FORUM

AI and the Human

To the Editor:

Wai Chee Dimock's timely editor's column "AI and the Humanities" expands on a "first of its kind" MLA convention session that included four Microsoft researchers (449; vol. 135, no. 3, May 2020, pp. 449–54). I want to thank Dimock and the MLA for initiating this important conversation. But how might we prepare literature scholars to evaluate specialist presentations on AI like the one Dimock discusses? Widely touted as a fourth industrial revolution, AI is notoriously subject to hype, clickbait, and misinformation. Its loudest proponents include the world's most profitable companies, as well as start-ups keen to attract investors. Widespread confusion stems partly from AI's technical jargon: terms like "deep learning" and "neural networks" suggest that today's technology reproduces the human brain. In fact, the reigning AI software architectures work by mining huge troves of data at unprecedented speed. This approach favors large companies with vast resources at their disposal.

AI research will likely continue to make impressive strides. But the data-centric technologies that dominate the field are fundamentally narrow (excelling at particular tasks), not general (in the manner of human intelligence). Lacking sentience, emotion, common sense, imagination, and a model of the world, these powerful pattern-finders cannot cognize the data points they extrapolate. Translating French into English, they do not "understand" either language; defeating the world's best Go player, they have no sense of "winning" or "play"; identifying some cancer, they do not perceive a disease. As the computer scientist Judea Pearl emphasizes, narrow AI cannot ask "Why?" or "What if?" Still less can it adjudicate the social or moral consequences of its findings or activities.

As AI has become implicated in surveillance, racial bias, political polarization, and environmental harm, discussions over how to govern it have begun to proliferate. Yet, while there is increasing talk of making AI "ethical," "democratic," and "human-centered," scholars in the humanities seldom shape these discussions. Ideally, conversation would focus on democratic decision-making: for instance, Should we pump public money into infrastructures for driverless cars? Should proprietary

317

algorithms be allowed to evaluate "employability" based on pseudoscientific analysis of human facial expressions? But in a media climate that favors sensational stories (which is itself a byproduct of profit-driven algorithms), exaggerations of AI prowess and the potential for *Terminator*-like scenarios take root.

As Dimock reports, Stephen A. Schwarzman, a multibillionaire and ally of Donald Trump, has funded research centers at MIT and Oxford University that purport to ensure that AI will "complement rather than replace human beings." But what powers this vision of data-mining tools as replacements for people? How can a technology that cannot advance a hypothesis, grasp cause and effect, or enjoin moral imagination "render educated human beings superfluous" (450)? The question, of course, is not whether AI has the ability to foment inequality, concentrate power, or harm biological life—which it clearly does. It is about whose interests are served when our conversations engender technological determinism and fear of the future. Whatever the impact of AI, the world today demands an educated citizenry: ready to advance racial justice, develop green technologies, enforce antitrust legislation, enact collective bargaining rights, regulate the gig economy, monitor global supply chains, and enable individuals to control the use of their data.

As Dimock builds on Microsoft's presentation on large language models (LLMs), she cites an op-ed (453n2) in which a start-up promises that, within five years, AIs will write screenplays deemed "better than human writing" (Richard Lea; "If a Novel Was Good, Would You Care If It Was Created by Artificial Intelligence?"; *The Guardian*, 27 Jan. 2020, www.theguardian.com/commentisfree/2020/jan/27/ artificial-intelligence-computer-novels-fiction-write -books). But the text generator in question (OpenAI's GPT-2) is nowhere near able to deliver on that forecast. Rather, despite millions of dollars, "breathtaking amounts of carbon emissions," and 450 gigabytes of data, the latest GPT model, according to the cognitive psychologist Gary Marcus and the computer scientist Ernest Davis, is a "fluent spouter" of statistically probable language, not a "reliable interpreter of the world" ("GPT-3, Bloviator: OpenAI's Language Generator Has No Idea What It's Talking About"; *MIT Technology Review*, 22 Aug. 2020, www.technologyreview.com/ 2020/08/22/1007539/gpt3-openai-language-generator -artificial-intelligence-ai-opinion/)—and still less a font of fabulous screenplays. In the future, such programs

might consistently answer factual questions, or produce something akin to a Wikipedia article mined from existing Wikipedia articles. To the extent that such potential technology helps generate new ideas or spur critical thinking, it will do so by aiding human users.

In November 2020, Timnit Gebru, the influential computer scientist, cofounder of Black in AI, and coleader of Google's ethics team, submitted a research paper titled "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" Gebru and her coauthors used the term "stochastic parrot" to describe LLMs like GPT-3, which parrot language based on stochastic models but do not understand it. LLMs, they argued, are unreliable, environmentally irresponsible, and subject to the biases of huge and undocumented training sets. What is more, these resource-intensive programs favor tech behemoths like Google while diverting research from creative approaches that might depend less on data (and could certainly screen out biased language [*University of Washington*, faculty .washington.edu/ebender/papers/Stochastic_Parrots.pdf]). Google's response was to order Gebru to retract the paper: when she demanded to know why, the company summarily fired the woman of color it had hired to showcase its commitment to ethics (Cade Metz and Daisuke Wakabayashi; "Google Researcher Says She Was Fired over Paper Highlighting Bias in A.I."; *The New York Times*, 3 Dec. 2020, www.nytimes.com/ 2020/12/03/technology/google-researcher-timnit-gebru .html?searchResultPosition=2). Yes, one sometimes imagines that robots would make better decisions.

Like most technologies, AI has the potential to serve democracy, inclusion, and environmental sustainability. But for Silicon Valley, "democratizing" AI means encouraging businesses and consumers to adopt new products in exchange for data and fees. Indeed, OpenAI—no longer very open—has now commercialized GPT-3 through an exclusive license with Microsoft.

In our time of crisis, tech companies and investors may dream of replacing schools with chatbots, just as their precursors envisioned swapping out the brick-and-mortar university for CD-ROMS, MOOCs, and (most recently) Zoom. If it happens, the reason will not be students' desire to learn from stochastic parrots. Nor will it be because technology made educated human beings superfluous. It will be because democracy failed.

I hope the MLA continues to foster conversations on AI and spur humanist understanding and advocacy. Perhaps a next invitation might go to Gebru—or,

maybe, Rediet Abebe, Joy Buolamwini, Cathy O'Neil, or Meredith Whittaker.

Lauren M. E. Goodlad
Rutgers University, New Brunswick

Reply:

I thank Lauren Goodlad for opening up a conversation about AI and the humanities, and can't agree more that algorithms have become deeply "implicated in surveillance, racial bias, [and] political polarization," developments too worrisome to be ignored. To her list of proposed speakers at future MLA sessions, I would like to add a few more. In *Race after Technology: Abolitionist Tools for the New Jim Code* (Polity Press, 2019), Ruha Benjamin discusses the knee-jerk racial profiling enforced by predictive analytics, algorithms that prejudge whole segments of the population based on facile stereotypes. Shoshana Zuboff shows that these predictive algorithms are the linchpin of a data economy still more insidious and encompassing, what she calls "surveillance capitalism." Predicting our consumer choices from the information collected online, this data economy thrives on "behavioral futures," so lucrative that nearly all the market capitalization of Google and Facebook (roughly eighty-eight percent and ninety-eight percent, respectively) comes from this revenue source (*The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* [PublicAffairs, 2019] pp. 328–50).

The effect on democratic institutions could not be more destructive. In "The Coup We Are Not Talking About," an opinion piece in *The New York Times* on 29 January 2021, Zuboff specifically links the insurrection at the Capitol to an "epistemic coup" twenty years in the making, based on the "profit-driven algorithmic amplification, dissemination and microtargeting of corrupt information." The truth or falsehood of online content makes no difference to this computational regime. Optimized only for user engagement, algorithms will promote anything that serves that end: they "splinter shared reality, poison social discourse, paralyze democratic politics and sometimes instigate violence and death" (www.nytimes.com/2021/01/29/opinion/sunday/facebook-surveillance-society-technology.html?searchResultPosition=1).

Field reports by Eli Parisier, Sheera Frenkel, Kevin Roose, and many others support this claim. Search engines and social media are extremist incubators, designed to prioritize self-replicating, self-reinforcing content. Creating filter bubbles and echo chambers for the like-minded, they fuel group polarization and advertising revenue in the same click. There is a reason that misinformation and monetization are so optimally aligned. The belated gestures of Amazon, Apple, and Google, as well as Twitter and Facebook, to deny service to a few obvious targets have not stopped extremists from regrouping across platforms and networking globally, a viral divisiveness likely to be with us for a long time. In the light of this, humanists can no longer afford to look away. We need to be at the table, learning as much as we can about artificial intelligence, so that democratic decision-making will reflect our input, based on hard-won knowledge rather than rampant misinformation.

Wai Chee Dimock
Yale University