



**Stanford University**  
Human-Centered  
Artificial Intelligence

# Stanford HAI Artificial Intelligence Bill of Rights

**A WHITE PAPER FOR STANFORD'S INSTITUTE FOR  
HUMAN-CENTERED ARTIFICIAL INTELLIGENCE**

Co-led by  
Michele Elam  
Rob Reich

## **DISCLAIMER**

*The Stanford Institute for Human-Centered Artificial Intelligence (HAI) is a nonpartisan research institute, representing a range of voices. The views expressed in this White Paper reflect the views of the authors.*

---

**JANUARY 2022**

# Contributors

## PRINCIPAL AUTHORS

**Michele Elam** is the William Robertson Coe Professor of Humanities in the English department at Stanford University, a faculty associate director of the Institute for Human-Centered Artificial Intelligence, and a race & technology affiliate at the Center for Comparative Studies in Race and Ethnicity.

**Rob Reich** is a professor of political science and, by courtesy, professor of philosophy at Stanford University. He is the director of the Center for Ethics in Society and co-director of the Center on Philanthropy and Civil Society (publisher of the Stanford Social Innovation Review), and a faculty associate director of the Institute for Human-Centered Artificial Intelligence.

## ACKNOWLEDGMENTS

We thank Daniel Zhang, Russell Wald, Michael Sellitto, Justin Sherman, and Benjamin Bronkema-Bekker for their contribution to this white paper, Dan Ho, Jennifer King, and Tina Huang for their helpful comments, and Jeanina Casusi, Joe Hinman, Nancy King, Shana Lynch, Stacy Peña, and Michi Turner for their help in preparing this publication.



## **Stanford HAI Artificial Intelligence Bill of Rights**

The Stanford Institute for Human-Centered Artificial Intelligence (HAI) offers the following submission for consideration in response to the [Request for Information](#) (RFI) by the White House Office of Science and Technology on public and private sector uses of biometric technologies. While our intention with this response is to examine the uses of biometric technologies as per the RFI, we also take the implications of broader artificial intelligence (AI) technologies into consideration. Biometrics and AI are uniquely intertwined. We can decouple these two technologies only to an extent and must understand both the full impacts of AI and how the biometric paradigm disproportionately affects marginalized groups and exacerbates inequities. Following Dr. Eric Lander and Dr. Alondra Nelson's recent call for a bill of rights to safeguard the American public against powerful technologies in an opinion piece for *Wired*,<sup>1</sup> we produce this set of six principles that recommends:

- Ensuring AI-powered biometric systems are developed and deployed in a manner that supports fundamental democratic values with respect to the rule of law, basic civil liberties, and universal human rights.
- Safeguarding fairness and rights to nondiscrimination.
- Ensuring transparency and explainability during development and due process rights in application.
- Strengthening participation of civil society organizations in important AI use and governance conversations.
- Embedding accountability measures into system design.
- Enhancing citizen education in relation to AI and its impacts.

---

### **Ensuring AI-powered biometric systems are developed and deployed in a manner that supports fundamental democratic values with respect to the rule of law, basic civil liberties, and universal human rights**

AI is arming governments with unprecedented capabilities to track, surveil, and monitor individuals.<sup>2</sup> For example, the pervasive tracking of individuals in public spaces via surveillance cameras, voice recognition systems, and social media not only interferes with individuals' rights to privacy, but also affects their rights to freedom of speech, expression, and association.<sup>3</sup> The

---

<sup>1</sup> Eric Lander and Alondra Nelson, "Americans Need a Bill of Rights for an AI-Powered World," *Wired* (2021), <https://www.wired.com/story/opinion-bill-of-rights-artificial-intelligence/>.

<sup>2</sup> Steven Feldstein, "The Road to Digital Unfreedom: How Artificial Intelligence Is Reshaping Repression," *Journal of Democracy* 30, no. 1 (2019): pp. 40-52, <https://doi.org/10.1353/jod.2019.0003>.

<sup>3</sup> Toni M Massaro, Helen Norton, and Margot E Kaminski, "SIRI-OUSLY 2.0: What Artificial Intelligence Reveals About the First Amendment," *Minnesota Law Review* 101 (2017): p. 2481, <https://scholarship.law.umn.edu/mlr/179>; Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown Books, 2016); Jeramie D. Scott,

use of such technologies by law enforcement potentially alters how individuals exercise such rights, including fundamental democratic guarantees of lawful social and political participation and protest. In short, biometric tracking can chill the potential of participatory democracy.<sup>4</sup> Moreover, nation-states with known human rights abuses and unaccountable government institutions can exploit the power of AI and biometrics, which poses a direct threat to open democratic societies.<sup>5</sup>

Therefore, incorporating democratic and human-centered values in both public and private uses of AI should be the central concern in developing and deploying AI-powered biometric technologies. First, it is important to identify democratic principles that developers of AI systems should abide by, namely the rule of law, civil liberties, and universal human rights. The benchmark of democratic principles rather than democratically sanctioned policy is essential here. All governments must recognize that “democracy supporting technologies” does not simply mean developers complying with legislation or rules passed by a government of a democratic society, which could include illiberal democracies or democracies facing significant stress and decline.<sup>6</sup> The key tension is to navigate between companies adhering to the decisions made by governments in democratic societies and adherence to a set of democratic ideals. Building institutional arrangements for the latter is far more difficult and far more important.

In the United States, the commercial sector has been the primary incubator of AI innovation, but leaving AI technology free to develop without guardrails presents a wide range of dangers. Steps to embed democratic and human-centered values into AI development could include updating existing regulations on antitrust and nondiscrimination or enhancing rule-of-law practice such as designating oversight agencies to monitor and audit AI systems.<sup>7</sup> Moreover, funding academic research or public-private partnerships with an emphasis on incorporating democratic principles could foster the protection of individual liberties, human rights, and data privacy during the development phase.

Finally, building a pro-democracy technology alliance to coordinate strategies against the abuse of AI-powered biometric technologies can slow the development of AI-enabled digital authoritarianism.<sup>8</sup> President Biden convened the inaugural Summit for Democracy in December 2021 that brought together leaders from government, civil society, and the private sector to tackle

---

“Social Media and Government Surveillance: The Case for Better Privacy Protections for Our Newest Public Space,” *Journal of Business & Technology Law* 12 (2017): p. 151, <http://digitalcommons.law.umaryland.edu/jbtl/vol12/iss2/2>; Joel R. Reidenberg, “Privacy in Public,” *University of Miami Law Review*, 69, no.1 (2014): <https://repository.law.miami.edu/umlr/vol69/iss1/6>; Woodrow Hartzog and Evan Selinger, “Surveillance as Loss of Obscurity,” *Washington and Lee Law Review*, 72, no.3 (2015): pp.1343-87, <https://scholarlycommons.law.wlu.edu/wlulr/vol72/iss3/10>.

<sup>4</sup> Feldstein, 2019.

<sup>5</sup> Nicholas Wright, “How Artificial Intelligence Will Reshape the Global Order,” *Foreign Affairs*, November 24, 2021, <https://www.foreignaffairs.com/articles/world/2018-07-10/how-artificial-intelligence-will-reshape-global-order>.

<sup>6</sup> For example, the Freedom House published democracy ratings and scores for more than 100 countries, which show some so-called democratic countries with lower scores due to the absence of or decline in political rights and civil liberties: <https://freedomhouse.org/report/summit-democracy/2021/summit-democracy-ratings-scores>.

<sup>7</sup> Karl Manheim and Lyric Kaplan, “Artificial Intelligence: Risks to Privacy and Democracy,” *Yale Journal of Law & Technology* 21 (2019): p. 106, [https://yjolt.org/sites/default/files/21\\_yale\\_j.l.\\_tech\\_106\\_0.pdf](https://yjolt.org/sites/default/files/21_yale_j.l._tech_106_0.pdf).

<sup>8</sup> This is part of the recommendations in the National Security Commission on Artificial Intelligence (NSCAI) final report: National Security Commission on Artificial Intelligence, *Final Report* (Washington, D.C.: 2021), <https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf>.

challenges confronting democracies and bolster democratic institutions. Specifically, the summit participants recommended we “invest in the development, use, and governance of technology that advances democracy and human rights.”<sup>9</sup> Failure to bolster democratic institutions in the face of rapidly developing AI would diminish their relevance, and the private sector and authoritarian governments would become more powerful.

### *Ideas to Explore*

- Update antitrust and civil rights regulations and enhance rule-of-law practice to compel the private sector to adopt a democracy-supporting and individual-rights-protecting vision of AI science and technology.
- Increase investment and fund academic research or public-private partnerships in AI-powered biometric innovations infused with democratic values.
- Support norm-setting efforts on AI-powered biometric technologies at the international level to support human rights and rule of law.
- Coordinate with other democratic countries and global civil society organizations to promote democratic values in AI and place export controls and human rights sanctions with its allies to deny authoritarian regimes resources used to develop mass surveillance technologies.

### **Safeguarding fairness and rights to nondiscrimination**

From healthcare and education to finance and the environment, governments and industries around the world are embracing AI-powered approaches to doing business and solving social problems. AI applications have the potential to reduce discrimination caused by structural and systemic inequities as well as human subjectivity, yet they can also introduce or exacerbate bias leading to discriminatory decisions that create injustice and undermine public trust in AI. Research has shown that AI-powered biometric technologies have disproportionate impacts on women, underrepresented racial and ethnic groups, the LGBTQIA+, the economically disadvantaged, and people with disabilities.<sup>10</sup>

For example, a hiring algorithm used by Amazon to screen résumés scored the applications of men higher than those of women—including by downgrading résumés that contained words like “women’s” and degrees from all-female colleges.<sup>11</sup> In addition, researchers have found major commercial facial recognition systems, built on the processing and analysis of biometric data (specifically, face images), often perform far more accurately on lighter-skinned faces compared

---

<sup>9</sup> “Summit for Democracy Summary of Proceedings,” The White House (The United States Government, December 23, 2021), <https://www.whitehouse.gov/briefing-room/statements-releases/2021/12/23/summit-for-democracy-summary-of-proceedings/>.

<sup>10</sup> David Danks and Alex John London, “Algorithmic Bias in Autonomous Systems,” *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, 2017, <https://doi.org/10.24963/ijcai.2017/654>; Joy Buolamwini and Timnit Gebru, “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification,” *Proceedings of Machine Learning Research* 81 (2018): pp. 1-15, <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>; Ben Hutchinson et al., “Unintended Machine Learning Biases as Social Barriers for Persons with Disabilities,” *ACM SIGACCESS Accessibility and Computing*, no. 125 (October 2019): p. 1, <https://doi.org/10.1145/3386296.3386305>.

<sup>11</sup> Jeffrey Dastin, “Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women,” Reuters, October 10, 2018), <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>; Rob Reich, Mehran Sahami, and Jeremy M. Weinstein, *System Error: Where Big Tech Went Wrong and How We Can Reboot* (New York, NY: Harper, an imprint of HarperCollinsPublishers, 2021).

to darker-skinned faces and on faces of male individuals compared to faces of female individuals.<sup>12</sup> They performed worst of all on the faces of Black women.<sup>13</sup> Moreover, it is important to acknowledge that discrimination is not simply non-recognition, or being unseen. In other words, it is not just a problem of data or solving accuracy. We must address issues outside of the technology realm, such as certain populations being overly-surveilled by institutions of power.<sup>14</sup>

State governments are increasingly allowing or requiring AI-powered risk assessment tools in criminal justice sentencing decisions,<sup>15</sup> which raises important concerns about explainability, equal protection, and due process. Due process is a constitutional right provided by the Fifth and Fourteenth Amendments that ensures fairness in legal proceedings and safeguards defendants against erroneous deprivations of life, liberty, and property. Using such tools may introduce bias into the decision-making of sentencing without any potential for explaining a black-box decision, thereby undermining the due process right.<sup>16</sup> For example, if a dataset is missing input from particular populations or infected by past discriminatory practices, using that dataset to build an AI-powered sentencing tool may yield results that are unfair or inequitable to certain underrepresented or protected groups.

Therefore, designing fair algorithms is more important than ever. One of the challenges in making AI systems fair lies in deciding how to make mathematically tractable the ideal of “fairness.” The case of the COMPAS algorithm, a proprietary algorithmic-based risk assessment tool developed to help parole boards assess recidivism risks, provides an example of such a challenge. A study by ProPublica found that while the tool correctly predicted recidivism for Black and white defendants at roughly the same rate, it misclassified the Black and white defendants over a two-year follow-up period.<sup>17</sup> Developers of the COMPAS technology argued that the algorithm was fair since it predicted the same likelihood of recidivism across all groups regardless of race.<sup>18</sup> But the ProPublica study shows that Black defendants are more likely to be classified as higher recidivist risks than white defendants, and Black defendants who do not re-offend are predicted to be riskier than white defendants who do not re-offend.

While there has been much research on how to quantitatively evaluate fairness, there is no general consensus.<sup>19</sup> The problem is twofold. First, defining fairness is not an easy task—there can be multiple reasonable conceptions of fairness. The normative approach to ensure fairness treats everyone the same. But such an approach fails to take into account that not everyone starts

---

<sup>12</sup> Buolamwini and Gebru, 2018.

<sup>13</sup> Buolamwini and Gebru, 2018.

<sup>14</sup> Ruha Benjamin, 2019, *Race After Technology: Abolitionist Tools for the New Jim Code* (Polity, 2020); Catherine D'Ignazio and Lauren F. Klein, *Data Feminism* (Cambridge, MA: The MIT Press, 2020)

<sup>15</sup> John Lightbourne, “Damned Lies & Criminal Sentencing Using Evidence-Based Tools,” *Duke Law & Technology Review* 15 (May 14, 2017): p. 327, <https://scholarship.law.duke.edu/dltr/vol15/iss1/16>.

<sup>16</sup> Black-box here refers to the system or model not being able to explain its output or decision, see: Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge, MA: Harvard University Press, 2016).

<sup>17</sup> Jeff Larson and Julia Angwin, “How We Analyzed the COMPAS Recidivism Algorithm,” ProPublica, May 23, 2016, <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>.

<sup>18</sup> Julia Dressel and Hany Farid, “The Accuracy, Fairness, and Limits of Predicting Recidivism,” *Science Advances* 4, no. 1 (2018), <https://doi.org/10.1126/sciadv.aao5580>.

<sup>19</sup> Mulligan et al., “This Thing Called Fairness: Disciplinary Confusion Realizing a Value in Technology,” *Proceedings of the ACM on Human-Computer Interaction* 3, no. 119 (2019): p.119, <https://dl.acm.org/doi/10.1145/3359221>.

off on equal ground, with equal resources and access to the same opportunities.<sup>20</sup> For example, consider the question of public school funding. It might seem that equal per pupil funding is what fairness requires. But do children with special needs—physical or cognitive handicaps—deserve identical funding as children without special needs? Fairness here seems to require greater funding for those with special needs. So to promote fairness and achieve equity, one should not aim to treat everyone the same, but to remove structural barriers to equal access to the opportunities, with the recognition that even this alone is insufficient to mitigate systemic inequities that persist. In other words, equal access is but one, if essential, step towards “fairness.”

Second, fairness could be construed as the property of an individual or of groups.<sup>21</sup> In the example of criminal justice sentencing, the former means defendants with identical criminal histories receive the same predictive score when reviewed by algorithms, whereas the latter refers to the share of ethnic minority defendants who are rated with the same level of recidivist risk as that of ethnic majority ones. While both conceptions of fairness seem reasonable and both have limitations, it is challenging to adopt both at the same time when designing algorithms.<sup>22</sup>

Where there is no singular definition of fairness, AI systems designed to maintain the status quo—as the COMPAS algorithm is set up to do—have the potential to perpetuate and exacerbate inequality. Therefore, to build fair AI systems, we need to define and quantify what we mean by a fair outcome—that is, truly equal access to resources at the beginning and attention to the backend impacts, whether or not there were claims a process was “fair.”

### *Ideas to Explore*

- Identify fairness considerations and approaches up-front, and involve multi-stakeholders, such as experts in the relevant domain and across disciplines, in the conversation.
- Explore a legally viable path for algorithmic fairness under current constitutional doctrines.<sup>23</sup>
- Develop testing and monitoring mechanisms to detect and mitigate fairness-related harms.

### **Ensuring transparency and explainability during development and due process rights in application**

Many elements of the rule of law and democratic societies rely on access to information. Transparency facilitates public discourse, evidence-based policymaking, regulatory oversight,

---

<sup>20</sup> Benjamin, 2019; Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Profile Books, 2019); Safiya Umoja. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018); Neda Atanasoski and Kalindi Vora, *Surrogate Humanity: Race, Robots, and the Politics of Technological Futures* (Durham: Duke University Press, 2019).

<sup>21</sup> Reich, Sahami, and Weinstein, 2021.

<sup>22</sup> Reich, Sahami, and Weinstein, 2021.

<sup>23</sup> Daniel E. Ho and Alice Xiang, “Affirmative Algorithms: The Legal Grounds for Fairness as Awareness,” *The University of Chicago Law Review Online*, October 30, 2020, <https://lawreviewblog.uchicago.edu/2020/10/30/aa-ho-xiang/>.

judicial review, and journalistic scrutiny.<sup>24</sup> In the case of AI-powered biometric technologies, understanding how AI systems work and how algorithms arrive at their decision promotes public trust in the responsible use of such technologies and ensures democratic norms during development and application.

However, private companies often keep crucial information about the inner workings of AI systems under wraps. This makes it difficult to understand exactly how biometric information is processed and where errors in biometric data analysis may originate. The black-box nature of AI systems, or not knowing how an algorithm reaches its conclusion, only adds to this opacity. This results in an information gap that prevents the public from knowing or responding to any AI missteps. Employers, for example, are increasingly using video surveillance and webcams that collect biometric information like eye movements or facial expressions to track worker performance.<sup>25</sup> Without understanding how those technologies work, employees are unable to get a justification or explanation of an AI-based decision and lose their ability to appeal to such a decision as a result.

While we might not know how an AI system produces unintended results, we can know what was intended for the system in the first place. Requiring the transparency of training data and procedure, documentation detailing performance characteristics, and iterations of algorithms, for example, sheds light on the explainability of AI systems. Moreover, access to reliable explanations for different audiences is equally important. The “right to explanation” clause in the European Union General Data Protection Regulation (GDPR) requires businesses to provide rationales for decisions made by the AI system.<sup>26</sup> Such an explanation could be used, for example, as a basis for the public’s right to appeal against an automated decision.

### *Ideas to Explore*

- Develop audit trail requirements and documentation of AI-powered biometric systems that cover all steps of the AI development process, which could include model architecture, training data, records of exhibited bias and previous predictions, etc.
- Require private companies to provide a right to explanation of decisions made by automated or AI systems.
- Implement executive and legislative actions to mandate developers of AI systems to provide access for auditing via independent regulatory agencies, such as the Federal Trade Commission (FTC), or third-party organizations.
- Implement bias and safety bug-bounty programs, allowing individuals to report algorithmic bias or security vulnerabilities to an organization and receive rewards or compensation, for AI systems to increase incentives for broader scrutiny of AI systems.

---

<sup>24</sup> Beth Simone Noveck, “Rights-Based and Tech-Driven: Open Data, Freedom of Information, and the Future of Government Transparency,” *Yale Human Rights and Development Law Journal*, 2017, [https://digitalcommons.nyls.edu/cgi/viewcontent.cgi?article=2208&context=fac\\_articles\\_chapters](https://digitalcommons.nyls.edu/cgi/viewcontent.cgi?article=2208&context=fac_articles_chapters).

<sup>25</sup> Darrell M. West, “How Employers Use Technology to Surveil Employees,” Brookings Institution, January 5, 2021, <https://www.brookings.edu/blog/techtank/2021/01/05/how-employers-use-technology-to-surveil-employees/>.

<sup>26</sup> “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,” EUR-Lex, May 4, 2016, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>.



## **Strengthening participation of civil society organizations in important AI use and governance conversations**

To help address the challenges AI-powered biometric technologies raise, the role of civil society is more important than ever. Civil society consists of nongovernmental and not-for-profit organizations outside the public and private sectors “that have a presence in public life [and] express the interests and values of their members and others, based on ethical, cultural, political, scientific, religious or philanthropic considerations.”<sup>27</sup> Such organizations bring important perspectives to discussions over how to harness the power of AI to benefit all parts of society while also minimizing any negative impacts on protected groups. Civil society also forms an important bulwark against authoritarian and unaccountable governments.

Civil society has a history of advocating a human-centered approach on behalf of underrepresented or marginalized populations.<sup>28</sup> For example, in the face of new emerging risks during past industrial revolutions, organized associations, such as faith-based charities and labor unions, and other friendly societies, helped improve worker conditions and reduce new emerging risk.<sup>29</sup> Today, civil society can play a similar role to shape the future of biometric technologies.

As governments and private companies confront the issue of engaging with the public and earning their trust when deploying AI-powered biometric technologies in sensitive sectors like healthcare and finance, civil society can bring balance and perspective to the conversation. It can fill blind spots and ease fears among those who stand to lose the most in a world dominated by unchecked algorithms and data.

### *Ideas to Explore*

- Mandate federal and local governments to consult representatives from community and civil society organizations when developing rules and regulations related to AI.
- Require private companies to conduct stakeholder consultations with civil societies and seek out their guidance during the development of AI technologies to augment human capabilities and ensure the inclusion of diverse developers of AI systems, representative datasets, and nondiscriminatory practices.

## **Embedding accountability measures into system design**

The democratic governance of AI-powered biometric technologies requires strong accountability mechanisms to keep these systems in check—for instance, so others can intervene in the event of dangerously incorrect decisions. This goes hand in hand with transparency: While the preceding transparency and explainability point focuses on developers, it is also essential for third parties to be able to access that information and hold systems accountable in response. A lack of

---

<sup>27</sup> “Civil Society,” World Bank, n.d., <https://www.worldbank.org/en/about/partners/civil-society/overview>.

<sup>28</sup> Wendy Pojmann, *Migration and Activism in Europe since 1945* (Palgrave Macmillan, 2016).

<sup>29</sup> World Economic Forum, *Civil Society in the Fourth Industrial Revolution: Preparation and Response*, 2019, <https://www.weforum.org/whitepapers/civil-society-in-the-fourth-industrial-revolution-preparation-and-response>.

accountability can be incredibly dangerous in certain settings, such as with opaque AI applications making recommendations or decisions in a medical setting.<sup>30</sup>

Part of this is technical. Independent, third-party organizations need access to AI system data and code to run technical audits or impact assessments. New York City, for example, adopted a law to require third-party bias audits of algorithms used by employers in hiring or promotions.<sup>31</sup> Third parties like researchers, civil society groups, community organizations, and regulators also need access to transparency and explainability information used internally by an organization, so they can fully understand and assess the design and deployment of AI-powered biometric technologies. Industry can also embed accountability mechanisms into system design, such that human interventions are possible (e.g., in clinical settings and law enforcement settings) as necessary.

Other parts of internal accountability regimes, however, are not technical. Companies developing AI tools need not only internal ethical guidelines (e.g., policies on transparency and explainability) but also human involvement in accountability structures, such as “rank-and-file employee representation on the board of directors, external ethics advisory boards, and the implementation of independent monitoring and transparency efforts.”<sup>32</sup> Congress, for example, has introduced bills in the past few years to mandate companies to conduct annual audits or impact assessments of AI algorithms.<sup>33</sup>

#### *Ideas to Explore*

- Implement executive and legislative actions to allow third-party auditor access to AI data and source code, as well as other transparency and explainability information, for the purposes of external researcher, civil society, and regulator assessments.
- Consider best practices that could be recommended to industry for embedding accountability mechanisms to test the outcome or results of its AI systems.

#### **Enhancing citizen education in relation to AI and its impacts**

Understanding of AI is becoming a more important part of modern civic participation as AI-powered biometric applications are already shaping how citizens receive political advertisements, political news, and voting information around elections through social media platforms. Akin to the rapid emergence of computer science (CS) coursework in secondary schools, developing and enhancing education around AI and biometrics would empower citizens to better participate in the emerging policy and civic discourse around these technologies.

---

<sup>30</sup> Helen Smith, “Clinical AI: Opacity, Accountability, Responsibility and Liability,” *AI & SOCIETY* 36, no. 2 (2020): pp. 535-545, <https://doi.org/10.1007/s00146-020-01019-6>.

<sup>31</sup> “Automated Employment Decision Tools,” The New York City Council, December 11, 2021, <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9&Options=ID>.

<sup>32</sup> Meredith Whittaker et al., “AI Now Report 2018,” AI Now Institute, December 2018, [https://ainowinstitute.org/AI\\_Now\\_2018\\_Report.pdf](https://ainowinstitute.org/AI_Now_2018_Report.pdf).

<sup>33</sup> Congress.gov. “H.R.2231 - 116th Congress (2019-2020): Algorithmic Accountability Act of 2019.” April 11, 2019. <https://www.congress.gov/bills/116th-congress/house-bill/2231>; Congress.gov. “Text - S.2968 - 116th Congress (2019-2020): A bill to provide consumers with foundational data privacy rights, create strong oversight mechanisms, and establish meaningful enforcement.” December 3, 2019. <https://www.congress.gov/bills/116th-congress/senate-bill/2968/text>.

Notably, this can also occur through government agencies doing public outreach, exploring curricula in public schools, and supporting related research in higher education. It is incumbent upon the federal government to initiate, design, and promote educational programs to better inform the American public on the ethical and rights-based challenges of AI and biometric technologies.

More than half of US high schools now offer CS courses, in light of the growing role of software in daily life and in recognition of the growing value of exposure to programming.<sup>34</sup> There are still troubling inequities in the quality, accessibility, and tailoring of this education across racial, gender, and socioeconomic class lines. For instance, one study found that significant performance gaps in CS coursework emerge between male and female students as early as 10 years old.<sup>35</sup> Nonetheless, this kind of education exposes citizens at an early age to CS, which touches every aspect of their lives. This can increase familiarity in general by engaging with powerful technologies, and it also gives future voters substantive exposure to software that shapes daily life. It must be done, however, with substantial focus on technology ethics—something missing in many current CS education programs—and in this case, with a focus on the ethics of designing, developing, and deploying AI and biometrics systems. This should include understanding the ethical implications of these technologies, the ethical frameworks and policies in place around their development, and what future ethics should look like, from engineers to regulators.

CS and AI coursework are also becoming increasingly important in higher education. According to the 2021 AI Index, the number of undergraduates completing CS degrees in 2019 is three times higher than the number in 2010.<sup>36</sup> On the graduate level, the number of AI and machine learning-specialized CS PhD graduates among all new CS PhDs in 2020 is 8.6 percentage points larger than in 2010—the most significant growth, relative to 18 other specializations.<sup>37</sup>

AI is a multidisciplinary field. It is important to develop and enhance AI education programs outside of the CS domain, including humanities, arts, and social sciences. Citizen education is increasingly necessary to enable democratic participation in the age of AI. Understanding these technologies is increasingly vital for everyday life.

#### *Ideas to Explore*

- Build AI themes, including AI ethics and the technology’s impact on society, into Department of Education recommendations on technology and CS education.
- Engage with civil society to understand best practices and substantive options for AI educational programs.

---

<sup>34</sup> Alyson Klein, “More than Half of High Schools Now Offer Computer Science, But Inequities Persist,” *Education Week* (November 5, 2021), <https://www.edweek.org/teaching-learning/more-than-half-of-high-schools-now-offer-computer-science-but-inequities-persist/2021/11>.

<sup>35</sup> Jennifer Tsan, Kristy Elizabeth Boyer, and Collin F. Lynch, “How Early Does the CS Gender Gap Emerge?,” *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 2016, <https://doi.org/10.1145/2839509.2844605>.

<sup>36</sup> Daniel Zhang et al., “The AI Index 2021 Annual Report,” March 2021, [https://aiindex.stanford.edu/wp-content/uploads/2021/11/2021-AI-Index-Report\\_Master.pdf](https://aiindex.stanford.edu/wp-content/uploads/2021/11/2021-AI-Index-Report_Master.pdf).

<sup>37</sup> Zhang et al., 2021.

- Conduct a study on existing CS and other multidisciplinary education programs related to AI nationally to understand equity and inclusion issues so as to inform the development of AI education.
- Ensure executive branch policies on AI are communicated clearly to the public, with discussion of their possible impacts on society and on communities.

As lead authors, we proudly submit this response on behalf of our colleagues and the Stanford Institute for Human-Centered Artificial Intelligence (HAI).



Michele Elam  
William Robertson Coe Professor of  
Humanities, Stanford University  
Faculty Associate Director, Stanford Institute  
for Human-Centered Artificial Intelligence  
(HAI)



Rob Reich  
Professor of Political Science, Stanford  
University  
Faculty Associate Director, Stanford Institute  
for Human-Centered Artificial Intelligence  
(HAI)