Whoever Controls Language Models Controls Politics

Hannes Bajohr

The world of artificial intelligence thinks both big and simple at the same time – and has done so from the very beginning. When the workshop that launched the concept and the field of 'artificial intelligence' was held at Dartmouth College in the summer of 1956, the self-imposed task was to figure out "how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves" (McCarthy et al. 2006). The expected duration of the project: two months.

Almost seventy years later – on March 22, 2023 – an open letter was published on the website of the longtermist Future of Life Institute, with, at the time of writing, more than thirty-three thousand signatures ('Pause Giant AI Experiments: An Open Letter' 2023). It included the likes of Elon Musk and many renowned AI researchers, and called for a moratorium on the development of large language models (LLMs) for at least six months. Systems such as ChatGPT, the authors claim, have now become too powerful and too dangerous, and "profound risks to society and humanity" are posed by "human-competitive AI" (ibidem). Until there is agreement on how to regulate this complex, all AI labs should refrain from further research.

If Dartmouth spectacularly underestimated how difficult the automation of intelligence would prove to be, the open letter is equally bombastic in drawing the wrong conclusions from the power of current language technology. First, even today, Dartmouth's goal remains unmet: for all their successes, ChatGPT and other LLMs do not operate at "human level" by any meaningfully robust measure (Mitchell and Krakauer 2023; Floridi 2023). Such fantasies fuel AI hype,* a tendency that has afflicted the industry since its early days (Aggarwal 2018), and which ultimately serves the companies that sell it (Luitse and Denkena 2021). What better proof of a developer's power than their ability to distribute a product that could potentially destroy the world? Insiders soon speculated that the goal was in fact to subvert the industry's long-held rule of open research and to continue working in secret for the suggested six months. And indeed, Musk announced the launch of his own AI company, called X.AI, on April 14; his signature very soon no longer seemed to count for much (Jin and Seetharaman 2023).

Second, however, and more importantly, the letter also speaks to a disastrous understanding of the interplay between technology and politics, both in terms of its dangers and the means to address them. While the fear that Al-generated text could flood information channels with falsehoods and "propaganda-as-a-service" (Bagdasaryan and Shmatikov 2022) is entirely valid, the letter is otherwise driven by apocalyptic fantasies about the total replacement of humans by machines and the "loss of control of our civilization" ('Pause Giant AI Experiments: An Open Letter' 2023).

This "x-risk" (the "x" standing for "existential," Bostrom 2002) is the primary concern of 'longtermists'– a libertarian, transhumanist, and ultra-utilitarian school of thought that gives possible future humans an incomparably greater moral weight than actual present ones (Torres 2021). Its proponents, to whom Musk, too, feels a close affinity, think in millennia and in terms of narrow utility maximization. For this reason, the threat of a hyper-intelligent machine – "we are all going to die" stated one particularly alarmist op-ed (Yudkowsky 2023) – worries them much more than, for example, the immediate damage caused by climate change, social injustice, or poverty – problems that, to them, are either non-issues or will be resolved through the very AI they perceive as existential threat (Klein 2023).

A Democratic Disaster

The risk LLMs like ChatGPT pose, however, is not so much the technical catastrophe of malicious computers. Much more concretely, language models threaten democratic disaster – through the privatization of language technologies as the future site of political public spheres, and the purely technocratic approach to solving its problems. This is where politics and civil society need to step up, and where democratic political theory needs to shift its focus.

Technological progress over the last few years has shown that the more data an AI system is fed, the more powerful it becomes – but also the more expensive it is to

Especially dubious but popular is Bubeck et al. (2023): AI hype at its finest that is methodologically and rhetorically flawed, but embraced by some media and researchers.

develop. While it is difficult to predict future trends, it is not impossible that competition to build ever larger models could result in only a handful of companies remaining in the race (Vincent 2023a), such as OpenAI/Microsoft, Google's Deepmind, or Anthropic. And while there are open-source efforts to 'democratize' language models, they have yet to prove successful compared to Big Tech (after all, most people will not train their own LLMs on their home computers, but will rely on a large company's packaged and serviced product). And at least for now, smaller non-commercial ventures and universities play a negligible role in achieving current scale and performance records.**

Two issues appear to be the most worrisome. What we may be facing, and this is the first issue, is an oligopoly that concentrates language technologies in the hands of a few private companies. These powerful players do not exert dominance over any old product. Rather, it is the future of political opinion-forming and deliberation that will be decided in LLMs, which poses a direct challenge to democratic political theory.

Why this is so can be shown by looking at what until now was seen as the biggest political problem with AI systems, their biases (Bender et al. 2021). LLMs model their output on the texts they have been trained on, which largely comprise the writing found across the Internet and other sources – including the prejudices, racism, and sexism that constitute much of this content. Because "raw data' is an oxymoron" (Gitelman 2013) – that is, data is always situated within a particular context, made for specific purposes, and shaped by the tools and systems used to generate, capture, and represent it – and because reality itself harbors a "world bias" (Pasquinelli 2019, 9) – that is, inequities in society are simply reflected, repeated, and reinforced in data meant to represent a 'neutral' stance – language models are inherently ideological, even in their ostensibly 'innocent' state of conception. This is so even if there is no conscious or malicious intent on the side of their creators.

But there is more: attempts at debiasing to achieve a 'neutral' outcome – trying to 'de-ideologize' LLMs, as it were – are always in vain, regardless of which end of the process is addressed. One can either censor the output, as is done (to some degree) with ChatGPT and its implementation in Bing (running the risk of rendering it unusable; Apprich forthcoming). Or, as is also practiced, one can sift through the input – the dataset – to remove undesirable components before training commences (Miller et al. 2022). Both filtering the results and curating the dataset amount to correcting the model based on a norm, a view of a better world, which is an eminently political choice. De-ideologizing AI thus necessarily involves formulating a social vision – and is thus again ideological.

ChatGPT happens to represent more progressive values, and conservative media have been quick to get excited about 'woke AI'"*** It is not unlikely that this 'progressiveness' is just due to PR considerations: sexist insults, extremist political positions, or racist output simply have a negative impact on tech companies' bottom lines. But even if one buys into the idea that true conviction stands behind OpenAI's choices, neither the uncurated nor the curated versions of LLMs are 'value-free.' To repeat: AI is always ideological (Bajohr 2021; Weatherby 2023).****

For this reason, it should concern us that decisions about the social vision that language models articulate are in the hands of a few companies not subject to democratic control and accountable to no one but their shareholders. They thus become, to misappropriate a term from philosopher Elizabeth Anderson, "private government"

^{**} At the time of writing, there are at least some rumblings that the big players are concerned about open-source models. It is still hard to say whether these concerns are justified. While it is correct that models like LLaM A have produced some encouraging results, there are two facts that should make one somewhat skeptical about a golden open-source future: The first is that LLaMA was leaked from Meta, and is thus only secondarily a product of free development; instead, it is "stand[ing] on the shoulders of giants" who are ultimately able to provide the necessary groundwork (Heaven 2023). The second is that "foundation models" (Bommasani et al. 2021) are increasingly part of a licensing economy in which the responsibility and 'servicing' for the underlying model will be a selling point, favoring large providers over open-source ones. At the U.S. Senate hearing on AI regulation on May 16, 2023, OpenAI CEO Sam Altman agreed that "there will be a relatively small number of providers that can make models," suggesting, however, that this would be positive for effective regulation: "The fewer of us that you really have to keep an eye on ... there's benefits there" (Zakrzewski et al. 2023).

This concept was also invoked by Musk as a reason for starting his own AI firm (Perrigio 2023a).

[&]quot;Ideology," here, is not to be understood as value judgment. It simply refers, as political philosopher Judith N. Shklar put it, "to political preferences, some very simple and direct, others more comprehensive.... In no case is there any effort to use the word 'ideology' as one of simple opprobrium. On the contrary, it may well be doubted whether political theory ... can be written without some sort of ideological impetus. Nor is there any reason to feel that the expression of personal preferences is an undesirable Haw. It must seem so only to those who equate objectivity with remoteness from their own experiences and especially from those they share with their contemporaries. However, if one thinks of ideology as merely a matter of emotional reactions, both negative and positive, to direct social experiences and to the views of others, it is clear that ideology is as inevitable as it is necessary in giving any thinking person a sense of direction" (Shklar 1986,4; see also Bajohr 2020).

(Anderson 2017).***** At first blush, this may not sound so new. "Artifacts have politics," as Landon Winner put it (1980), and so do digital ones: simply through the way it makes information accessible, Google has already had an outsized influence on what appears as reality to users (Noble 2018). However, with the emergent private government that is capitalized machine learning, an even deeper capture has taken hold. For the product of AI companies is the main resource that makes for a vital democracy: language. It is language through which we negotiate political alternatives at the only level where this is possible – the political public sphere. With LLMs, instead of debating what kind of world we want to live in, that decision is already made even before a single word has been uttered, because the language at one's disposal has itself already been subjected to a preliminary political decision. The more language produced by such models permeates the finest capillaries of everyday life in the future, the more dire such an outlook must seem.

Machines of Epistemic Injustice

It provides no comfort that such LLMs can of course also be steered toward the right, as computer scientist David Rozado recently showed by creating *RightWingGPT* (Rozado 2023). In fact, this points to a second worry from the vantage point of democratic theory. A future in which a conservative language AI coexists with a progressive one would not lead to some kind of balance or trajectory toward ever more nuanced positions. Nor would many factions represented by LLMs constitute a wholesome "variety of sects dispersed over the entire face" of the Internet, to cite James Madison's republican panacea against the dominance of one group over another (Madison 2003,45). For it would no longer be sects or factions talking to each other, but modeled speech itself. Immediately, political LLMs would eliminate the discussion among social groups whose conflicts ideally contribute to the formation of the opinions of an informed public. Instead of exchange, there would be only the reinforcement of already existing opinions; unlike the much-vaunted echo chambers of social media, it would not even be people who set the parameters of that discussion, but a complex system of natural language processing and profit-driven private corporations.

The detrimental effects on democratic politics in particular can be illustrated by an argument posited by political philosopher Judith Shklar. She held that the most important duty of liberal democracies is to structure their public institutions and public forums so that the voices of the marginalized can be heard. This is not simply a moral imperative, but a democratic one. Since hegemonic notions of justice, equity, fairness, and so on are positive concepts, and, as such, limited to cases these concepts explicitly 'allow for' – that is, cases the majority can imagine as relevant – it is essential to listen to negative insights, that is, cases that slip through the cracks of the official conceptual matrix. Sometimes something is not even thought of as an injustice until it is pointed out as such by those affected by it.

The "sense of injustice" (Shklar 1990, 83) that the marginalized articulate – the immediate feeling of injury preceding all explicitly formulated concepts of justice – is not only a cry for its concrete alleviation; it also expands the democratic project by broadening the notion of what is understood as a possible injustice. Hearing the marginalized means taking them seriously as an epistemic resource and at the same time including them as citizens and thus representing more people in the polity. Miran-da Fricker (2007) has termed the phenomenon whereby citizens cannot shift the frame of what constitutes justice in a given democracy because they remain unheard and thus unincluded "epistemic injustice." The hegemonic view that LLMs formulate and encode does not allow for this epistemic correction that comes from listening to the sense of injustice. Thus, LLMs are by their very structure machines of epistemic injustice. This turns into a serious problem if, in the long run, LLMs themselves become a surrogate, a 'synthetic' public sphere. As AI systems generate more and more of the texts that populate our discourse – which seems highly likely – the proportion of discourse produced by humans may steadily decrease. Because language models are difficult to

I say "misappropriate" because Anderson was referring to companies' power to regulate the lives of their employees when they, for instance, set standards for their speech even in their private time (Anderson 2017, 39). With LLMs, one need not work for a company to be affected. Nevertheless, the limits on and redirection of speech have a similarly quasi-governmental quality.

change once trained – and because they infer norms about the future from facts of the past (O'Neil 2016; Eubanks 2017), there is a danger of what Bender et al. have called a "value lock" (2021). This means that opinions, values, norms, and tendencies that are otherwise open to modification through discourse or ideological combat, including minorities' sense of injustice, become fixed in place due to the system's inability to adapt. No amount of discussion or hegemonic struggle can alter these baked-in values; the result is a technologically-induced political stagnation that includes an increasingly narrow epistemic horizon. Thus, the content of contributions to such a synthetic public sphere is not only predetermined by technological systems and capital interest, but also bound to remain the same, as whatever engagement they encounter runs parallel to and is unaffected by non-LLM public discourse.

For these considerations, it does not matter whether one follows what political philosophy calls the deliberative or the agonistic approach to democratic theory. From the standpoint of deliberative democratic theory – which understands politics as the process of collective reasoning among citizens – the "requirement of free deliberation" depends on the "discursive quality of the contributions" to the public sphere and the possible inclusion of all citizens (Habermas 2022, 150), both of which would be at stake in a synthetic public sphere generated by ideologically predetermined and value-locked LLMs. From the standpoint of an agonistic political theory – which thinks of politics as a domain of struggle and contestation that might not result in any clear-cut consensus – such a synthetic public sphere would likewise eliminate the "legitimate political channels for dissenting voices" that translate the "struggle between opposing hegemonic projects" into the channels of democratic agonistics (Mouffe 2005, 21).

Again, this need not be the result of malicious intent. A good example of how the engineering spirit of 'fixing things with technology' may in fact squash democratic debate, both deliberative and agonistic, is the study by Argyle et al. (2023). The engineers attempted to "improve the quality of divisive conversations" by interposing an AI system between the exchanges of two debating parties that restated their positions in more "neutral" language. As the authors believe in seemingly Habermasian fashion, "improving the quality of political discourse" – that is, an enforced civil tone – "will have broader benefits related to social cohesion and democracy" (Argyle et al., 3). Not only is this solution an example of techno-paternalism, however, since it does not respect the deliberate communicative choices of the participants in the debate but simply 'fixes' them. It is also, projected onto a larger scale, for instance, as a feature in messengers or discussion forums, very much not neutral, but rather again the result of a prior decision about what neutrality and civility – i.e., the limits and conditions of discourse – entail.****** And this, again, is a deeply political choice.

Last Resort: Communization

Whoever controls language models controls politics. The regulation of AI – which Big Tech ostensibly calls for, but only as voluntary self-supervision by the industry as a type of "regulatory capture" (Vincent 2023b)****** – cannot settle for mere ethical guidelines (Stark 2023). As the firing of Timnit Gebru and Margaret Mitchell from Google in late 2020 shows, "ethics departments" are, at best, a fig leaf of accountability that companies can discard at will (Simonite 2021).

To be sure, it is absolutely necessary to create legal regulations (Noble 2018), beginning with prohibiting using AI for deceptive purposes and banning LLM training without the data source owner's consent (AI Now Institute 2023). Moreover, antitrust law could allow breaking up large companies (Srnicek 2017; Zuboff 2019). It might,

^{******} This is not a hypothetical worry, see Jakesch et al. 2023.

Former Google CEO Eric Schmidt admitted as much in no uncertain terms in an interview with NBC leading up to the Senate AI hearings (in which OpenAI CEO Sam Altman suggested a pro-business regulatory framework to senators). Schmidt not only advocated what amounts to self-regulation through a business-to-law pipeline – that is, a legal privilege – but also claimed for his industry the sole competence to grasp the intricacies of what is to be regulated – that is, an epistemic privilege. Conveniently, the rule of experts coincides with the rule of Big Tech. "Eric Schmidt: When this technology becomes more broadly available, which it will, very quickly, the problems will get much worse. I would much rather have the companies define reasonable boundaries.

Reporter: It shouldn't be a regulatory framework, it maybe shouldn't even be a sort of a democratic vote, it should be the expertise within the industry to help to sort that out?

Schmidt: The industry will first do that, because there is no way a non-industry person can understand what is possible. It's just too new, too hard, there's not the expertise. There is no one in the government who can get it right, but the industry can roughly get it right, and then the government can put a regulatory structure around it' (NBC News 2023, sc. 9:03).

for instance, be desirable to keep dataset collecting and the training process apart in two separate legal entities. The EU's Digital Markets and Digital Services Acts are better positioned than current US efforts in this regard. The oft-heard argument that any regulation would hamper 'innovation' is misplaced, too: local regulatory efforts within strong economic blocks exert global effects, and the EU's aggressive competition laws have often given it an advantage over US efforts and have led them to be emulated abroad (Bradford et al. 2019). However, as a report in Time magazine has shown, Open-AI lobbyists have already succeeded in influencing the EU AI Act legislation in their favor (Perrigio 2023b).

For this reason, it is necessary to think big here as well. If these more traditional regulatory measures are ineffective and AI systems become the site of articulating social visions, a dominant factor in the make-up of the public sphere, or even a political infrastructure themselves, there is much to be said for actually subjecting LLMs to public control as well. If this is taken to its logical conclusion, the last resort would be communization.

If one accepts the idea that the infrastructures of technologically mediated communication are infrastructures just like any other – water, electricity, roads – and that their construction depends to a considerable extent on direct or indirect public funding (Zuboff 2019) as well as on "data as labor" from unpaid users (Posner and Weyl 2018, 205; Whittaker 2021), then the socialization of certain technologies appears less a shocking overreach of the state than the forceful realization that public goods and services should also be in the hands of a self-governing public (Taylor 2014).

Seen in this way, the open letter appears in a different light: not merely as the technological catastrophism of a group of fearmongering 'longtermists,' but as an attempt to distract from the political consequences of this technology. For those consequences are far more concrete than a robot takeover, even in the stark, heuristically overstated way I have presented them here. Regulating these technologies poses a much more dangerous threat to the companies and individuals profiting most from the hype around AI. But from the standpoint of democratic theory, this is precisely what is needed now.

This essay first appeared as "Das Ende der menschlichen Politik" in *Neue Zürcher Zeitung*, April 25, 2023. The present version has been substantially expanded and updated.

Bibliography

Aggarwal, Alok. 2018. "The Birth of AI and the First AI Hype Cycle." KDnuggets (blog), February 13, 2018. https://www.kdnuggets.com /the-birth-of-ai-and-the-first-ai-hype-cycle.html.

AI Now Institute. 2023. "Antitrust and Competition: It's Time for Structural Reforms to Big Tech." *AI Now Institute* (blog), April 11, 2023. https://ainowinstitute.org/publication/antitrust -and-competition.

Anderson, Elizabeth. 2017. Private Government: How Employers Rule Our Lives (and Why We Don't Talk About It). University Center for Human Values Series. Princeton, Oxford: Princeton University Press.

Apprich, Clemens, forthcoming. "Ja-Sager." In ChatGPT und andere "Quatschmaschinen": Gespräche mit Künstlicher Intelligenz, edited by Anna Tuschling, Andreas Sudmann, and Bernhard J. Dotzler. Bielefeld: Transcript.

Argyle, Lisa P., Ethan Busby, Joshua Gubler, Chris Bail, Thomas Howe, Christopher Rytting, and David Wingate. 2023. "AI Chat Assistants Can Improve Conversations about Divisive Topics." arXiv. https://arxiv.org/abs/2302.07268.

Bagdasaryan, Eugene, and Vitaly Shmatikov. 2022. "Spinning Language Models: Risks of Propaganda-As-A -Service and Countermeasures." In 2022 IEEE Symposium on Security and Privacy (SP), 769-86. https://doi.org/10.1109/SP46214.2022 .9833572.

Bajohr, Hannes. 2020. "On Liberal Disharmony: Judith N. Shklar and the 'Ideology of Agreement."" Journal of the History of Ideas Blog, May 6, 2020. https://jhiblog.org/2020/05/06 /on-liberal-disharmony-judith-n-shklar-and-the -ideology-of-agreement/.

---. 2021. "Wer sind wir? Warum künstliche Intelligenz immer ideologisch ist." *Republik*, April 6, 2021. https://www.republik.ch /2021/04/06/warum-kuenstliche-intelligenz-immer -ideologisch-ist.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–23. Association for Computing Machinery.

Bommasani, Rishi, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, et al. 2021. "On the Opportunities and Risks of Foundation Models." arXiv. http://arxiv.org/abs/2108.07258.

Bostrom, Nick. 2002. "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards." *Journal of Evolution and Technology* 9. http://jetpress.org/volume9/risks.html

Bradford, Anu, Adam Chilton, Katerina Linos, and Alexander Weaver. 2019. "The Global Dominance of European Competition Law Over American Antitrust Law." *Journal of Empirical Legal Studies* 16(4): 731-66. https://doi.org/10.1111/jels.12239.

Bubeck, Sebastien, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, et al. 2023. "Sparks of Artificial General Intelligence: Early Experiments with GPT-4." arXiv. https://doi.org/10.48550 /arXiv.2303.12712.

Eubanks, Virginia. 2017. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. First Edition. New York, NY: St. Martin's Press.

Floridi, Luciano. 2023. "AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models." *Philosophy & Technology* 36 (15). https://doi.org/10.1007/s133347-023-00621-y.

Fricker, Miranda. 2007. Epistemic Injustice: Power and the Ethics of Knowing. Oxford: Oxford University Press. Gitelman, Lisa, ed. 2013. "Raw Data" Is an Oxymoron. Cambridge, MA: MIT Press. https://doi .org/10.1080/1369118x.2014.920042.

Habermas, Jürgen. 2022. "Reflections and Hypotheses on a Further Structural Transformation of the Political Public Sphere." *Theory, Culture & Society* 39 (4): 145–71. https://doi.org/10.1177 /02632764221112341.

Heaven, Will Douglas. 2023. "The Open-Source AI Boom Is Built on Big Tech's Handouts. How Long Will It Last?" *MIT Technology Review*, May 12,2023. https://www.technologyreview.com /2023/05/12/1072950/open-source-ai-google-openai -eleuther-meta/.

Jakesch, Maurice, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. 2023. "Co-Writing with Opinionated Language Models Affects Users' Views." In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 1–15. Hamburg, Germany: ACM. https://doi.org/10.1145/3544548.3581196.

Jin, Berber, and Deepa Seetharaman. 2023. "Elon Musk Creates New Artificial Intelligence Company X.AI." *The Wall Street Journal*, April 14, 2023. https://www.wsj.com/articles/elon-musks-new -artificial-intelligence-business-x-ai-incorporates-in -nevada-962c7c2f.

Klein, Naomi. 2023. "AI Machines Aren't 'Hallucinating'. But Their Makers Are." *The Guardian*, May 8, 2023. https://www.theguardian .com/commentisfree/2023/may/08/ai-machines -hallucinating-naomi-klein.

Luitse, Dieuwertje, and Wiebke Denkena. 2021. "The Great Transformer: Examining the Role of Large Language Models in the Political Economy of AI." *Big Data & Society* 8 (2): 205395172110477. https://doi.org/10.1177/20539517211047734.

Madison, James. 2003. "The Federalist No. 10." In *The Federalist: With Letters of "Brutus,"* edited by Terence Ball. Cambridge: Cambridge University Press. 40–46.

McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. 2006. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *AI Magazine* 27 (4): 12–14.

Miller, Erin, Roudabeh Kishi, Clionadh Raleigh, and Caitriona Dowd. 2022. "An Agenda for Addressing Bias in Conflict Data." *Scientific Data* 9 (1): 593. https://doi.org/10.1038/s41597-022-01705-8.

Mitchell, Melanie, and David C. Krakauer. 2023. "The Debate Over Understanding in Al's Large Language Models." *Proceedings of the National Academy of Sciences* 120 (13). https://doi.org /10.1073/pnas.2215907120.

Mouffe, Chantal. 2005. *On the Political.* London: Routledge.

Noble, Safiya Umoja. 2018. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: New York University.

NBC News, dir. 2023. "Life and Death Decisions Are Being Made' by Artificial Intelligence." YouTube, May 15, 2023. Video, 10:42. https://www.youtube.com/watch?v=VgJnqJ9WxAw. O'Neil, Cathy. 2016. Weapons of Math

Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown.

Pasquinelli, Matteo. 2019. "How a Machine Learns and Fails—A Grammar of Error for Artificial Intelligence." *Spheres*, no. 5. http://spheres-journal .org/how-a-machine-learns-and-fails-a-grammar-of -error-for-artificial-intelligence/.

"Pause Giant AI Experiments: An Open Letter." 2023. Future of Life Institute (blog), March 22, 2023. https://futureoflife.org/open-letter /pause-giant-ai-experiments/.

Perrigio, Billy. 2023a. "Elon Musk Is Bringing the Culture Wars to AI." *Time*, March 3, 2023. https://time.com/6260185/elon-musk-ai -culture-wars/.

- ---. 2023b. "Exclusive: OpenAI Lobbied the E.U. to Water Down AI Regulation." *Time*, June

20, 2023. https://time.com/6288245/openai-eu -lobbying-ai-act/.

Posner, Eric A. und E. Glen Weyl. 2018. Radical Markets: Uprooting Capitalism and Democracy for a Just Society. Princeton: Princeton University Press.

Rozado, David. 2023. "RightWingGPT— An AI Manifesting the Opposite Political Biases of ChatGPT." Substack newsletter. *Rozado's Visual Analytics* (blog), February 16, 2023. https:// davidrozado.substack.com/p/rightwinggpt.

Shklar, Judith N. 1986. Legalism: Law, Morals, and Political Trials. Cambridge: Harvard University Press.

---.1990. *The Faces of Injustice*. New Haven: Yale University Press.

Simonite, Tom. 2021. "What Really Happened When Google Ousted Timnit Gebru." *Wired*, June 8, 2021. https://www.wired.com/story /google-timnit-gebru-ai-what-really-happened/. Srnicek, Nick. 2017. *Platform Capitalism*.

London: Polity.

Stark, Luke. 2023. "Breaking Up (with) AI Ethics." *American Literature* 95, no. 2: 365-79. https://doi.org/10.1215/00029831-10575148.

Taylor, Astra. 2014. The People's Platform: Taking Back Power and Culture in the Digital Age. Toronto: Random House Canada.

Torres, Emile P. 2021. "Against Longtermism." *Aeon*, October 19, 2021. https://aeon .co/essays/why-longtermism-is-the-worlds-most -dangerous-secular-credo.

Vincent, James. 2023a. "AI Is Entering an Era of Corporate Control." *The Verge*, April 3, 2023. https://www.theverge.com/23667752/ai-progress -2023-report-stanford-corporate-control.

. 2023b. "The Senate's Hearing on AI Regulation Was Dangerously Friendly." *The Verge*, May 19,2023. https://www.theverge.com/2023/5 /19/23728174/ai-regulation-senate-hearings-regulatory -capture-laws.

Whittaker, Meredith. 2021. "The Steep Cost of Capture." Interactions 28 (6): 50–55. https://doi .org/10.1145/3488666.

Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus* 109, no. 1: 121-36.

Weatherby, Leif. 2023. "ChatGPT Is an Ideology Machine." *Jacobin*, April 17, 2023. https:// jacobin.com/2023/04/chatgpt-ai-language-models -ideology-media-production.

Yudkowsky, Eliezer. 2023. "The Open Letter on Al Doesn't Go Far Enough: We Need to Shut It All Down." *Time*, March 29, 2023. https://time.com /6266923/ai-eliezer-yudkowsky-open-letter -not-enough/.

Zakrzewski, Cat, Nitasha Tiku, Cristiano Lima, and Will Oremus. 2023. "OpenAI CEO Tells Senate That He Fears AI's Potential to Manipulate Views." *Washington Post*, May 16, 2023. https:// www.washingtonpost.com/technology/2023/05/16 /ai-congressional-hearing-chatgpt-sam-altman/.

Zuboff, Shoshana. 2019. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: Public Affairs.