AI Policy Matters

Larry Medsker (The George Washington University; lrm@gwu.edu)

DOI: 10.1145/3465074.3465079

Abstract

AI Policy Matters is a regular column in *AI Matters* featuring summaries and commentary based on postings that appear twice a month in the *AI Matters* blog (https://sigai.acm.org/aimatters/blog/). We welcome everyone to make blog comments so we can develop a rich knowledge base of information and ideas representing the SIGAI members.

FR and Bad Science: Should some research not be done?

Facial recognition issues continue to appear in the news, as well as in scholarly journal articles, while FR systems are being banned and some research is shown to be bad science. AI system researchers who try to associate facial technology output with human characteristics are sometimes referred to as machine-assisted phrenologists. Problems with FR research have been demonstrated in machine learning research such as work by Steed and Caliskan in "A set of distinct facial traits learned by machines is not predictive of appearance bias in the wild." Meanwhile many examples of harmful products and misuses have been identified in areas such as criminality, video interviewing, and many others. Some communities have considered bans.

Yet, journals and conferences continue to publish bad science in facial recognition.

Some people say the choice of research topics is up to the researchers – the public can choose not to use the products of their research. However, areas such as genetic, biomedical, and cybersecurity R&D do have limits. Our professional computing societies can choose to disapprove research areas that cause harm. Sources of mitigating and preventing irresponsible research being introduced into the public space include:

Peer pressure on academic and corporate research and development

Public policy through laws and regulations

Corporate and academic self-interest – organizations' bottom lines can suffer from bad behavior and publicity

Vigilance by journals about publishing papers that promulgate the misuse of FR.

A recent article by Matthew Hutson in The New Yorker discusses "Who should stop unethical AI." He remarks that "Many kinds of researchers—biologists, psychologists, anthropologists, and so on—encounter checkpoints at which they are asked about the ethics of their research. This doesn't happen as much in computer science. Funding agencies might inquire about a project's potential applications, but not its risks. University research that involves human subjects is typically scrutinized by an I.R.B., but most computer science doesn't rely on people in the same way. In any case, the Department of Health and Human Services explicitly asks I.R.B.s not to evaluate the possible long-range effects of applying knowledge gained in the research, lest approval processes get bogged down in political debate. At journals, peer reviewers are expected to look out for methodological issues, such as plagiarism and conflicts of interest; they haven't traditionally been called upon to consider how a new invention might rend the social fabric."

Big Issues

Big Tobacco, Big Oil . . . and Big Tech

A larger discussion is growing out of the recent news about Dr. Timnit Gebru and Google. Big Tech is having a huge impact on individuals and society both for the many products and services we enjoy and for the current and potential cases of detrimental effects of unethical behavior or naiveté regarding AI ethics issues. How do we achieve AI ethics responsibility in all organizations, big and small? And,

not just in corporations, but governmental and academic research organizations?

Some concerned people focus on regulation, but for a variety of reasons public and community pressure may be quicker and more acceptable. This includes corporations earning reputations for ethical actions in the design and development of AI products and systems. An article in MIT Technology Review by Karen Hao discusses a letter signed by nine members of Congress that "sends an important signal about how regulators will scrutinize tech giants." Ideally our Public Policy goal is strong AI Ethics in national and global communities that self-regulate on AI ethical issues, comparable to other professional disciplines in medical science and cybersecurity. Our AI Ethics community, as guidelines evolve, could provide a supportive and guiding presence in the implementation of ethical norms in the research and development in AI. The idea of a global AI Ethics community is reflected also in a recent speech by European Union President Ursula von der Leyen at the World Leader for Peace and Security Award ceremony. She advocates for transatlantic agreements on AI.

What Can Biden Do for Science?

A *Science—Business* Webcast presented a forum of public and private sector leaders discussing ideas about the need for the president-elect to convene world leaders to re-establish "rules of engagement" on science. Participants in the Webcast urged that a global assembly "should press leaders of the big industrial nations to open – or reopen – their research systems, while also ensuring that COVID-19 vaccines are freely available to everyone in the world." About an international summit, Robert-Jan Smits, former director-general of the European Commission's research and innovation directorate said it "would really show that senior leaders are turning the page."

HCAI for Policymakers

"Human-Centered AI" by Ben Shneiderman was recently published in *Issues in Science and Technology* 37, no. 2 (Winter 2021): 56–61. A timely observation is that Artificial Intelligence is clearly expanding to include human-centered issues from ethics, explainability, and trust to applications such as user interfaces for self-driving cars. The importance of the HCAI fresh approach, which can enable more widespread use of AI in safe ways that promote human control, is acknowledged by the appearance in NAS Issues in Science and Technology. An implication of the article is that computer scientists should build devices to enhance and empower—not replace—humans.

HCAI as described by Prof. Shneiderman represents a radically different approach to systems design by imagining a different role for machines. Envisioning AI systems as comprising machines and people working together is a much different starting point than the assumption and goal of autonomous AI. In fact, a design process with this kind of forethought might even lead to a product not being developed, thus preventing future harm. One of the many interesting points in the NAS Issues article is the observation about the philosophical clash between two approaches to gaining knowledge about the world—Aristotle's rationalism and Leonardo da Vinci's empiricism—and the connection with the current perspective of AI developers: "The rationalist viewpoint, however, is dominant in the AI community. It leads researchers and developers to emphasize data-driven solutions based on algorithms." Data science in particular unfortunately often focuses on the rationalist approach without including the contributions from and protection of the human experience.

From the NAS article, HCAI is aligned with "the rise of the concept of design thinking, an approach to innovation that begins with empathy for users and pushes forward with humility about the limits of machines and people. Empathy enables designers to be sensitive to the confusion and frustration that users might have and the dangers to people when AI systems fail. Humility leads designers to recognize the inevitability of failure and inspires them to be always on the lookout for what wrongs are preventable." Policymakers need to "understand HCAI's promise not only for our machines but for our lives. A good starting place is an appreciation of the two competing philosophies that have shaped the development of AI, and what those imply for the

design of new technologies ... comprehending these competing imperatives can provide a foundation for navigating the vast thicket of ethical dilemmas now arising in the machine-learning space." An HCAI approach can incorporate creativity and innovation into AI systems by understanding and incorporating human insights about complexity into the design of AI systems and using machines to prepare data for taking advantage of human insight and experience. For many more details and enjoyable reading, see the NAS Issues article.

AI Centre of Excellence (AICE)

AICE conducted their inaugural celebration in December, 2020. Director John Kamara founded the AI Centre of Excellence in Kenya and is passionate about creating value and long term impact of AI and ML in Africa. The Centre aims to accomplish this by providing expert training to create skilled and employable AI and ML engineers. The Centre dives into creating sustainable impact through Research and Development. AI research and products are estimated to contribute over 13 trillion dollars to the global economy by 2030. This offers the Centre an opportunity to carry out research in selected sectors and build products based on the research. The world has around 40K AI experts in the world, with nearly half in the US and less than 5 percent in Africa. Oxford Insights estimates that Kenya ranks first in Africa, and AICE aims to leverage this potential and transform AICE into a full blown Artificial Intelligence Centre of Excellence. Please keep your eyes on Africa and ways our public policy can assist efforts there to grow AI in emerging education and research.

Data for AI: Interview with Dr. Eric Daimler

I recently spoke with Dr. Eric Daimler about how we can build on the framework he and his colleagues established during his tenure as a contributor to issues of AI policy in the White House during the Obama administration. Eric is the CEO of the MIT-spinout Conexus.com and holds a PhD in Computer Science from Carnegie Mellon University. Here are the interesting results of my interview with him. His ideas are important as part of the basis for ACM SIGAI Public Policy recommendations.

LRM: What are the main ways we should be addressing this issue of data for AI?

EAD: To me there is one big re-framing from which we can approach this collection of issues, prioritizing data interoperability within a larger frame of AI as a total system. In the strict definition of AI, it is a learning algorithm. Most people know of subsets such as Machine Learning and subsets of that called Deep Learning. That doesn't help the 99 percent who are not AI researchers. When I have spoken to non-researchers or even researchers who want to better appreciate the sensibilities of those needing to adopt their technology, I think of AI as the interactions that it has. There is the collection of the data, the transportation of the data, the analysis or planning (the traditional domain in which the definition most strictly fits), and the acting on the conclusions. That sense, plan, act framework works pretty well for most people.

LRM: Before you explain just how we can do that, can you go ahead and define some of your important terms for our readers?

EAD: AI is often described as the economic engine of the future. But to realize that growth, we must think beyond AI to the whole system of data, and the rules and context that surround it: our data infrastructure (DI). Our DI supports not only our AI technology, but also our technical leadership more generally; it underpins COVID reporting, airline ticket bookings, social networking, and most if not all activity on the internet. From the unsuccessful launch of healthcare.gov, to the recent failure of Haven, to the months-long hack into hundreds of government databases, we have seen the consequences faulty DI can have. More data does not lead to better outcomes; improved DI does.

Fortunately, we have the technology and foresight to prevent future disasters, if we act now. Because AI is fundamentally limited by the data that feeds it, to win the AI race, we must build the best DI. The new presidential administration can play a helpful role here, by defining standards and funding research into data technologies. Attention to the need for better DI will speed responsiveness to future crises (consider COVID data delays) and es-

tablish global technology leadership via standards and commerce. Investing in more robust DI will ensure that anomalies, like ones that would have helped us identify the Russia hack much sooner, will be evident, so we can prevent future malfeasance by foreign actors. The US needs to build better data infrastructure to remain competitive in AI.

LRM: So how might we go about prioritizing data interoperability?

EAD: In 2016, the Department of Commerce (DOC) discovered that on average, it took six months to onboard new suppliers to a midsize trucking company—because of issues with data interoperability. The entire American economy would benefit from encouraging more companies to establish semantic standards, internally and between companies, so that data can speak to other data. According to a DOC report in early 2020, the technology now exists for mismatched data to communicate more easily and data integrity to be guaranteed, thanks to a new area of math called Applied Category Theory (ACT). This should be made widely available.

LRM: And what about enforcing data provenance?

EAD: As data is transformed across platforms—including trendy cloud migrations—its lineage often gets lost. A decision denying your small business loan can and should be traceable back to the precise data the loan officer had at that time. There are traceability laws on the books, but they have been rarely enforced because up until now, the technology hasn't been available to comply. That's no longer an excuse. The fidelity of data and the models on top of them should be proven—down to the level of math—to have maintained integrity.

LRM: Speaking more generally, how can we start to lay the groundwork to reap the benefits of these advancements in data infrastructure?

EAD: We need to formalize. When we built 20th century assembly lines, we established in advance where and how screws would be made; we did not ask the village blacksmith to fashion custom screws for every home repair. With AI, once we know what we want to have automated (and there are good reasons to not to automate everything!), we should then define in advance how we want it to behave. As you read this, 18 million programmers are already formalizing rules across every aspect of technology. As an automated car approaches a crosswalk, should it slow down every time, or only if it senses a pedestrian? Questions like this one—across the whole economy—are best answered in a uniform way across manufacturers, based on standardized, formal, and socially accepted definitions of risk.

LRM: In previous posts, I have discussed roles and responsibilities for change in the use of AI. Government regulation is of course important, but what roles do you see for AI tech companies, professional societies, and other entities in making the changes you recommend for DI and other aspects of data for AI?

EAD: What is different this time is the abruptness of change. When automation technologies work, they can be wildly disruptive. Sometimes this is very healthy (see: Schumpeter). I find that the "go fast and. . . " framework has its place, but in AI it can be destructive and invite resistance. That is what we have to watch out for. Only with responsible coordinated action do we encourage adoption of these fantastic and magical technologies. Automation in software can be powerful. These processes need not be linked into sequences just because they can. That is, just because some system can be automated does not mean that it should. – Too often there is absolutism in AI deployments when what is called for in these discussions is nuance and context. For example, in digital advertising my concerns are around privacy, not physical safety. When I am subject to a plane's autopilot, my priorities are reversed.

With my work in the US Federal Government, my bias remains against regulation as a first-step. Shortly after my time with the Obama Whitehouse, I am grateful to have participated with a diverse group for a couple of days at the Halcyon House in Washington D.C. We created some principles for deploying AI to maximize adoption. We can build on these and rally around a sort of LEED-like standard for AI deployment.

–

Dr. Eric Daimler is CEO and Founder of Conexus and Board Member of Petuum and WelWaze. He was a Presidential Innova-

tion Fellow, Artificial Intelligence and Robotics. Eric is a leading authority in robotics and artificial intelligence with over 20 years of experience as an entrepreneur, investor, technologist, and policymaker. Eric served under the Obama Administration as a Presidential Innovation Fellow for AI and Robotics in the Executive Office of President, driving the agenda for U.S. leadership in research, commercialization, and public adoption of AI and Robotics. His newest venture, Conexus, is a groundbreaking solution for what is perhaps today's biggest information technology problem — data deluge. Eric works to empower communities and citizens to leverage robotics and AI to build a more sustainable, secure, and prosperous future. His academic research has been at the intersection of AI, Computational Linguistics, and Network Science (Graph Theory). He has studied at the University of Washington-Seattle, Stanford University, and Carnegie Mellon University, where he earned his Ph.D. in Computer Science.

Please join our discussions at the

SIGAI Policy Blog

Larry Medsker is a Research Professor at The George Washington University, where he was founding director of the Data Science graduate program. He is currently a faculty member in the GW Human-Technology Collaboration Lab and Ph.D. program. His research in AI includes work on artificial neural networks, hybrid intelligent systems, and the impacts of AI on society and policy. He is Co-Editor-in-Chief for the journal *AI and Ethics* and the Public Policy Officer for the ACM SIGAI.