the dataset. This typology at once suggests that the goal of transparency may be impossible, but this impossibility enriches the possibilities for developing Stader's concept of reflected prejudiced use.

Keywords Algorithms · Gadamer · Artificial Intelligence · Judgement · Philosophical hermeneutics · Large language models

1 Introduction

Daniel Stader's recent article, "Algorithms Don't Have a Future," takes a stand against longstanding fears that the calculative reasoning of artificial intelligence (AI) can and will displace human judgement, instead suggesting that algorithmic calculation alters "the range, the transparency, and the possibilities of judgement" (Stader, 2024, p. 3). Stader argues that while calculations can operate analogously to ossified prejudices, they cannot displace judgement itself because calculation and judgement exist according to alternate ways of being. Stader defends his claim by enlisting a particular reading and application of the role of judgement in the work of Arendt and Gadamer, particularly as Heideggerian phenomenologists operating in the wake of a Kantian framework. Stader argues that algorithms operate analogously to non-algorithmic prejudices of human experience, ultimately advocating for the transparency of the prejudices embedded into algorithms.

In keeping with general formulation of Stader's piece, I affirm his position that calculation does not replace judgement or prejudice but changes my relation with

✉ Matthew S. Lindia
lindiam@duq.edu

¹ Duquesne University, 600 Forbes Ave, Pittsburgh, PA 15282, United States

them. However, departing from Stader, I suggest that machine learning algorithms achieve this effect through an essential plurality of judgements that are ossified from *no judgement in particular*. As such, if algorithmic output is not the presentation or repetition of any singular or original prejudice, Stader's call for transparency becomes complicated because it only attends to one plane of alterity within the multiplex ossifications inherent to algorithmic output. As such, I complement his argument by elaborating a typology of algorithmic prejudices, consisting of structural prejudices of the programmer and exemplary prejudices of the dataset. This typology, then, is shown to extend Stader's concept of reflected prejudiced use.

2 Stader in Review

Stader's (2024) fundamental contributions have to do with his claims that (1) "judgement increasingly relies on calculative patterns such as algorithms" (p. 3), and (2) "that AI does not simply do what humans have done before, but that it is embedded by human judgement, thereby changing it and with it the way humans relate to themselves and others" (p. 3). Following a Kantian framework inherited and extended by Arendt and Gadamer, Stader understands judgement in the broadest sense as the subsumption of a particular under a general, not merely as a logical operation, but as an act embedded in language, human experience or temporality, and sociality. Drawing on Arendt and Gadamer, Stader then elaborates the relation of judgement and prejudice. Though Gadamer and Arendt thematize prejudice differently, their analyses share in being characterized as non-reflective in contrast to actively made judgements and their resistance to characterizing prejudice in purely negative terms as simply a harmful stereotype.

For Gadamer, pre-judgements (for sake of clarity, I will refer to Gadamer's concept as pre-judgements and Arendt's as prejudices) are that stockpile of experience that enables judgement in the first place. Pre-judgements are both an extension of Heidegger's ontological analysis of Dasein and the temporality of being, as well as "the non-given totality [of tradition] to which every act of judgement refers" (Stader, 2024, p. 10). For Arendt, a prejudice is a judgement "which originally had its own appropriate and legitimate experiential basis and which evolved into a prejudice only because it was dragged through time without its ever being reexamined or revised" (Arendt, 2005, p. 101, as quoted in Stader, 2024, p. 21). In Stader's terms, Arendt's conception of prejudice is as an *ossified judgement*; a judgement that was once made actively but has since become automatic. Though both characterizations of prejudice are nonreflective, Gadamer's pertains to the ontological condition of interpreting experience according to my projects and my historically-effected-consciousness, whereas Arendt's pertains to the transformation of an act of judgement into an enduring or static evaluation. The prefix, *pre-*, then, relates to these notions in different ways. For Gadamer, the *pre-* of pre-judgement refers to the ontological conditions that condition my initial understanding preceding the judgement that make judgement possible, whereas for Arendt, the prefix refers to the judgement that has already been made and is now simply unreflectively assumed. Stader (2024) draws the connection between these two concepts by highlighting how in both,

“interpretive frameworks must be taken for granted,” but not necessarily so (p. 11); for Stader, both Arendtian and Gadamerian prejudices can be brought to light and active judgement by way of reflection.

Stader's (2024) argument proceeds by way of three theses on algorithms: First, “Algorithms are always embedded in purposeful contexts and cannot be defined or understood without external references that provide a basis for extensional judgements about whether something is an algorithm, what an algorithm is or what its output actually means” (p. 13), second, “Algorithms, as purpose-embedded entities, emerge from clusters of judgements” (p. 14), and third “The outputs of algorithms can only be used in a prejudiced way” (p. 15). Stader's (2024) article closes with a call for transparency, earlier defined as “the disclosure of the basic assumptions as well as the selection and the extent of data that form the algorithm” (p. 4). For the remainder of this commentary, I complicate the possibility of this sort of transparency by highlighting the notion of a cluster of judgements in a different way than in done in Stader's piece. At the same time, Stader's turn to philosophical hermeneutics and development of reflected prejudiced use can be further developed as a way to dis-cover algorithmic prejudices, even without making them transparent.

3 Calculation and Plurality

If the outputs of algorithms emerge from clusters of judgements, and not as the ossification of a single judgement, then the prejudices that appear in machine learning outputs are not *simply* the prejudices of the programmer/programming; they ossify no judgement in particular, which is not quite to say that they ossify judgements that nobody has made. This is an implication of statistical and inductive architecture underlying machine learning algorithms. The output of the algorithm effectively says, *‘The most likely judgement would have been as such.’* The judgements that the programmer makes are different from the judgements that appear as ossified in the algorithmic output, which are different than the judgements that informed the training data. In Stader's example of the Delphi AI tool, which provides users to responses for ethical inquiries, this can be seen in that the programmers have precisely never programmed a judgement about the morality of killing a tyrant. As Stader (2024) aptly explains, the creators of the program have only ossified “assumption that there *are* common and underlying rules of human moral judgement that can be made operable (and thus in principle explicit) for a machine by training it with large amounts of data” (p. 24). As such, the output that appears to me as a judgement concerning my inquiry, is not, in a specific way the ossified judgement that the programmers have made about my question. But neither is it the judgement ossified from the crowdsourced data. Stader (2024) makes this clear: “the collection of such judgements does not represent people's actual moral actions, but only their judgement in hypothetical cases” (p. 24). Even more, the output does not represent any single judgement, but as Stader notes, the *collection* of judgements. So, the algorithm does not simply make a calculation that relates to multiple judgements but turns its relation *to a plurality of judgements* into an ossification. From the start, the output is not a repetition of a judgement that has already been made,

never simply the re-presentation of the programmer, the data set, or the individual judgements from which the data set grows. Following work from a poststructuralist perspective (Coeckelbergh & Gunkel, 2023; Gunkel, 2024) that suggests the inadequacy of searching for the origin of algorithmic outputs, we see even from a hermeneutic perspective that the massive plurality of algorithms precludes their conceptualization as any simple repetition. There is no original judgement to which the output of the algorithm can be traced; in this sense, the output is an ossification of a matrix of judgements, not any single item in particular. In this sense, the algorithmic judgement is a simulacrum, the production of a copy without original.

By giving this kind of attention to the ossified judgements, not only of a programmer, but also of the judgements that produce data for the algorithm, one of Stader's main themes has been strangely inverted: calculation does not give the algorithmic output its definiteness, but rather its plurality. Only because the collection of judgements are calculated can they be taken as neither an authentically 'original' judgement nor a re-presentation. This calls into question the possibility of the axiomatic transparency that Stader calls for: even *if* a programmer could honestly and effectively disclose their prejudices, this does not bring the prejudices of the algorithm into full transparency because these prejudices are only one dimension of the manifold of prejudices implicated in the calculation; the programmer's prejudices that determine the scope of calculations only effect, but do not determine the prejudices that have produced the dataset themselves.

4 Otherwise than Transparency: Structural and Exemplary Prejudices

If algorithms are characterized by a relative (i.e., not absolute) alterity to their ossifications, then algorithmic outputs can both coincide with prior judgements, as well as produce judgements that have not appeared in their specificity before. This means that to the extent that Stader correctly calls attention to the human judgements involved in programming, transparency of these judgements alone cannot exhaust a hermeneutic for algorithmic output because they do not exhaust the 'past' of algorithms themselves. Algorithms, in other words, can surprise their programmers, not only because the programmer's biases might not be fully transparent to her, but also because algorithms ossify a plurality of judgements, only one level of which includes the judgements of the programmer.

In closing, I return to Stader's discussion of reflected prejudiced use, which I believe can be further developed to "be sensitive to the text's alterity" as Gadamer (2013, p. 282) would say, by attending to the different kinds of prejudices ossified in the algorithm, not simply recollecting the 'original' judgements of a programmer. Let us return to Stader's call for transparency. He begins by noting that "[r]eflection on conventional prejudices works by intervening in their function, in their formation, their concepts, in the connection with their basic experience; this is not possible with algorithms because the algorithmic formation is not a process that originates from the conduct of human life" (p. 23). Here, I amend Stader slightly: the difficulty of reflecting on the prejudices of algorithms does not arise because they do not originate from the conduct of human life (indeed, if they operationalize

prior judgements, then they have a necessary connection to human life), but rather, because, as Stader even writes, “[c]onsidering the assumed amount of data, they seem to be superior to our limited [human] realm of experience” (p. 24).

In Stader’s call for transparency of algorithmic prejudices, he thematizes the possibility of reflected prejudiced use according to one dimension of algorithmic alterity, which can be extended by attending more intentionally to the issue of plurality. If the ossification of human judgements in algorithms can be taken as a plurality in relative alterity to the prior judgements of the programmer and the prior judgements that produced the dataset, then we can begin to see a typology of algorithmic prejudices emerge that enriches the possibility of reflected prejudiced use. On the one hand, the algorithm ossifies the *structural prejudices* of the programmer, whereas on the other hand, it ossifies the *exemplary prejudices* of the dataset. Stader’s call for transparency is addressed toward the structural prejudices of the programmer. These prejudices set at least initial conditions of possibility for the process by which an algorithm operationalizes and calculates relations amongst data. But these conditions are only significant by their putting into play actual data, which in many cases, are themselves ossifications of judgements. Again, Stader’s Delphi example is helpful here. The prejudices that are ossified from the programmer are of a different kind than the prejudices ossified from the crowdworkers who produce the data that the algorithm operationalizes. If the prejudices of the programmer are structural, by defining initial conditions of possibility for the algorithm, then the prejudices of the crowdworkers are exemplary, in that they are taken as examples as the kind of judgement that a user wants to make in a particular instance.

Stader’s implicit emphasis on structural prejudices seems warranted, in that the programmer can make judgements about what data to include as exemplary for a dataset in the first place. And yet, it is precisely by way of the multitude of exemplary that the algorithm exceeds the scope of human judgement, by implicating a massive plurality of judgements as exemplars for my use of the algorithm. Insofar as Stader calls for reflected prejudiced use as a way of attending to, and thus actively reintegrating ossifications into the human lifeworld, this kind of reflection ought to be conceptualized in a continuous deferral that accounts for the prejudices of the programming as well as the prejudices that have produced the dataset. This multiplicity of prejudices makes transparency an impossible task, but an impossible task that correlates to the more-than-human scope of algorithmic processes.

Funding There was no funding to disclose for this project—ML

Declarations

Ethical Approval The study includes no human or animal subjects that require ethical approval—ML.

Consent to Participate The study includes no subjects who needed to give consent to participate—ML.

Consent for Publication This study includes no human participant who needed to give consent to publish—ML.

Competing Interests The author declares no competing interests—ML

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arendt, H. (2005). Introduction into Politics. In H. Arendt, *The Promise of Politics* (pp. 93–200). Ed. and intr. J. Kohn. Schocken Books.
- Coeckelbergh, M., & Gunkel, D. J. (2023). ChatGPT: Deconstructing the debate and moving it forward. *AI & Society*, 1–11. <https://doi.org/10.1007/s00146-023-01710-4>
- Gadamer, H.-G. (Ed.). (2013). *Truth and Method* ((Reprint edition). ed.). Bloomsbury Academic.
- Gunkel, D. (2024). Literary Theory for LLMs. *Newhouse Impact Journal*, 1(1), 6. <https://doi.org/10.14305/jn.29960819.2024.1.1.02>
- Stader, D. (2024). Algorithms Don't Have A Future: On the Relation of Judgement and Calculation. *Philosophy & Technology*, 37(1), 21. <https://doi.org/10.1007/s13347-024-00705-3>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.