AI Policy Matters

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

DOI: 10.1145/3430581.3430587

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.

AI Data

Confusion in the popular media about terms such as algorithm and what constitutes AI technology cause critical misunderstandings among the public and policymakers. More importantly, the role of data is often ignored in ethical and operational considerations. Even if AI systems are perfectly built, low quality and biased data cause unintentional and even intentional hazards.

Language Models and Data

A generative pre-trained transformer GPT-3 is currently in the news. For example, James Vincent in the July 30, 2020, article in The Verge writes about GPT-3, which was created by OpenAI. Language models, GPT-3 the current ultimate product, have ethics issues on steroids for products being made. Inputs to the system have all the liabilities discussed about Machine Learning and Artificial Neural Network products. The dangers of bias and mistakes are raised in some writings but are likely not a focus among the wide range of enthusiastic product developers using the open-source GPT-3. Language models suggest output sequences of words given an input sequence. Thus, samples of text from social media can be used to produce new text in the same style as the author and potentially can be used to influence public opinion. Cases have been found of promulgating incorrect grammar and misuse of terms based on poor quality inputs to language models. An article by David Pereira includes examples and comments on the use of GPT-3. The article "GPT-3: an AI Game-Changer or an Environmental Disaster?" by John Naughton gives examples of and commentary on results from GPT-3.

Data Governance

A possible meta solution for policymakers to keep up with technological advances and AI data issues is discussed by Alex Woodie in "AI Ethics and Data Governance: A Virtuous Cycle." He quotes James Cotton, who is the international director of the Data Management Centre of Excellence at Information Builders' Amsterdam office: "as powerful as the AI technology is, it can't be implemented in an ethical manner if the underlying data is poorly managed and badly governed. It's critical to understand the relationship between data governance and AI ethics. One is foundational for the other. You can't preach being ethical or using data in an ethical way if you don't know what you have, where it came from, how it's being used, or what it's being used for."

USTPC in the News

The ACM's US Technology Policy Committee (USTPC) was very active in July, 2020! The contributions and visibility of USTPC as a group and as individual members are very welcome and impressive. The following list has links to highly-recommended reading.

Amicus Brief: USTPC Urges Narrower Definition of Computer Fraud and Abuse Act

ACM's USTPC filed an amicus curiae ("friend of the court") brief with the United States Supreme Court in the landmark case of Van Buren v. United States. "Van Buren marks the first time that the US Supreme Court has reviewed the Computer Fraud and Abuse Act (CFAA), a 1986 law that was originally intended to punish hacking. In recent years, however, the CFAA has been used to criminally prosecute both those who access a computer system without permission, as well as those who have permission but exceed their authority to use a database once logged in."

USTPC Statement On Face Recognition

(USTPC) has assessed the present state of facial recognition (FR) technology as applied by the government and private sector. The Committee concludes that, "when rigorously evaluated, the technology too often produces results demonstrating clear bias based on ethnic, racial, gender, and other human characteristics recognizable by computer systems. The consequences of such bias, USTPC notes, frequently can and do extend well beyond inconvenience to include profound injury, particularly to the lives, livelihoods and fundamental rights of individuals in specific demographic groups, including some of the most vulnerable populations in our society." See the NBC news article.

Barbara Simons Recipient of the 2019 ACM Policy Award

USTPC's Barbara Simons, founder of USTPC predecessor USACM, is the recipient of the 2019 ACM Policy Award for "long-standing, high-impact leadership as ACM President and founding Chair of ACM's US Public Policy Committee (USACM), while making influential contributions to improve the reliability of and public confidence in election technology. Over several decades, Simons has advanced technology policy by founding and leading organizations, authoring influential publications, and effecting change through lobbying and public education." Congratulations, Barbara!

Potential New Issues

ACM urged Preservation of Temporary Visa Exemptions for Nonimmigrant Students, and Harvard filed a complaint for declaratory and injunctive relief. This issue may have dramatic impacts on university research and teaching.

Thank you, USTPC, for your hard work and representation of ACM to policymakers!

AI in Congress

Politico reports on two separate bills introduced on June 2. (See the section entitled "Artificial Intelligence: Let's Do the Thing".)

The National AI Research Resource Task Force Act. "The bipartisan, bicameral bill introduced by Reps. Anna Eshoo, (D-Calif.), Anthony Gonzalez (R-Ohio), and Mikie Sherrill (D-N.J.), along with companion legislation by Sens. Rob Portman (R-Ohio) and Martin Heinrich (D-N.M.), would form a committee to figure out how to launch and best use a national AI research cloud. Public and private researchers and developers from across the country would share this cloud to combine their data, computing power and other resources on AI. The panel would include experts from government, academia and the private sector."

The Advancing Artificial Intelligence Research Act. "The bipartisan bill introduced by Senate Commerce Chairman Roger Wicker (R-Miss.), Sen. Cory Gardner (R-Colo.) and Gary Peters (D-Mich.), a founding member of the Senate AI Caucus, would create a program to accelerate research and development of guidance around AI at the National Institute of Standards and Technology. It would also create at least a half-dozen AI research institutes to examine the benefits and challenges of the emerging technology and how it can be deployed; provide funding to universities and nonprofits researching AI; and launch a pilot at the National Science Foundation for AI research grants."

AI and Facial Recognition

Concerns About Facial Recognition: Discrimination, Privacy, and Democratic Freedom

While including ethical and moral issues, a broader list of issues is concerning to citizens and policymakers about face recognition technology and AI. Areas of concerns include accuracy; surveillance; data storage, permissions, and access; discrimination, fairness, and bias; privacy and video recording without consent; democratic freedoms, including right to choose, gather, and speak; and abuse of technology such as non-intended uses, hacking, and deep fakes. Used responsibly and ethically, face recognition can be valuable for finding missing people, responsible policing and law enforcement, medical uses, healthcare, virus tracking, legal system and court uses, and advertising. Various guidelines by organizations such as the AMA and legislation like S.3284 – Ethical Use of Facial Recognition Act are being developed to encourage the proper use of AI and face recognition. Some of the above issues do specifically require ethical analysis as in the following by Yaroslav Kuflinski:

1. Accuracy — FR systems naturally discriminate against non-whites, women, and children, presenting errors of up to 35% for non-white women.
2. Surveillance issues — concerns about “big brother” watching society.
3. Data storage — use of images for future purposes stored alongside genuine criminals.
4. Finding missing people — breaches of the right to a private life.
5. Advertising — invasion of privacy by displaying information and preferences that a buyer would prefer to keep secret.
6. Studies of commercial systems are increasingly available, for example an analysis of Amazon Rekognition.
7. Biases deriving from sources of unfairness and discrimination in machine learning have been identified in two areas: the data and the algorithms. Biases in data skew what is learned in machine learning methods, and flaws in algorithms can lead to unfair decisions even when the data is unbiased. Intentional or unintentional biases can exist in the data used to train FR systems.
8. New human-centered design approaches seek to provide intentional system development steps and processes in collecting data and creating high quality databases, including the elimination of naturally occurring bias reflected in data about real people.

Bias That Pertains Especially to Facial Recognition (Mehrabi, et al. and Barocas et. al.)

1. Direct Discrimination: “Direct discrimination happens when protected attributes of individuals explicitly result in non-favorable outcomes toward them”. Some traits like race, color, national origin, religion, sex, family status, disability, exercised rights under CCPA, marital status, receipt of public assistance, and age are identified as sensitive attributes or protected attributes in the machine learning world.
2. Indirect Discrimination: Even if sensitive or protected attributes are not used against an individual, indirect discrimination can still happen. For example, residential zip code is not categorized as a protected attribute, but from the zip code one might infer race, which is a protected attribute. So, “protected groups or individuals still can get treated unjustly as a result of implicit effects from their protected attributes”.
3. Systemic Discrimination: “policies, customs, or behaviors that are a part of the culture or structure of an organization that may perpetuate discrimination against certain subgroups of the population”.
4. Statistical Discrimination: In law enforcement, racial profiling is an example of statistical discrimination. In this case, minority drivers are pulled over more than compared to white drivers — “statistical discrimination is a phenomenon where decision-makers use average group statistics to judge an individual belonging to that group.”
5. Explainable Discrimination: In some cases, discrimination can be explained using attributes like working hours and education, which is legal and acceptable. In “the UCI Adult dataset, a widely-used dataset in the fairness domain, males on average have a higher annual income than females; however, this is because, on average, females work fewer hours than males per week. Work hours per week is an attribute that can be used to explain low income. If we make decisions without considering working hours such that males and females end up averaging the same income, we could lead to reverse discrimination since we would cause male employees to get lower salary than females.
6. Unexplainable Discrimination: This type of discrimination is not legal as explainable discrimination because “the discrimination toward a group is unjustified”.

How to Discuss Facial Recognition

Recent controversies about FR mix technology issues with ethical imperatives and ignore that people can disagree on which are the “correct” ethical principles. A recent ACM tweet on FR and face masks was interpreted in different ways and ACM issued an official clarification. A question that emerges is if AI and other technologies should be, and can be, banned rather than controlled and regulated. In early June, 2020, IBM CEO Arvind Krishna said in a letter to Congress that IBM is exiting the facial recognition business and asking for reforms to combat racism: "IBM no longer offers general purpose IBM facial recognition or analysis software. IBM firmly opposes and will not condone uses of any technology, including facial recognition technology offered by other vendors, for mass surveillance, racial profiling, violations of basic human rights and freedoms, or any purpose which is not consistent with our values and Principles of Trust and Transparency," Krishna said in his letter to members of congress, "We believe now is the time to begin a national dialogue on whether and how facial recognition technology should be employed by domestic law enforcement agencies."

Policy and AI Ethics

The Alan Turing Institute Public Policy Programme

Among the complexities of public policy making, the new world of AI and data science requires careful consideration of ethics and safety in addressing complex and far-reaching challenges in the public domain. Data and AI systems lead to opportunities that can produce both good and bad outcomes. Ethical and safe systems require intentional processes and designs for organizations responsible for providing public services and creating public policies. An increasing amount of research focuses on developing comprehensive guidelines and techniques for industry and government groups to make sure they consider the range of issues in AI ethics and safety in their work.

An excellent example is the Public Policy Programme at The Alan Turing Institute under the direction of Dr. David Leslie. Their work complements and supplements the Data Ethics Framework, which is a practical tool for use in any project initiation phase. Data Ethics and AI Ethics regularly overlap. The Public Policy Programme describes AI Ethics as "a set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies. These values, principles, and techniques are intended both to motivate morally acceptable practices and to prescribe the basic duties and obligations necessary to produce ethical, fair, and safe AI applications. The field of AI ethics has largely emerged as a response to the range of individual and societal harms that the misuse, abuse, poor design, or negative unintended consequences of AI systems may cause." They cite the following as some of the most consequential potential harms:

- Bias and Discrimination
- Denial of Individual Autonomy, Recourse, and Rights
- Non-transparent, Unexplainable, or Unjustifiable Outcomes
- Invasions of Privacy
- Isolation and Disintegration of Social Connection
- Unreliable, Unsafe, or Poor-Quality Outcomes

The Ethical Platform for the Responsible Delivery of an AI Project, strives to enable the "ethical design and deployment of AI systems using a multidisciplinary team effort. It demands the active cooperation of all team members both in maintaining a deeply ingrained culture of responsibility and in executing a governance architecture that adopts ethically sound practices at every point in the innovation and implementation lifecycle." The goal is to "unite an in-built culture of responsible innovation with a governance architecture that brings the values and principles of ethical, fair, and safe AI to life." Useful references:

1 Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute.

2 Data Ethics Framework (2018).

Principled Artificial Intelligence

In January, 2020, the Berkman Klein Center released a report by Jessica Fjeld and Adam Nagy "Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI", which summarizes contents of 36 documents on AI principles. This work acknowledges the surge in frameworks based on ethical and human rights to guide the development and use of AI technologies. The authors focus on understanding ethics efforts in terms of eight key thematic trends:

- Privacy
- Accountability
- Safety and security
- Transparency and explainability
- Fairness and non-discrimination
- Human control of technology
- Professional responsibility
- Promotion of human values

They report "our analysis examined the forty-seven individual principles that make up the themes, detailing notable similarities and differences in interpretation found across the documents. In sharing these observations, it is our hope that policymakers, advocates, scholars, and others working to maximize the benefits and minimize the harms of AI will be better positioned to build on existing efforts and to push the fractured, global conversation on the future of AI toward consensus."

Human-Centered AI

Prof. Ben Shneiderman's research emphasizes human autonomy as opposed to the popular notion of autonomous machines. The ideas are now available in the International Journal of Human–Computer Interaction. The abstract is as follows: "Well-designed technologies that offer high levels of human control and high levels of computer automation can increase human performance, leading to wider adoption. The Human-Centered Artificial Intelligence (HCAI) framework clarifies how to (1) design for high levels of human control and high levels of computer automation so as to increase human performance, (2) understand the situations in which full human control or full computer control are necessary, and (3) avoid the dangers of excessive human control or excessive computer control. The methods of HCAI are more likely to produce designs that are Reliable, Safe and Trustworthy (RST). Achieving these goals will dramatically increase human performance, while supporting human self-efficacy, mastery, creativity, and responsibility."

COVID AI

AI is in the news and in policy discussions regarding COVID-19, both about ways to help fight the pandemic and in terms of ethical issues that policymakers should address. Michael Corkery and David Gelles in the NY Times article "Robots Welcome to Take Over, as Pandemic Accelerates Automation", suggest that "social-distancing directives, which are likely to continue in some form after the crisis subsides, could prompt more industries to accelerate their use of automation." An MIT Technology Review article by Genevieve Bell, "We need mass surveillance to fight COVID-19—but it doesn't have to be creepy" looks at the pros and cons of AI technology and if we now have the chance to "reinvent the way we collect and share personal data while protecting individual privacy."

Public Health and Privacy Issues

Liza Lin and Timothy W. Martin in "How Coronavirus Is Eroding Privacy" write about how technology is being developed to track and monitor individuals for slowing the pandemic, but that this "raises concerns about government overreach."

Here is an excerpt from that WSJ article: "Governments worldwide are using digital surveillance technologies to track the spread of the coronavirus pandemic, raising concerns about the erosion of privacy. "

"Many Asian governments are tracking people through their cellphones to identify those suspected of being infected with COVID-19 without prior consent. European countries are tracking citizens' movements via telecommunications data that they claim conceals individuals' identities; American officials are drawing cellphone location data from mobile advertising firms to monitor crowds, but not individuals. The biggest privacy debate concerns involuntary use of smartphones and other digital data to identify everyone with whom the infected had recent contact, then testing and quarantining at-risk individuals to halt the further spread of the disease. Public health officials say surveillance will be necessary in the months ahead, as quarantines are relaxed and the virus remains a threat while a vaccine is developed."

"In South Korea, investigators scan smartphone data to find within 10 minutes people who might have caught the coronavirus from someone they met. Israel has tapped its Shin Bet intelligence unit, usually focused on terrorism, to track down potential Coronavirus patients through telecom data. One U.K. police force uses drones to monitor public areas, shaming residents who go out for a stroll."

"The COVID-19 pandemic is ushering in a new era of digital surveillance and rewiring the world's sensibilities about data privacy. Governments are imposing new digital surveillance tools to track and monitor individuals. Many citizens have welcomed tracking technology intended to bolster defenses against the novel coronavirus. Yet some privacy advocates are wary, concerned that governments might not be inclined to unwind such practices after the health emergency has passed."

"Authorities in Asia, where the virus first emerged, have led the way. Many governments didn't seek permission from individuals before tracking their cellphones to identify suspected coronavirus patients. South Korea, China and Taiwan, after initial outbreaks, chalked up early successes in flattening infection curves to their use of tracking programs."

"In Europe and the U.S., where privacy laws and expectations are more stringent, governments and companies are taking different approaches. European nations monitor citizen movement by tapping telecommunications data that they say conceals individuals' identities."

"American officials are drawing cellphone location data from mobile advertising firms to track the presence of crowds—but not individuals. Apple and Google recently announced plans to launch a voluntary app that health officials can use to reverse-engineer sickened patients' recent whereabouts—provided they agree to provide such information."

NSF Program on Fairness in Artificial Intelligence (FAI) in Collaboration with Amazon

A new National Science Foundation solicitation NSF 20-566 has been announced by the Directorate for Computer and Information Science and Engineering, Division of Information and Intelligent Systems, Directorate for Social, Behavioral and Economic Sciences, and Division of Behavioral and Cognitive Sciences.

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.