Algorithmic Structural Epistemic Injustice

Artificial Intelligence, Unjust Regimes of Epistemic Recognition, and the Reinforcement of Ignorance[1]

Fabian Schuppert

Introduction

Contemporary discussions suggest that we currently witness an epistemic revolution in the form of widely available artificial intelligence (AI) systems. AI systems are increasingly used in a variety of contexts, shaping the way people communicate, access information, and make decisions. This essay focuses on chatbots such as ChatGPT, which are large language models that are using a form of self-supervised machine learning (which is the artificial intelligence aspect of contemporary chatbots) in order to perform natural language processing tasks, such as natural language understanding, text classification, knowledge summaries, and natural language generation. For reasons of simplicity, for the remainder of this essay the terms artificial intelligence (AI), large language models (LLMs), natural language processing (NLP) shall all refer to particular aspects of contemporary chatbots.

Nowhere is the impact of AI-based generative chatbots more evident than in education and publishing, where these tools are deployed to write essays, recommend and summarise articles, moderate debates, grade coursework, or suggest improvements to people's writing. This means that these large language models influence which knowledge is used in certain conversations and how, as well as that they pass "judgment" on which forms of expression are "better" or "more scientific".

1 I am extremely grateful to Kalia Barkai, Hilkje Hänel, Nikolaus Hoffmann, Börries Nehe, Judith Renner, and Janina Walkenhorst for insightful comments which helped improve the argument in the paper.

According to proponents of AI-based LLMs, these technologies may seem neutral, even democratizing, as they supposedly open up knowledge to everybody, promise to remove human bias and ensure a "level playing field." However, a closer examination reveals that LLMs systematically reproduce—and even amplify—epistemic injustices and stereotypes about what counts as relevant knowledge, good writing, and rational argumentation. In addition, particular fields of inquiry are presented in very particular ways, meaning that existing stereotypes about what particular forms of knowledge can do and be, are also systematically reproduced. In doing so, AI tools perpetuate existing power structures, making the privileged even more privileged, and foster a particularly insidious form of epistemic ignorance and epistemic domination.

In their existing critiques of AI, LLMs, and algorithmic processing, scholars of epistemic injustice have primarily focused on the way that social media algorithms filter public opinion and privilege certain content over other (Stewart, Cichocki, and McLeod 2022), or the role of primarily white male programmers inscribing their own biases into the code of algorithms (Hoffmann 2019; Noble 2018), as well as how digital environments shape individual's identity formation (Origgi and Ciranna 2017). While all of these issues are serious and merit further investigation, I want to focus on a slightly different aspect, namely, the internal workings of LLMs and how their in-built regimes of recognition lead to particular forms of epistemic oppression, an issue also explored by recent contributions by Miragoli (2024) and McInerney (2024).

The Hidden Curricula of LLMs: Learning the Norms and Stereotypes of an Unjust World

Prior to being released for its intended use, all LLMs are trained on vast datasets, often scraped from published material, publicly available databases, and other forms of data that have – in one way or another – already been filtered by human hands. Instead of being a neutral pool of resources to draw from, these datasets reflect the dominant conceptions of how our world works, including problematic assumptions about what matters in the world, what typical gender roles are, and the qualities of different ethnic groups. Because of where LLMs like the chatbots this essay is concerned with are developed, there is a distinct Western bias in these models. This is hardly surprising as the LLMs are simply trained on the hegemonic norms, beliefs, and worldviews of their time, which means that also the stereotypes and blind spots of the

hegemonic view are part of the training curricula. Through training, the LLMs are also supposed to be taught what counts as "valuable" knowledge, "proper" language use, "rational" argumentation, and "scientific" rhetorical style.

In other words, AI models are trained and developed on particular regimes of recognition, which bestow a particular status on some forms of knowledge, while other forms of knowledge are treated differently. What do I mean by this?

A regime of recognition is a (more or less) fixed order of who or what will get what kind of recognition for what kind of property, status or action. So, children in Germany for instance often get recognition for being "well-behaved", which is a normative marker of esteem recognition, when they say please and thank you (the actions which trigger the ascription of a form of recognition), no matter whether that really tells one anything about the kids overall behaviour. Not all forms of recognition regimes are per se morally problematic. In fact, bestowing esteem recognition and respect onto others is a central feature of human interaction (McBride 2013). But many reasons for bestowing a particular form of recognition onto someone or something are morally problematic, as in the statement "This was a really good throw for a girl". The actual esteem given initially seems positive, i.e. "a good throw", but it immediately becomes clear that the qualifier "for a girl" not only raises questions about the reason underlying the given esteem recognition, but also changes the nature of the esteem given: it is not a "good throw" simpliciter, it is a "good throw *for a* girl", which sends all kinds of messages about the status of girls when it comes to throwing. To be clear, this does not mean that this is a question of distributing recognition equally, since unequal recognition fulfils an important function in human interaction, as it is used to signal one's feelings and beliefs.

When it comes to how LLMs are trained, it is by no means the standard case that some forms of knowledge get a lot of recognition, while others get very little. While this might be a case of problematically recognising different forms of knowledge, it is much more common that different forms of knowledge also get different kinds of recognition, which reflect and perpetuate problematic stereotypes about what specific forms of knowledge are and can be.

It is therefore not the case that AI generated essays on the status of "indigenous knowledge" would be either using overtly negative language or directly express the view that indigenous forms of knowledge are worth less than other

forms of knowledge.[2] Instead statements on indigenous knowledge regularly included adjectives such as "spiritual", "traditional", and "cultural", while essays on feminism include terms such as "care", "subjective", and "emotion".

Needless to say, all of the above terms could form part of excellent treatments of indigenous knowledge or feminism, since indigenous scholars do indeed often highlight the role of traditions, cultural values, and spiritual worldviews, while feminists have long advocated for the importance of subjective experiences of oppression, the centrality of care work, and the role of emotions in reasoning. The problem with the examined essays is that these terms are used in juxtaposition to terms like "rational", "scientific", "established" and with insufficient critical contextualisation, so that implicitly indigenous knowledge apparently never can be "rational". Put simply, indigenous knowledge is put into a box with a very small number of labels attached, labels which reproduce hierarchies of esteem, as long as being rational, scientific and established continue to be seen as the ideal norm that all other forms of knowledge and research aspire to. The issue is thus not that indigenous knowledge is not esteemed or recognised in LLM-generated content, but that it is esteemed and recognised only in very particular ways and that it is portrayed in reference to what counts as "normal proper knowledge".

Interestingly a similar pattern can be observed when LLMs are tasked with "improving the writing" of a particular text or assessing the "rationality of an argument". "Good writing," as assessed by LLMs, is often synonymous with the conventions of standardized English and the pseudo-academic prose found in executive summaries of scientific reports, which try to sell a particular solution. Hence, there is a particular tone of expression, somewhere between marketing pitch and policy report summary that is seen as the gold standard of "good writing". Norms of rational argumentation are typically equated with forms of argumentation that privilege deductive reasoning and that use marker terms such as "objective" and "evidence-based". Expressions such as "subjective experience" and "informal norms" were taken out of pre-written inputs and replaced by LLMs by expressions such as "empirical evidence" and "well-known facts", which clearly express a different thing than the initial expressions.

2 The following empirical observations are based on a three-month trial using different prompts for three widely used chatbots: ChatGPT, Google Gemini, and Microsoft Co-Pilot.

When it comes to hidden training curricula of standard LLMs, two things happen: first, the algorithms are trained on hegemonic norms and worldviews, which come with an in-built regime of recognition that distributes esteem markers to all sorts of agents, views, events, and states of affairs. While a critic might point out that implicitly passing on markers of esteem is a necessary aspect of any curriculum, it makes a huge difference on which data and knowledge a model was trained, and which recognition order has been passed on. As is well-established, predominant views of history, politics, and society are often racist and sexist, which means that algorithms trained on this data suffer from the same problem. LLMs do not merely reflect but encode and reinforce prevailing norms—norms that are themselves the product of long-standing histories of inclusion, marginalisation, exclusion, and epistemic silencing.

Taking privileged norms and worldviews as the "neutral" knowledge base for the most widely used chatbots and LLMs leads to a perpetuation of privilege. There is a distinct epistemic injustice already at the training and input level of most LLMs, since the in-built recognition regime that structures what counts as what expresses a whole range of identity stereotypes and problematic value judgments which affects the normative status of different kinds of knowledge and agents as knowers.

The internal workings of LLMs: Recognition orders, social capital, and epistemic authority without sufficient reason

The epistemic injustice at the training stage is not the only input injustice of most LLMs. The second input injustice occurs through how LLMs like ChatGPT actually work. While many people believe that the latest generation of LLMs blaze through the entire vast data that can be found on the internet in order to generate an answer to a query within seconds, the truth is actually quite different. Similar to how search engines like google operate, ChatGPT, Gemini and co. use shortcuts, by operating with directories of webpages, which through web-crawling have been deemed to be statistically particularly relevant. In other words, when one sends a short prompt to a chatbot (e.g. "What is climate change?") the program does not search the entire web and then passes judgment on which information to provide as an answer; the search in the web would not only take much too long, but also the program is not able to pass an evaluative judgment that would be sensitive to all sort of different queries.

The algorithms at work in contemporary LLMs of course do pass value judgments in some sense, since what the algorithms are trained to do is to draw their material from the pages/sources of information which the underlying metric (which was already used as part of the web crawling) to be statistically the most relevant. All that matters during the process of web crawling and during the actual query to the LLM is that the information on a website has been "easily discoverable" through web search optimization and that the site is "well networked", that is it is connected by a myriad of links from other webpages.

There are two distinct problems here with the internal workings of LLMs: first, the *regime of recognition* that underpins the determination of which sources of information are relevant, is based on a form of *privilege and social capital*; second, *the ascription of epistemic authority* that AI-based LLMs happens *without an appropriate set of good reasons*.

Let me start with the first problem, the regime of recognition which is based on a form of privilege and social capital. As mentioned above, LLMs rely on a register of webpages that were discovered through web crawl and that were deemed to be statistically particularly relevant. This means that these webpages get a lot of hits during the web crawl process, in part because these webpages were optimised for this kind of crawl (e.g. search engine optimisation – SEO). It is the same system that advanced search engines like google rely on. What this means is that you can game the system and make sure that your webpages get a lot of hits. One particularly easy way to do this is by spending money on it, which is unsurprisingly something that privileged actors can do much more easily than other actors. It also helps if you have a large portfolio of sites which constantly refer to each other, but which do not look like they come from the same tree. Springer Academic Publishing is really good in this area, which means that articles published in their journals, or chapters published in their books get a lot of hits on google and they also feature heavily when chatbots are tasked with writing a scientific essay on a particular topic. The actual quality of the article or chapter is irrelevant, since the LLMs' esteem recognition for a particular source of knowledge is based on its statistical hit-rate.

This is where the second part of the first problem comes in, namely the importance of "social" networks, that is, many external sites which refer to the page one wants to optimise for discovery. For external sites to refer to one's own page, though, one needs to be well connected. Just like with human beings then, social networks matter, which means this case could be interpreted as a case related to what Bourdieu called social capital, namely, the benefits in-

dividuals and groups derive from social connections, networks, and relationships (Bourdieu 1986).

In order to be statistically relevant, then, it helps if one is economically privileged, as well as privileged regarding the social capital one has at one's disposal. Because of how LLMs are designed, they rate these privileges highly and reward well-connected and search optimised sites with more esteem recognition, which makes it more likely that any future query will be answered by a chatbot using the information provided by privileged agents.

This reproduction of privilege is in many ways subtle but extremely profound. Precisely because LLMs are built to award esteem recognition to those pages that get a lot of hits in their metrics, LLMs are extremely liable to be tricked into assuming that the sites which are best at being highlighted by a web crawl and a chatbot prompt, are also the best sources of knowledge. This is precisely the second problem, i.e. the ascribing of epistemic authority without sufficient reason.

When processing a prompt, chatbots do not care about the academic credentials of the author of the webpage the chatbot draws the answer from. This can be on the one hand refreshing, since not everything coming from an Oxbridge educated person gets automatically treated differently than everyone else's views. But on the other hand, since LLMs *only care* about statistical relevance there is no quality control outside of statistical relevance. There is no quality control mechanism in the sense of an algorithm passing evaluative judgment on what a most complete and sophisticated answer should look like. Instead, whatever is statistically most relevant is taken to be objectively good information, simply because it is statistically most relevant. Statistical relevance on its own, though, is not a good enough reason for ascribing something epistemic authority, especially in light of the reproduction of privilege and social capital, which I described above.

The Epistemic Injustices of LLM Outputs

For a long time, the importance of statistical relevance within the internal workings of LLMs was directly reflected in the output of Generative Pre-Trained Transformers (GPTs). As Emily Bender and colleagues (2021) provocatively asked as recently as 2021, will even the most sophisticated GPTs ever be more than "stochastic parrots" that – while being able to process natural language in such a way that one can have a conversation with them – will

only ever repeat information back to their human interlocutors that they were initially trained with or that they copied and pasted from a webpage?

As Konstantine Arkoudas (2023) argues, Bender's question has been answered with a resounding "yes", since ChatGPT and the latest generation of chatbots are much more than stochastic parrots. The latest GPTs can come up with their own views and arguments, combining different sources of information and developing (relatively) independent conclusions, which goes even as far as inventing concepts or pieces of literature that the GPT has inferred should exist. As Arkoudas (2023) points out, this is a huge technological advancement, but it does not mean that the latest GPTs are indeed "intelligent reasoners", since the latest chatbots still struggle with exercises in logic, text-based maths problems, and the identification and application of norms. At their core, the latest chatbots are still LLMs, which try to identify the correct answer by relying on statistical relevance and the regimes of recognition described above.

Considering the different epistemic injustices that we could identify at the input stage of LLMs, it is hardly surprising that we can also observe epistemic injustices at the output stage. In the following, I want to focus on three issues, all of which are directly related to the perpetuation of privilege and ignorance discussed above: epistemic oppression, silencing, and toxic deficiency.

Epistemic oppression: Building on Kristie Dotson's (2014) account of epistemic oppression, one can identify how the recognition order underlying statistical relevance which is tied up with privilege, social capital, and the unjust background of a long history of biased framing and naming, leads to a situation in which marginalised groups and forms of knowledge are systematically oppressed. Because of how LLMs are set up and work, with their unjust input and their focus on producing answers that reflect statistical relevance, the output of the latest chatbots fails to adequately reflect and incorporate the knowledge and insights of many marginalised groups. For example, it is not the case that on the internet there are no sources on Black feminist thought, but when one looks at the results of chatbot queries regarding important 20th century social thought, or even feminist thought, these sources – most often provided by Black female writers – are strangely absent. This means that in the realm of LLMs the contribution to knowledge by Black feminist writers is made virtually impossible.

Toxic deficiency: Toxic deficiency is a phenomenon aptly described by Martin Miragoli (2024, p. 9) who points out marginalised groups are not just harmed as knowers (as in the case of epistemic oppression) but also as knowledge

seekers. Because of the epistemic injustices occurring at the input stage of LLMs, LLMs have in-built "hermeneutical lacunae" (Miragoli 2024, p. 10), in that within the shared hermeneutical resources of the LLM universe concepts and understandings pertinent to the lived experiences of marginalised groups are missing. This leads to a toxic deficiency in the outputs of LLMs, in which the hermeneutical lacunae are translated into "answers" which are deficient conceptually and which try to reinforce hegemonic norms onto the person who seeks knowledge. A young trans-person trying to make sense of their identity and feelings will often encounter said deficiency, while at the same time being confronted with heteronormative binary gender stereotypes, which can aptly be described as toxic, since they hurt the young person's sense of self. The same is true for a Black person, who wants to know more about the history of Black civilisation which in mainstream history-writing has been rendered invisible. This invisibility is on the one hand a deficiency and on the other hand it is covered by a toxic array of presentations of history in which entire parts of the world are treated as blank spaces which only come into view when European imperialism demands it. Black civilisation is not only conspicuously absent, it is negated through toxic colonial imaginaries.

Silencing: A third issue is silencing (Dotson 2011), which happens mainly when we turn our attention to the linguistic preferences of contemporary chatbots, which do not give all speakers the same recognition, because some expressions and ways of conversing are deemed to be less desirable and sophisticated. By privileging dominant norms of writing, speaking, and rationality, AI systems help maintain the social and epistemic dominance of those who already wield it. Students, professionals, and creators who already write and argue in the expected style are further rewarded, while those who do not are penalized and silenced. The privileged thus become even more privileged, as their modes of expression are held up as universal standards.

With the spread of AI-supported LLMs into various areas, this problem becomes even more pronounced: automated grading systems in schools can systematically disadvantage students who speak in dialects or bring nontraditional modes of argumentation to their work. In publishing, algorithms that recommend or surface content based on "readability" or "argument quality" quietly filter out diverse voices.

Following the description of these three kinds of injustice, two very important points need to be made: first, the observed injustices are not accidents or a malfunctioning of otherwise well-working systems, but they are a systematic design-feature (Ruiz and Sertler 2024; Miragoli 2024); second, the kinds

of epistemic injustices observed are not based on personal interactions, but forms of systemic epistemic injustices.

Let me start by stressing that all three injustices at the output stage are systemic issues, as Miragoli (2024) convincingly argues regarding toxic deficiencies:

> "That is: because the hermeneutical lacuna present in our shared online resources is a direct consequence of the very functioning and training practices of ML-based AIs, epistemic injustices of a hermeneutical kind […], are not just unlucky byproducts of developers' biases, but a *systematic feature of the design of AI design*." (Miragoli 2024, p. 11)

The same is true when it comes to epistemic oppression, since certain groups and sources of knowledge are systematically excluded at both input stages, as well as at the output stage, making it thus impossible for some knowers to contribute to the shared wisdom of society. Contemporary chatbots/LLMs thus produce epistemic injustices *by design*, and not just as an unintended side product.

Secondly, the kinds of epistemic injustices observed are not based on personal interactions, but forms of systemic epistemic injustices, which is what distinguishes the cases described here from most of the cases described by Dotson (2011; 2014). Unlike in interpersonal cases, where for instance testimonial injustices are based on identity prejudices against a particular speaker who is harmed in their capacity as a knower, in the case of silencing LLMs simply process an available text trying to improve its writing or grade student essays, which cannot be clearly attributed to a particular group or speaker; LLMs thus commit an epistemic injustice in a slightly different way: LLMs come with in-built recognition hierarchies which assign epistemic status and epistemic validity based on token markers and statistical relevance, which are supposed to generally signal what should count a sound argument or as a proficient use of language. The same recognition regimes in combination with metrics of statistical relevance are used by LLMs to determine what an objective and truthful answer should be.

The Cultivation of Epistemic Ignorance

Ironically, in their quest to produce "objective" or "truthful" answers, LLMs actually generate the opposite; these tools cultivate a particular kind of epistemic ignorance—one that is both pervasive and difficult to detect, because it follows in the footsteps of hegemonic views on knowledge, norms, and history, which have long claimed to simply advance the truth and nothing but the truth.[3]

One particularly pernicious effect of widespread AI-based LLM use in contemporary chatbots is the homogenisation of knowledge. LLMs prune away stories and artefacts that sit uncomfortably with the established consensus. Therefore, chatbots prompted to write short five-page essays about African and European history, paint very different pictures of what happened in the 20th century. The essay on African history focused on civil strife, military coups and economic challenges in 20th century Africa, while Europe was presented as the beacon of human rights and solidarity, with a special positive mention of the European Union. This is of course how much of mainstream media represents both continents, so why should this be a particular concern for scholars of epistemic injustice? The reason is that we can observe here how AI tools perpetuate epistemic injustices and forms of misrecognition systematically, despite these tools having been trained to avoid identity prejudices based on superficial markers. Developers of LLMs forcefully argue that their algorithms avoid identity-based biases and use "neutral" metrics for assessing the "value" of a piece of information, that is, statistical relevance and a wide network. As shown above, however, these are not sufficiently good reasons to ascribe something general epistemic authority.

LLMs thus promise truth and objectivity, but what they generate is a false or weak objectivity (Harding 1995). Users may believe that algorithmic recommendations and assessments are neutral, when in fact they are deeply implicated in longstanding hierarchies of race, class, language, and knowledge, precisely because the markers of statistically relevant value and esteem are in themselves coded in racist, sexist, and classist ways.

This reproduction of privilege is in many ways subtle but extremely profound. It ensures that those at the centre of cultural, linguistic, and epistemic power have their perspectives normalized and reaffirmed, while those at the

3 Much of the "knowledge"-base of LLMs is itself guilty of what Gaile Pohlhaus Jr. (2012) has called wilful hermeneutical ignorance.

margins are forced to translate, assimilate, or remain unheard, while still being oppressed at the same time. This problem is particularly pronounced in cases where long-standing practices of silencing, cultural exclusion and historic whitewashing exist and have led to a state, in which alternative accounts of history have been systematically oppressed.

Because latest generation chatbots operate under the assumption that there always is one correct and objective answer, and because the algorithms equivocate statistical relevance with epistemic authority, contribute to a homogenization of knowledge. In this regard, Google and ChatGPT are very much alike, since they try to find "canonical" results, assuming that reducing complexity is in this case a good thing. This is precisely where LLMs ought to learn from Harding's (1995) nuanced account of strong objectivity, which aims to block "might is right" accounts of knowledge. Maximising objectivity requires a wide and diverse knowledge base, which is where the regimes of recognition of existing LLMs fall down, because they breed ignorance through advancing a narrow understanding of objectivity.

Reducing Epistemic Injustice through training LLMs?

At this point, critics might object that my view of LLMs is too negative, since ChatGPT and friends do bring a range of benefits: people do use chatbots for improving their English/their writing, and it is possible to train latest generation chatbots through prompting them repeatedly. While it is of course true that chatbots can help people in all sorts of ways, this does not change my primary point, namely, that LLMs by design bring a range of systemic epistemic injustices with them.

In addition, the idea that conversing with chatbots is a good way to train them to avoid future injustices is misguided and a bit naïve. The reason for this rather harsh assessment is again the way that chatbots actually work. Chatbots are indeed designed to be further trained through conversations with human end-users. It's what is called Reinforcement Learning from Human Feedback (Christiano et al., 2017), which means that chatbots adapt to what their interlocutor is searching for. Therefore, it is entirely possible that a Black feminist might make their chatbot sound more like a Black feminist chatbot, by asking all sorts of challenging questions and forcing the LLM running the chatbot to look for particular information. But what happens here has nothing to do with reducing epistemic injustice.

First of all, training a chatbot is only possible if one already possesses the relevant knowledge of what one is looking for. In other words, one needs to be aware of the deficiencies, in order to be able to train.

Second, in training the chatbot, one actually provides free labour and one offers one's epistemic capacities freely for capitalist extraction. Basically, the Black feminist would have turned themselves into a free resource.

Third, chatbots do not adjust their general answers, but only in this particular conversation, which means what we can observe here is more a case of mirroring or disingenuous talking back, which basically would function like a social media echo chamber.

What remains untouched is the underlying structure, in which hegemonic statistical dominance is what matters. This means that what really works best for changing chatbot outputs, is to "flood the zone with shit", which is a problem one can already observe. There have been repeatedly cases in which AI-supported LLMs have been tricked by AI-generated content, leading to a self-reinforcing circle of fake claims, masquerading as objective fact. As Microsoft's disaster with their chatbot Tay showed, feeding a chatbot huge quantities of certain forms of information can make a difference, but this works best if the information provided is extremely simple, which runs counter to what a feeding of information in the name of epistemic justice would require.

Conclusion

My aim in this short essay was to highlight a few of the distinct epistemic injustices that AI-supported LLMs raise. While contemporary LLMs are trained so as to avoid clearly racist, homophobic, and sexist language, the regimes of recognition with which LLMs operate at the input stage lead to the reproduction of privilege and ignorance. As a result of this unjust background, the output of LLMs often leads to forms of epistemic oppression, toxicity, and silencing. However, these pernicious effects are not unfortunate accidents or a malfunctioning of the highly sophisticated AI-supported LLMs, but they are systematic features by design (Miragoli 2024; Ruiz and Sertler 2024). Therefore, hopes that end users can train LLMs in such a way as to produce epistemically just chatbots are overly optimistic and naïve. While making LLMs' internal workings more transparent is certainly a step in the right direction, it is the systemic design and its regimes of recognition that really is the root problem.

References

Arkoudas, Konstantine (2023). "ChatGPT is no stochastic parrot. But it also claims that 1 is greater than 1." *Philosophy & Technology* 36.3: 54.

Bender, Emily M., et al (2021). "On the dangers of stochastic parrots: Can language models be too big?." *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. March. pp. 610–23. https://doi.org/10.1145/3442188.3445922.

Bourdieu, Pierre (1986). "The forms of capital". In: Richardson J (ed.) *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood, pp. 241–58.

Dotson, Kristie (2011). "Tracking epistemic violence, tracking practices of silencing." *Hypatia* 26.2: 236–257.

Dotson, Kristie (2014). "Conceptualizing epistemic oppression." *Social epistemology* 28.2: 115–138.

Harding, Sandra (1995). "'Strong objectivity': A response to the new objectivity question." *Synthese* 104.3: 331–349.

Hoffmann, Anna Lauren (2019). "Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse." *Information, Communication & Society* 22.7: 900–915.

McBride, Cillian (2013). *Recognition*. Polity.

Mclnerney, Tomás (2024). "The Algorithmic Construction of Epistemic Injustice." *Socio-Legal Studies of Epistemic Injustice and Spaces and Places*.

Miragoli, Martin (2024). "Conformism, Ignorance & Injustice: AI as a Tool of Epistemic Oppression." *Episteme*: 1–19.

Noble, Safiya Umoja. *Algorithms of oppression: How search engines reinforce racism*. New York University Press, 2018.

Origgi, Gloria, and Serena Ciranna (2017). "Epistemic injustice: The case of digital environments." *The Routledge handbook of epistemic injustice*. Routledge, 303–312.

Pohlhaus Jr, Gaile (2012). "Relational knowing and epistemic injustice: Toward a theory of willful hermeneutical ignorance." *Hypatia* 27.4: 715–735.

Ruíz, Elena, and Ezgi Sertler (2024). "Injustice by design." *Inquiry*: 1–24.

Stewart, Heather, Emily Cichocki, and Carolyn McLeod (2022). "A perfect storm for epistemic injustice: Algorithmic targeting and sorting on social media." *Feminist Philosophy Quarterly* 8.3/4.