A Policy Memorandum to the Partnership on AI: Accreditation and Educational Programs to Ensure Fairness in AI

Andy Sanchez, Kevin McDermott, Alicia Cintora
Cornell University, Ithaca, New York
Corresponding author: andysanchez2013@gmail.com
Keywords: artificial intelligence; AI; fairness; bias mitigation; regulation; accreditation; education; partnership on AI; technology

Executive Summary: Artificial intelligence (AI) technologies are set to revolutionize society and in many instances has already seen partial- or full-scale deployment. These AI models learn from large data sets, which are unfortunately susceptible to biases. Many examples—from gender-dependent error in facial recognition software to race-dependent error in bail assessment—demonstrate the urgent need for ensuring fairness as a central component in AI development. To tackle the problem on a systemic scale that can address current and emerging AI technologies, we propose for the Partnership on AI—a nonprofit consortium of leading tech companies, civil rights groups, and universities—to establish an independent nonprofit accreditation board, the Forum for Artificial Intelligence Regularization (FAIR). We further propose that their academic partners adopt major changes to the computer science education system, focusing on interdisciplinary discussion, ethics, and skill building to help minimize bias in AI. Our proposals are tailored to promote strong industry buy-in, boost innovation, and invite a new generation of conscientious engineers to help safely design our future. Ultimately, this work would optimize the intersection of product development and bias mitigation, benefiting companies and consumers alike.

I. Introduction

Artificial intelligence (AI) technologies, or systems that perceive their environment and "learn" to achieve specified goals, are set to revolutionize society (Russell and Norvig 2003). Such models are already active in everyday technologies, e.g. Siri, and ongoing developments can be expected to have far-reaching impacts. These models will change our technology use, our social systems, and even our basic understanding of what costs—be they social, political, or financial—are acceptable for different technological advancements. Look no further than the incredible investments into autonomous vehicles, not only by automotive giants like Ford and Tesla, but also transportation experts Uber and Lyft (Mercer and Macaulay 2019). These self-driving cars will substantially reduce travel times and improve fuel efficiency (Chen et al. 2019), and they also will mitigate one of the most significant modern public health concerns: vehicular deaths. Given that multiple studies cite human error as the leading cause of car accidents (National Highway Traffic Safety Administration 2008; National Highway Traffic Safety Administration 2015), the management consulting firm McKinsey & Company estimates that AI could reduce the number of global vehicular deaths by as much as 93% (Bertoncello and Wee 2015).

With such transformative potential benefits on the horizon, AI technologies are already being incorporated into our infrastructures in increasingly pervasive ways. International Data Corporation, a leader in global market analysis, estimates that by 2022 $158 billion will have been spent on Smart Cities (Shirer and Da Rold 2018). The Smart Cities Council defines these ventures as bridging AI and urban planning to promote "livability, workability, and sustainability" (Enbysk, Bane, and Cowan 2016). A leading example is Sidewalk Toronto, a $1 billion development organized by Alphabet, publicized as a neighborhood built "from the internet up" (George-Cosh and Brown 2017; Doctoroff 2016). In this project, artificial intelligence systems would control traffic lights for minimal pedestrian and autonomous vehicle congestion, monitor garbage chutes for "pay-as-you-throw" billing, and regulate building and sidewalk temperatures—all powered by extensive sensors that would feed back into the neighborhood's algorithms to enhance efficiency and productivity on a month-to-month basis (Fussell 2018; Scola 2018).

However, the predictive accuracy of AI models is dependent on the scope and complexity of the data they are initially trained on (Alpaydin 2004), a vulnerability that can affect the efficacy and downstream consequences of the technologies they support. Additionally, because these models learn in a way that eventually moves beyond human comprehension, direct intervention by developers becomes difficult; effectively, these models are a black box to programmer and consumer alike (Došilović, Brčić, and Hlupić 2018; Lipton 2016; Murdoch et al. 2019). Therefore, when scrutinizing an AI model, we should not only ask, "are the predictions accurate based on the data?" but also, "is the training data itself biased?" If data used to train the AI are biased, the algorithm will produce biased results (Barocas and Selbst 2016).

II. Exemplified bias

These vulnerabilities have already produced some notable examples of biased AI, highlighting the potential for negative impacts if these technologies are widely incorporated without appropriate oversight. One study quantified the failure rate of facial recognition software from Microsoft, IBM, and Megvii, and found that these failures are a function of both skin color and gender (Buolamwini and Gebru 2018). Lighter-skinned men were identified with an error rate below 1%; lighter-skinned women were misidentified up to 7% of the time. Darker-skinned men were incorrectly identified up to 12% of the time, while darker-skinned women had an error rate up to 35%.

Such biases and other oversights by the tech industry could have outsized impacts if incorporated into major societal changes like autonomous vehicles and smart cities. For instance, facial recognition biases could affect the decision processes of autonomous vehicles maneuvering in accidents, with the potential for racial or gendered discrepancies in the rates of death and injury. In fact, tech industry advancements have already permitted new mechanisms for marginalizing communities of color, such as the alleged housing discrimination practices by Facebook's advertising algorithms, which have elicited a lawsuit from the U.S. Department of Housing and Urban Development (Department of Housing and Urban Development 2019). It is plausible that these and related issues could propagate major discrimination if integrated into city-wide infrastructure, like in Sidewalk Toronto. That project is already facing resistance from the local community over the industry's lack of transparency and lack of appreciation for downstream consequences, issues that are directly correlated to bias in AI. Jim Balsillie, CEO of a major Canadian technology firm, referred to Sidewalk Toronto as "a colonizing experiment in surveillance capitalism attempting to bulldoze important urban, civic, and political issues" (Balsillie 2018). His concerns have been reinforced by prominent resignations from the project, including that of former Ontario Privacy Commissioner Ann Cavoukian, who left after learning third parties would be able to access identifiable information of residents (Kofman 2018). While the core aim of each of these technologies is to help their respective communities, these examples illustrate how often the tech industry fails to appreciate the pitfalls of AI, even propagating systemic bias.

In fact, the infrastructure ramifications of biased technology are already affecting citizens today. Courts in 14 states in the U.S. have adopted risk algorithms for bail assessment (Koepke and Robinson 2018). A ProPublica investigation demonstrated that one such algorithm is almost twice as likely to wrongly identify black defendants as likely to re-offend than white defendants (Larson et al. 2016). A preliminary evaluation of the use of another algorithm showed that when judges departed from its recommendations, they were more than three times as likely to do so in a punitive direction—a human response which could potentially compound the bias already within the algorithm (Santa Cruz County Probation Department 2015). Northpointe, the company responsible for the model examined by ProPublica, refused to release its data sets, citing concerns about compromising intellectual property security—a pervasive attitude in the tech community and one often used to avoid regulation.

It is important to acknowledge that any disparities or inequalities currently present in society will be inherently attached to AI systems and will be able to proliferate unless we make a conscious effort to eliminate them.

III. Policy Landscape

Government regulation of the tech industry is growing, though it is not primarily concerned with the question of bias. Europe has been particularly aggressive in developing a regulatory framework around consumer privacy, exemplified by the passage of the General Data Protection Regulation (GDPR) in April 2016 and implemented in May 2018. The law requires any company that collects or processes data on EU citizens to adhere to new standards, such as requiring companies to anonymize any personal data collected and provide transparent communication channels in the event of a data breach, or else face fines up to 4% of a company's global annual revenue (Regulation (EU) 2016/679 2016). To maintain access to regulated markets, U.S. companies have made some efforts to comply with GDPR and its closest U.S. equivalent, the California Consumer Privacy Act, which was passed in 2018 and will go into effect in 2020 (Assembly Bill 375 2018).

In the U.S. government, leaders in both political parties have called for major changes to the management of the tech industry. Democratic Senator Elizabeth Warren breaking up tech giants (Warren 2019) while Democratic Senator Amy Klobuchar proposed simply taxing their data use (Klobuchar 2019). Additionally, Republican Senator Ted Cruz proposed changes to Section 230 of the Communications Decency Act which would increase companies' liability for content posted by users (U.S. Code 47 Section 230 2018; Cruz 2019). Although these initiatives have been supported by other government officials and departments (Overly 2019; Wakabayashi, Benner, and Lohr 2019; McKinnon and Kendall 2019), none have come to fruition, illustrating the difficulty of implementing meaningful government regulation.

Notably, these regulatory efforts address the intersection of social media, privacy, and free speech. They are not designed—or even concerned with the need—to evaluate the fairness of algorithms. Earlier this year, President Trump signed an executive order making the National Institute of Standards and Technologies (NIST), responsible for setting "appropriate" AI technical standards (Executive Order 13859 2019), primarily for needs within the U.S. government and for promoting U.S. technological advancement. NIST coordinates with other agencies and non-government stakeholders as needed. However, a recent study has brought NIST's practices into question, finding it evaluates facial recognition models using photographic databases of arrested persons, of visa applicants, and even of children who have been exploited for pornography, without requesting permission from the individuals (Keyes, Stevens, and Wernimont 2019). This activity highlights two major conflicts within proper AI evaluation: whether the metrics used to evaluate algorithms were transparently and ethically sourced, as well as who is doing the evaluation. While the databases used by NIST are technologically appropriate (in that they include a sufficient amount of faces of color, a typical failure point for these algorithms), the acquisition process is ethically concerning and raises the question of whether an independent body would be better suited for the task. Additionally, NIST currently evaluates only facial recognition, a small portion of the overall AI landscape.

In a recent development, the Department of Energy just established the Artificial Intelligence and Technology Office on September 6, 2019. The stated mission of this office is to secure American dominance in AI and related research (Department of Energy 2019). Conspicuously, no statement by the office has included any mention of bias mitigation as a consideration for AI development.

Further complicating regulatory prospects, the Information Technology Industry Council (ITI), a major lobbying group representing the tech industry, advocated against over-regulation of AI, stating, "We believe governments should avoid requiring companies to transfer or provide access to technology, source code, algorithms, or encryption keys as conditions for doing business" (Information Technology Industry Council 2017). While ITI has championed regulations regarding data privacy (Meritalk 2018; Information Technology Industry Council 2017), the tech industry remains decidedly resistant to submit to regulations on AI, fearing the loss of competition and entrepreneurship.

This conflicted policy landscape demonstrates the need for an innovative, third-party mechanism to evaluate AI algorithms. To ensure buy-in from the tech community, such novel evaluation must be supplemented by support for continued technical innovation in AI; further, innovative efforts in education can also prepare a workforce attuned to the issue of bias in AI.

IV. Policy proposals

Because this issue involves civil rights, technological advancement, business interests, government action, and the protection of consumers, our solution must be as flexible as it is far-reaching. Success must balance innovation for the advancement of society, safeguards for vulnerable communities, and a conscious dialogue between experts covering a range of cross-sectional disciplines. We believe the Partnership on AI (PAI), a major nonprofit consortium of technology corporations, civil rights advocacy groups, and universities, is uniquely suited to develop a technologically proficient yet socially responsible system for bias mitigation. PAI's "thematic principles," which ground its research and outreach efforts, include "fair, transparent, and accountable AI" and "social and societal influences of AI" (Heer 2018). To this end, they have hosted conferences and sponsored research on AI fairness, including for risk assessment tools in the criminal justice system (Partnership on AI 2019).

Therefore, we propose that PAI launch the following initiatives: 1) a nonprofit accreditation board, independent from industry and government, to evaluate existing and forthcoming technologies, incentivized by tax breaks for successfully accredited products, and 2) expanded educational support to train leaders in the field, increasing the influence of those with technical and societal appreciation of the problem. While individually these proposals support meaningful change in developing unbiased technologies, together they can significantly shift the culture and products of the tech community.

i. Independent accreditation board (FAIR)

To address current and emerging AI technologies, we propose that PAI establish an independent nonprofit accreditation board: the Forum for Artificial Intelligence Regularization (FAIR). FAIR would provide a centralized platform for voluntary fairness evaluation of AI models, using standards specified by the corresponding market. This accreditation board could become analogous to consumer watchdogs launched by nonprofits, such as Consumer Reports, helping provide public information about which products include measures to prevent bias.

Given sufficient growth, FAIR would produce a positive feedback loop: the presence of the accreditation boosts public awareness of this issue, driving consumers to favor accredited products; this in turn creates a market and societal pressure for companies to pursue accreditation. Companies receiving accreditation would then advertise it, thereby further promoting awareness of the issue. This natural progression could be augmented by advertising campaigns from FAIR itself, stressing the public interest value of bias mitigation. For markets in which consumers are not active buyers, such as in the city-wide project of Sidewalk Toronto, this public interest could still provide a positive pressure for the use of FAIR-approved products, similar to the failure of Amazon HQ2 in New York City after residents determined the deal was not in the public interest. Ultimately, FAIR accreditation would become the norm for AI technology.

While there are many ways to evaluate AI fairness, we believe the best way to balance the public service mission of FAIR while maximizing industry participation is by directly evaluating finalized products. Using an emerging methodology, described in Murdoch et al. 2019 as "post-hoc interpretability," FAIR would use benchmarks to statistically determine which relationships a given model has learned and—essential for the context of fairness—what the model has failed to learn. This evaluation process would not require companies to hand over source code or training data sets, thereby protecting IP and promoting industry buy-in. FAIR would rely on robust and representative data sets to perform evaluations, collected in a transparent manner.

Given the wide range of applications of AI, establishing accurate metrics to evaluate different products is nontrivial; for instance, the technical and ethical considerations of facial recognition fairness are not equivalent to those for bail assessment. As a result, FAIR must be a positive collaboration of technical expertise across many AI tasks and the social considerations pertinent to applying this advanced technology, and must bring together industry professionals, academics, civil rights advocates, and government administrators. This partnership, with its transparent process for determining appropriate databases and its coverage of multiple markets, would significantly differentiate FAIR from NIST's current facial recognition evaluation.

Further, as an independent organization, FAIR would provide local, state, and federal governments an opportunity to define financial incentive for work in this area. A clear reference point lies in LEED property tax exemptions, as supported by the nonprofit organization behind LEED, the U.S. Council on Green Buildings (U.S. Green Building Council 2014); these financial incentives could support small and large companies alike in developing bias-mitigation technologies. Additionally, this independence would free legislators from the burden of crafting specific legislation that might limit innovation. Finally, as a nonprofit organization, FAIR would rely on grants for funding—perhaps from the newly established Artificial Intelligence and Technology Office, among other sources—which will minimize conflicts of interest that hinder effective regulation.

Given the rate at which these AI models are introduced and updated, accreditation would not be a one-time status; products would need to be evaluated continually, ideally at every major update and at least semiannually to account for smaller-scale changes. Therefore, we believe the successful implementation of FAIR will create jobs for experts from a wealth of fields: technology, academia, communication, law, and administration. Ultimately, FAIR would serve as an effective regulatory proxy, screening existing products and promoting awareness of a major public concern, all while encouraging the innovation of equitable technologies.

ii. Policy proposal education

Another crucial step in promoting the development of unbiased technologies lies in shifting some of the cultural components of the tech community. For example, two major tenets of the tech industry ethos are "modularity," the idea that a digital task can be distilled to a base set of code which is then used across many diverse applications, and "disruption," a common marketing term for how the tech industry can revolutionize traditional thinking in any non-technical industry. While modularity has been instrumental in faster, more efficient design, and while tech industry disruptions have developed exciting technologies, these philosophies can have major negative consequences when they are applied without due regulation to the complicated, high-consequence arena of social institutions, such as bail assessment (Selbst et al. 2018). Further, emerging research suggests many of the issues related to AI bias are directly linked to the lack of diversity in the tech community (Myers West, Whittaker, and Crawford 2019). These issues must be addressed to promote bias mitigation.

Complementing FAIR's downstream evaluation of completed products, educational reform provides an opportunity to ingrain fairness "upstream" as a central principle in developing AI technologies. To this end, we advocate for the universities within PAI to expand fairness education for computer science undergraduate programs. As members of PAI, these universities have an obligation not only to expand research but also to direct the incoming generations of coders. These universities can help stimulate the market for AI fairness by cultivating a diverse workforce that demands greater emphasis on bias mitigation from their employers. Similar to the LEED Green Associate courses available at many universities, these programs will develop both the regulatory workforce of FAIR itself and the programmers who design accredited products. Training at PAI universities could serve as the benchmark for computer science undergraduate programs across the country.

We can envision a wide range of pedagogical approaches to achieve these goals. Educational programs may focus on compliance-centric tracks, cultivating a particular set of standards akin to ethics training for lawyers. Others may utilize a hands-on approach that trains programmers with state-of-the-art tools for detecting bias in algorithms and subsequently retraining the models. However, we believe the best programs will pursue an interdisciplinary approach, partnering experts in law, ethics, and technology to develop and normalize a meaningful discussion on these complicated, non-modular issues.

We further hope that this culture change will invite engineers who are not currently represented in the homogenous tech community into the fold, by bringing inclusion explicitly to the forefront of their education. We believe that the interdisciplinary track, with its unique collaborations and emphasis on partnering necessary tools with meaningful dialogue, is the path most likely to accomplish this task. As these conscientious engineers enter the industry, they will inflect a meaningful ethical value system into product development, thereby applying a positive pressure that will normalize regulatory and evaluative efforts like FAIR.

Conclusions

Although society stands to benefit immensely from the implementation of AI technologies, we must be proactive in ensuring the equitable implementation of these seismic changes. We have already seen several unfortunate examples of biased models, some with reverberating negative consequences. Given the scant regulatory landscape and the inhibitions of various stakeholders, effective measures against this issue must incorporate adequate feedback channels to build a suite of flexible initiatives. We believe our proposals to better regulate and reward unbiased technologies and expand the educational landscape around fairness are uniquely suited to the skillset and collaboration within the Partnership on AI. Further, these efforts will appease competing stakeholders and effectively address this urgent need.

References

Alpaydin, Ethem. 2004. *Introduction to Machine Learning.* Cambridge: The MIT Press.

Barocas, Solon and Andrew D. Selbst. 2016. “Big Data’s Disparate Impact.” *California Law Review* 104: 671-732.

Balsillie, Jim. 2018. “Sidewalk Toronto has only one beneficiary, and it is not Toronto.” *The Globe and Mail,* October 5.

Bertoncello, Michelle and Dominik Wee. 2015. “Ten ways autonomous driving could redefine the automotive world.” McKinsey & Company. Last modified June 2015. https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/ten-ways-autonomous-driving-could-redefine-the-automotive-world.

Buolamwini, Joy and Timnit Gebru. 2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.” *Proceedings of Machine Learning Research* 81: 1-15.

California Legislature. 2018. “Assembly Bill No. 375, Chapter 35.” California Consumer Privacy Act of 2018. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375.

Chen, Yuche, Jeffrey Gonder, Stanley Young, and Eric Wood. 2019. “Quantifying autonomous vehicles national fuel consumption impacts: A data-rich approach.” *Transportation Research Part A* 122: 134-145.

Cruz, Rafael. 2018. “Facebook, Social Media Privacy, and the Use and Abuse of Data.” Filmed April 10 at Senate Committee on the Judiciary, Senate Committee on Commerce, Science, and Transportation Hearing, Washington, DC. Video, 2:25:08. https://www.judiciary.senate.gov/meetings/facebook-social-media-privacy-and-the-use-and-abuse-of-data.

Department of Energy. 2019. “Artificial Intelligence and Technology Office: Mission.” Accessed September 13, 2019. https://www.energy.gov/science-innovation/artificial-intelligence-and-technology-office.

Department of Housing and Urban Development. 2019. FHEO No. 01-18-0323-8. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375.

Doctoroff, Daniel L. 2016. “Reimagining cities from the internet up.” *Sidewalk Labs Blog*, November 30. https://www.sidewalklabs.com/blog/reimagining-cities-from-the-internet-up.

Došilović, Filip Karlo, Mario Brčić, and Nikica Hlupić. 2018. “Explainable artificial intelligence: A Survey.” *41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*: 210-215.

Enbysk, Liz, Philip Bane, and Stuart Cowan. 2016. “Smart Infrastructure Unlocks Equity and Prosperity for Our Cities and Towns.” *Smart Cities Council Policy Brief*: 1-11.

Fussell, Sidney. 2018. “The City of the Future Is a Data-Collection Machine.” *The Atlantic,* November 21.

George-Cosh, David, and Eliot Brown. 2017. “Google Parent Nears Deal to Build Its Vision of a City in Toronto.” *The Wall Street Journal,* October 4.

Heer, Jeffrey. 2018. “The partnership on AI.” *AI Matters* 4, no. 3: 25-26.

Information Technology Industry Council. 2017. “AI Policy Principles.” *AI Policy Principles Full Report*: 1-5. https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf.

Information Technology Industry Council. 2014. "Data & Privacy." Accessed May 6. https://www.itic.org/policy/data-privacy.

Keyes, Os, Nikki Stevens, and Jacqueline Wernimont. 2019. "The Government Is Using the Most Vulnerable People to Test Facial Recognition Software." *Slate*, March 17. https://slate.com/technology/2019/03/facial-recognition-nist-verification-testing-data-sets-children-immigrants-consent.html.

Klobuchar, Amy. 2019. "GDPR & CCPA: Opt-ins, Consumer Control, and the Impact on Competition and Innovation." Filmed March 12 at Senate Committee on the Judiciary Hearing, Washington, DC. Video, 1:34:17. https://www.judiciary.senate.gov/meetings/gdpr-and-ccpa-opt-ins-consumer-control-and-the-impact-on-competition-and-innovation.

Koepke, John Logan and David G. Robinson. 2018. "Danger Ahead: Risk Assessment and the Future of Bail Reform." *Washington Law Review* 93: 1725-1807.

Kofman, Ava. 2018. "Google's 'Smart City of Surveillance' Faces New Resistance in Toronto." *The Intercept*, November 13. https://theintercept.com/2018/11/13/google-quayside-toronto-smart-city.

Lipton, Zachary C. 2017. "The Mythos of Model Interpretability." *ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*: 1-9. https://arxiv.org/pdf/1606.03490.pdf.

Larson, Jeff, Surya Mattu, Lauren Kirchner, and Julia Angwin. 2016 "How We Analyzed the COMPAS Recidivism Algorithm." *ProPublica*, May 23. https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

McKinnon, John D. and Brent Kendall. 2019. "States to Move Forward With Antitrust Probe of Big Tech Firms." *The Wall Street Journal,* August 19.

Mercer, Christina and Tom Macaulay. 2019. "Which Companies Are Making Driverless Cars?" *TechWorld,* June 27. techworld.com/picture-gallery/data/-companies-working-on-driverless-cars-3641537.

Meritalk. 2018. "ITI Floats Priorities For Data Privacy Policy, Legislation." *Meritalk*, October 22. https://www.meritalk.com/articles/iti-floats-priorities-for-data-privacy-policy-legislation.

Murdoch, James, Chandan Singh, Karl Kumbier, Reza Abbasi-Asl, and Bin Yu. 2019. "Interpretable machine learning: definitions, methods, and applications." *Proceedings of the National Academy of Sciences*: 1-11.

Myers West, Sarah, Meredith Whittaker, and Kate Crawford.2019. "Discriminating Systems: Gender, Race, and Power in AI." *AINow Institute*: 1-33.

National Highway Traffic Safety Administration. 2008. "National Motor Vehicle Crash Causation Survey: Report to Congress." *U.S. Department of Transportation*, DOT HS 811 059: 1-47.

National Highway Traffic Safety Administration. 2015. "Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey." *U.S. Department of Transportation,* DOT HS 812 115: 1-2.

Overly, Steven. 2019. "Conservatives, liberals slam Facebook bias audit." *Politico*, August 20. https://www.politico.com/story/2019/08/20/facebook-audit-conservative-bias-amid-criticism-1469287.

"Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation)." 2016. *Official Journal of the European Union*.

Russell, Stuart J. and Peter Norvig. 2003. *Artificial Intelligence: A Modern Approach*. Edinburgh Gate: Pearson Education.

Santa Cruz County Probation Department. 2015. "Alternatives to Custody Report 2."

Scola, Nancy. 2018. "Google Is Building a City of the Future in Toronto. Would Anyone Want to Live There?" *Politico*, June 29. https://www.politico.com/magazine/story/2018/06/29/google-city-technology-toronto-canada-218841.

Selbst, Andrew D, Danah Boyd, Sorelle Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2018. "Fairness and Abstraction in Sociotechnical Systems." *ACM Conference on Fairness, Accountability, and Transparency* 1, no. 1: 1-17.

Shirer, Michael, and Serene Da Rold. 2018. "IDC Forecasts Smart Cities Spending to Reach $158 Billion in 2022, with Singapore, Tokyo, and New York City Among Top Spenders." *IDC Spending Guide Press Release,* July 23. https://www.idc.com/getdoc.jsp?containerId=prUS44159418.

Trump, Donald J. 2019. "Executive Order on Maintaining American Leadership in Artificial Intelligence." *United States Office of the Federal Registrar,* EO 13859.

United States Congress. 2006. "47 U.S. Code § 230: Protection for private blocking and screening of offensive material." *Code of Federal Regulations*, title 47, section 230. https://www.govinfo.gov/app/details/USCODE-2011-title47/USCODE-2011-title47-chap5-subchapII-partI-sec230.

U.S. Green Building Council. 2014. "Good to Know: Green Building Incentive Strategies." *U.S. Green Building*

Council, May 2. https://www.usgbc.org/articles/good-know-green-building-incentive-strategies-0.

Warren, Elizabeth. 2019. "Here's how we can break up Big Tech." *Medium,* March 8. https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c.

Wakabayashi, Daisuke, Katie Benner and Steve Lohr. 2019. "Justice Department Opens Antitrust Review of Big Tech Companies." *The New York Times*, July 23.

Andy Sanchez is a PhD candidate at Cornell University in chemical and biomolecular engineering. His research explores a new mechanism for antibody detection, particularly for use in Lyme disease diagnosis, and he has done extensive work in science communication advocacy. He obtained his Master of Fine Arts in creative writing from New York University and his Bachelor of Science in chemical engineering from Texas A&M University.

Kevin McDermott is a PhD candidate at Cornell University in physics, expecting to finish in the fall of 2019. His dissertation is an analysis of data collected by the CMS experiment at the CERN LHC, searching for long-lived particles using delayed photons. Kevin obtained a Bachelor of Science in physics with a minor in theology at the University of Notre Dame in the spring of 2013.

Alicia Cintora is a PhD candidate at Cornell University in materials science and engineering (MSE). Her dissertation is focused on developing synthetic methods for nitroxide polymers as well as finding novel solid-state applications. Alicia obtained a Bachelor of Science in MSE from the University of Illinois at Urbana-Champaign in 2014.

Andy, Kevin, and Alicia were partners on a policy project as part of a graduate level science policy course at Cornell University. This paper is an offshoot of that work.