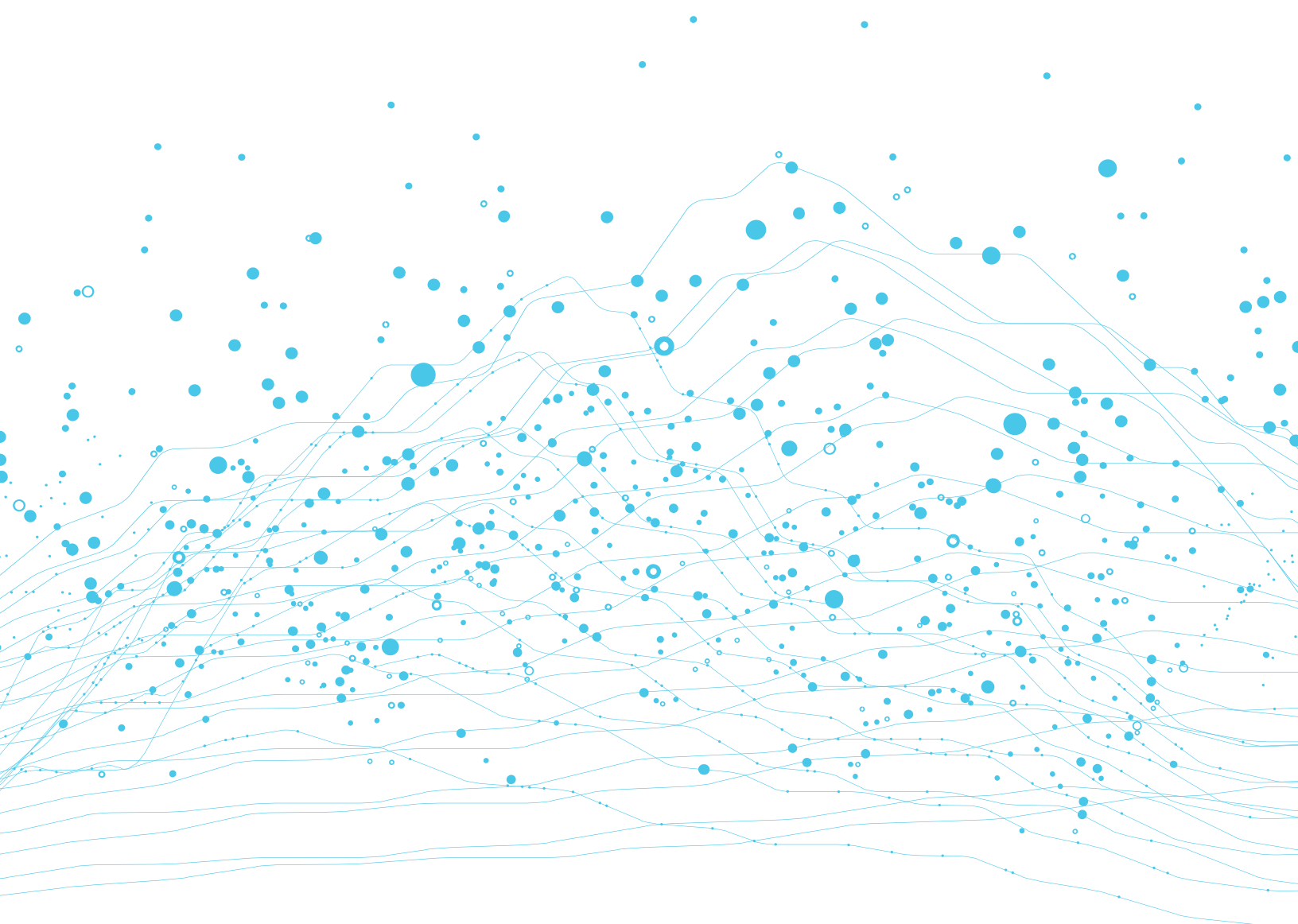


# Impact Assessments: Supporting AI Accountability & Trust

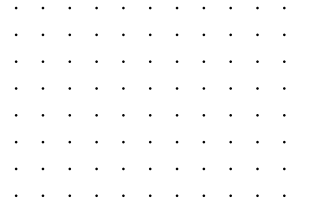


# Contents

Executive Summary .....	3
Introduction .....	6
Overview of AI Accountability Tools .....	8
1. Impact Assessments .....	8
2. Third-Party Auditing.....	14
3. Conformity Assessment.....	19
Recommendations.....	22
About The Authors.....	23



# Executive Summary



The deployment of artificial intelligence (AI) systems is increasing rapidly, with effects that are already beginning to transform society. AI systems offer vast potential to improve and streamline economic efficiency, support better decision-making, and provide data-driven predictions that lead to better outcomes.

However, there are persistent concerns related to the potential of these technological improvements to perpetuate some harms. To tap AI's full potential, there must be broad confidence that it has been developed ethically and is being used responsibly. Policymakers and leaders from industry, academia, and civil society are therefore looking to tools that support robust accountability and cultivate trust.

There are several approaches to algorithmic accountability. **Impact assessments** are already widely used in different fields, with privacy impact assessments being a well-established tool. For AI systems, algorithmic impact assessments (AIAs) help mitigate risk by enabling teams across an organization to review a system's objectives, design, and intended purpose before its use.



In recent years, leading companies have designed templates, frameworks, and interactive AIA tools, with civil society organizations and governments also making progress on AIAs. *This paper introduces the basic elements of AIAs and their benefits for AI accountability.*

In contrast, **third-party auditing** for AI systems is hindered by a lack of relevant technical and professional standards. While work towards standards is ongoing, the present quality of AI audits can vary wildly, potentially undermining trust in AI. Among other issues, some technology companies may seek out auditors offering more favorable methods, criteria, and scope, thus reducing accountability. Meanwhile, the European Union is considering a **conformity assessment** approach modeled on its rulebook for consumer product safety. While incrementally better than an auditing mandate, this approach still raises questions regarding standards and the possibility that prescriptive requirements will inhibit future innovation.

We conclude that AIAs have three important advantages over other AI accountability tools: (1) they are *familiar* to organizations already conducting impact assessments for privacy and data protection; (2) they are *practical* because they do not rely on technical standards, which are currently nascent; (3) they are *future-proof* because they can adapt as AI systems and AI governance evolve.

In the United States, several recent policy proposals would have significant implications for AI accountability. The leading federal privacy bill – the American Data Privacy and Protection Act – includes robust provisions for AI regulation including impact assessments. Notably, its authors removed a third-party audit requirement before the bill's approval by a House committee in July 2022. Other bills take different approaches although impact assessments are often favored.



## In the near term, we recommend the following:

- **Policymakers should encourage algorithmic impact assessments.** Governments, industry organizations, and companies have taken the first steps, with some governments already adopting AIAs as their preferred AI accountability tool in regulations and draft proposals. Meanwhile, governance frameworks, including NIST's AI Risk Management Framework, explicitly incorporate impact assessments or enable compatibility with them.
- **Policymakers should support harmonization.** With policymakers at various levels of government and around the globe looking to develop regulatory frameworks for AI technology, harmonization is essential. Policymakers should avoid creating a patchwork of inconsistent requirements by looking to best practices in the field and engaging in robust international cooperation.
- **Policymakers should support standards work related to AI systems.** Standards development is typically the culmination of years of consensus-building among experts from academia, the private sector, government, and civil society. Because standards development can be a lengthy process, policy support should include securing funding and resources for standards development, bolstering knowledge-sharing, and promoting strategic engagement with international bodies.
- **Policymakers may wish to establish regulatory backstops to undergird companies' self-conducted algorithmic impact assessments.** This could help promote oversight, accountability, and public trust, while retaining the adaptability and advantages of algorithmic impact assessments.



# Introduction

AI promises to accelerate economic, social, and technological progress. In the coming decades, AI will power advances in sectors ranging from health care to education to manufacturing, making the economy more efficient, promoting opportunity, and raising standards of living in the United States (U.S.) and around the world. By 2030, AI is expected to add \$15.7 trillion to the global economy.<sup>1</sup>

At its core, AI evaluates problems and provides insights to assist with decision-making. However, there are persistent concerns related to these technological advancements. There are concerns, for example, that AI systems can operate as “black boxes,” evading scrutiny with opaque inputs and operations.<sup>2</sup> Some of the potential harms and misuses of AI include well-publicized cases involving law enforcement,<sup>3</sup> mortgage lending,<sup>4</sup> video hiring practices,<sup>5</sup> and national security.<sup>6</sup>



1 Dr. Anand S. Rao & Gerard Verweij, *Sizing the prize What's the real value of AI for your business and how can you capitalise?*, PwC (2017), <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>.

2 Dillon Reisman, Jason Schultz, Kate Crawford, and Meredith Whittaker, *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AI Now Inst., at 6, (Apr. 2018), <https://ainowinstitute.org/aiareport2018.pdf>.

3 Kashmir Hill, *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*, N.Y. Times (Dec. 29, 2020), <https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html>.

4 Scott Simon, *How Some Algorithm Lending Programs Discriminate Against Minorities*, Nat'l Pub. Radio (Nov. 24, 2018), <https://www.npr.org/2018/11/24/670513608/how-some-algorithm-lending-programs-discriminate-against-minorities>.

5 Electronic Privacy Information Center, *Complaint and Request for Investigation, Injunction, and Other Relief (in re: HireVue, Inc.)* (Nov. 6, 2019), [https://epic.org/wp-content/uploads/privacy/ftc/hirevue/EPIC\\_FTC\\_HireVue\\_Complaint.pdf](https://epic.org/wp-content/uploads/privacy/ftc/hirevue/EPIC_FTC_HireVue_Complaint.pdf).

6 Atlantic Council Digital Forensic Research Lab, *Hacked new program and deepfake video spread false Zelenskyy claims* (Mar. 16, 2022), <https://www.atlanticcouncil.org/blogs/new-atlanticist/russian-war-report-hacked-news-program-and-deepfake-video-spread-false-zelenskyy-claims/#deepfake>.

To address these risks, policymakers are developing regulatory frameworks to build trust in AI systems by ensuring that they are transparent and that organizations using AI can be held accountable. While enforcing an accountability mechanism of some form is widely supported, efforts vary across jurisdictions, with some approaches creating unintended barriers to progress and beneficial AI applications.<sup>7</sup> For example, states like New York and California have proposed mandatory third-party auditing of AI systems as one solution.<sup>8</sup> Yet this approach faces significant hurdles including a lack of technical standards or widely accepted professional standards. Another model championed by the European Union (EU) is a “conformity assessment” approach derived from the EU’s consumer product safety regulations. While the EU’s model has some advantages in select contexts, it is not well-suited to AI software, which is ultimately quite different from a physical product.

**Algorithmic impact assessments** (AIAs) represent a tried-and-true accountability tool. Impact assessments have been leveraged for decades in other policy contexts, including privacy and data protection, and are commonplace across U.S. and global regulatory regimes. With AI adoption at an early stage, AIAs can contribute to the development of best practices and inform future policy discussions,

while remaining adaptable to technology that may vary across organizations and evolve over time. Ultimately, AI accountability remains central to the public’s trust in and use of AI systems—and AIAs promise to be an important part of achieving that public trust. We expand further on AIAs in the sections below.

This report describes and analyzes three different accountability tools: AIAs, third-party audits, and conformity assessments. We also describe emerging frameworks from both the public and private sectors to further illustrate the current AI accountability landscape. Finally, we provide recommendations for policymakers to build stronger and more transparent AI accountability frameworks and advance the benefits of AI in society.




<sup>7</sup> Kirsten Martin & Ari Waldman, *Are Algorithmic Decisions Legitimate? The Effect of Process and Outcomes on Perceptions of Legitimacy of AI Decisions*, *J. Bus. Ethics* (2022). <https://doi.org/10.1007/s10551-021-05032-7>.

<sup>8</sup> *A Local Law to amend the administrative code of the city of New York, in relation to automated employment decision tools*, enacted Dec. 11, 2021, NYC law no. 2021/144, <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9> (visited July 25, 2022); see also Workplace Technology Accountability Act, AB 1651, Chapter 5, §§ 1561 – 1562(a).

# Overview of AI Accountability Tools

In this section, we describe three distinct AI accountability tools: impact assessments, which we view as the most effective and pragmatic approach available today, third-party audits, and conformity assessments.



Each of these tools influence the design and deployment of AI in specific systems and implicate emerging policy debates. As policymakers' interest in AI accountability grows, it is crucial to understand which tools are most effective in practice. For example, policymakers could unintentionally damage consumer confidence and public trust in AI if they rush toward a third-party audit approach that lacks mature standards and faces hurdles in implementation, ultimately degrading the legitimacy of the accountability tool and undercutting innovation.<sup>9</sup>

## 1. Impact Assessments

An *impact assessment* is a type of accountability tool that defines the intended purpose of a specific project and considers potential unintended consequences and effects on individuals and society.<sup>10</sup> We begin by describing the development and use of impact

<sup>9</sup> Martin & Waldman, *supra* note 7, at 5.

<sup>10</sup> *What is Impact Assessment*, Org. for Econ. Co-operation & Dev <https://www.oecd.org/sti/inno/What-is-impact-assessment-OECDImpact.pdf>. (last visited Jul 22 2022)



assessments in existing contexts, before turning to algorithmic impact assessments.

Impact assessments are widely used in different fields like city planning, criminal sentencing reform, and public health community programs, among others. An environmental impact assessment, for example, considers the effects of a project or policy on the natural environment and certain populations of animals.<sup>11</sup> State governments and local municipalities also issue fiscal impact assessments for proposed programs.

Within the technology field, *privacy impact assessments* (PIAs) are an effective, well-established tool utilized by public and private sector organizations. When conducting a PIA, organizations evaluate their management of personally identifiable information in the context of a digital product or service. After considering potential impacts on users' privacy, PIAs catalog compliance with relevant data collection, retention, and protection standards and legal requirements.<sup>12</sup> PIAs are commonly required by U.S. states as well, with privacy laws in Virginia<sup>13</sup> Colorado<sup>14</sup> and Connecticut<sup>15</sup> requiring companies to carry out “data protection assessments” for activities that demonstrate a “heightened risk of harm to a consumer,” such as using sensitive data for targeted advertising.<sup>16</sup>

To ensure accountability, impact assessments are often buttressed by regulatory oversight. PIAs are occasionally required by the Federal Trade Commission (FTC) in consent decrees reached with companies that have engaged in unfair or deceptive privacy practices.<sup>17</sup> In some instances, the FTC may also compel a company to complete a PIA after a settlement, which the regulator can retain for potential future enforcement action.<sup>18</sup> In addition, the U.S. Environmental Protection Agency has enforcement power over environmental impact assessments, acting as a source of legitimacy to ensure the assessments are valid and that organizations took steps to mitigate identified risks.<sup>19</sup>



<sup>11</sup> See *Health impact assessment*, World Health Org., <https://www.who.int/health-topics/health-impact-assessment> (last visited Jul. 13, 2022).

<sup>12</sup> Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, Madeleine Claire Elish, and Jacob Metcalf, *Assembling Accountability: Algorithmic Impact Assessment for the Public Interest*, Data & Soc’y, at 33 (June 29, 2021), <https://datasociety.net/wp-content/uploads/2021/06/Assembling-Accountability.pdf>; see also Marguerite Reardon, *Facebook’s FTC consent decree deal: What you need to know*, CNET (Apr. 14, 2018), <https://www.cnet.com/news/politics/facebook-s-ftc-consent-decree-deal-what-you-need-to-know/>.

<sup>13</sup> See S.B. 1392 § 591-576(A)-(F), <https://lis.virginia.gov/cgi-bin/legp604.exe?212+ful+CHAP0036+pdf> (specifying a risk-based approach to Virginia’s data protection assessments which should identify and balance the benefits from data processing against the ‘potential risks to the rights of the consumer’).

<sup>14</sup> See Sec. 7(1), Colorado Privacy Act, Senate Bill 21-190, 73d Leg., 2021 Regular Sess. (Colo. 2021), to be codified in Colo. Rev. Stat. (“C.R.S.”) C.R.S. §§ 6-1-1302(c)(II)(C), 6-1-1309(4) (outlining that the Attorney General can request the data protection assessment upon request, with enforcement power to ‘impose penalties where violations occur’).

<sup>15</sup> See Connecticut Data Privacy Act (“CTDPA”), S.B. 6, 2022 Gen. Assemb., Reg. Sess. §§ 6(1)–(3); 8 (Conn. 2022).

<sup>16</sup> See Colorado Privacy Act, § 6-1-1309(2)(a); Connecticut Data Privacy Act § 8; S.B. 1392 § 591-576(A)(1)-(5).

<sup>17</sup> *Id.* at 33.

<sup>18</sup> *Id.* at 20 (explaining that a regulator may retain the quasi-private PIA “for potential future action, thus standing as a proxy for the public”).

<sup>19</sup> Moss et al., *supra* note 12, at 14.



## *Example: GDPR data protection IAs.*

Article 35 of Europe’s General Data Protection Regulation (GDPR) requires companies to perform Data Protection Impact Assessments (DPIAs) whenever data processing “is likely to result in a high risk to the rights and freedoms of natural persons.”<sup>20</sup> High-risk scenarios include, for example, when a hospital collects health data in its information system or when sensitive data is archived from a clinical trial. Notably, because companies complying with GDPR are familiar with DPIAs, most companies developing or deploying AI will likely already be familiar with impact assessments.

Under GDPR, a DPIA must include:

*“a description of the ‘processing operations’ and the purpose of the processing; an assessment of the necessity of processing in relation to the purpose; an assessment of the risks to individual rights and freedoms; and importantly, the measures a company will use to address these risks and demonstrate GDPR compliance, including security measures....”<sup>21</sup>*

The European Data Protection Board’s DPIA guidance reflects an understanding of the dynamic and fast-paced nature of technology products and services.<sup>22</sup> For example, the guidance suggests that DPIAs be “updated throughout the lifecycle of the project” and that “carrying out a DPIA is a continual process, not a one-time exercise.”<sup>23</sup>

<sup>20</sup> Article 35(1) GDPR.

<sup>21</sup> Article 35(7) GDPR; see also Margot E. Kaminski & Gianclaudio Malgieri, Algorithmic Impact Assessments Under the GDPR: Producing Multi-Layered Explanations, 11 INT’L DATA PRIVACY L. 125, 130 (prefacing the DPIA requirements prior to expanding on EDPB guidance).

<sup>22</sup> See generally Kaminski & Malgieri, *supra* note 21, at 126 (detailing that the EDPB serves as an advisory body and provides guidelines for companies and regulators; while the guidelines are non-binding, Article 70 of the GDPR “states that the EDPB is required to issue opinions, guidelines, and recommendations in order to ensure the consistent application of the GDPR”).

<sup>23</sup> See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is ‘likely to result in high-risk’ for the purposes of Regulation 2016/679, 14. See also Kaminski & Malgieri, *supra* note 21, at 130.



## Algorithmic impact assessments.

Companies conduct an AIA to document how they identify, test for, and mitigate the risks posed by an AI system they are using. AIAs typically take the form of a list of questions (see example below) and may be completed by multiple teams within an organization.

The goal of AIAs is to consider different risks posed by the technology in question. AIAs help mitigate risk by enabling teams across an organization to review an AI system’s objectives, design, and intended purpose before its use.<sup>24</sup> This analysis is cost-effective but also paves the way for meticulous review of the details and outcomes of a complex project.<sup>25</sup> The earlier that problems, such as unintended bias, are uncovered and reviewed by companies using AI systems, the sooner they can be remediated.<sup>26</sup> AIAs also promote accountability by requiring documentation and evidence of the decision-making processes. They ask open-ended questions with a bottom-up reporting structure,<sup>27</sup> requiring different stakeholders within a company to describe their decision-making and the steps they took to address potential risks to end users.

In recent years, companies and other industry leaders have designed and suggested templates, frameworks, and interactive AIA tools. These developments demonstrate that leaders in the AI field recognize they must move (and are moving) AIAs from high-level principles to practice. For example, Workday has called for risk-based regulation and established a “Trustworthy by Design” AI policy proposal.<sup>28</sup> The Workday proposal outlines basic elements of a successful AI governance framework, including engaging the C-suite, compliance personnel devoted to AI, and robust AIAs with proper documentation, training, and cross-company resources.<sup>29</sup> Microsoft also released a responsible AI framework, incorporating guidance for AI tools like speech-to-text technology and facial recognition technology.<sup>30</sup> Its impact assessment template contains six core pillars:

<sup>24</sup> *Id.*

<sup>25</sup> *Id.*

<sup>26</sup> *Id.* at 123.

<sup>27</sup> *Id.* at 148.

<sup>28</sup> *Building Trust in AI and ML*, Workday whitepaper (2021) <https://www.workday.com/content/dam/web/en-us/documents/whitepapers/building-trust-in-ai-ml-principles-practice-policy.pdf>.

<sup>29</sup> *Id.* at 10-11.

<sup>30</sup> Natasha Crampton, *Microsoft’s framework for building AI systems responsibly*, Microsoft On the Issues (June 21, 2022), <https://blogs.microsoft.com/on-the-issues/2022/06/21/microsofts-framework-for-building-ai-systems-responsibly/>.

accountability, transparency, fairness, reliability and safety, privacy and security, and inclusiveness.<sup>31</sup> These company proposals complement BSA | The Software Alliance’s “Confronting Bias Framework,” released in 2021, which sets out a process for organizations to perform an AI impact assessment.<sup>32</sup>

Below are selected questions from an AIA template developed by the Ada Lovelace Institute, a non-profit based in the United Kingdom, to illustrate what an AIA may look like in practice.<sup>33</sup> The Ada Lovelace template is tailored to AI used in healthcare. Although impact assessments share common requirements, AIAs can be tailored to the context where AI is used, including credit, education, employment, and housing.

## 1. High-level project information

This section asks background information on your project for viewers of this AIA who may not be familiar with it. It also covers high-level context questions to inform thinking on ethical considerations and potential harms in the latter sections of this template.

### Your project

**1.a.i) Describe the purpose of your project.** This should be a concise summary, of no more than 250 words. You can write this in the form of an abstract for a paper. Assume your audience doesn't have much technical knowledge—perhaps you are explaining this to a stranger

**1.a.ii) Describe the intended uses of your project.**

**2.a. Could this project lead to the creation of exacerbation of inequalities of unlawful discrimination against particular communities?** For example, through worsening differential access to care? What might your current plans for evaluating or monitoring bias and fairness overlook?

Reflexive exercise	Participatory workshop	Synthesis

**3.b. What kind of socio-environmental requirements are necessary for the success of this system in operation?** E.g. stable connection to the internet, training for doctors and nurses, collaboration between particular clinical administration staff etc.

In answering this question, consider which stakeholders will use this system, how they would optimally interact or work together for the system to succeed, how information would be shared (and with whom), and what social, technical and workflow dependencies may need to exist. You might also consider what kinds of infrastructure stakeholders will need to use this system successfully.

Reflexive exercise	Participatory workshop	Synthesis

**3.c. What are likely challenges/hurdles to achieving the best-case scenario?**

Reflexive exercise	Participatory workshop	Synthesis

<sup>31</sup> Microsoft Responsible AI Impact Assessment Template (June 2022), <https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2022/06/Micro-soft-RAI-Impact-Assessment-Template.pdf>.

<sup>32</sup> BSA – The Software Alliance, *Confronting Bias: BSA’s Framework to Build Trust in AI*, <https://ai.bsa.org/wp-content/uploads/2021/06/2021bsaaibias.pdf>.

<sup>33</sup> NMIP algorithmic impact assessment (AIA) template, Ada Lovelace Inst., [https://docs.google.com/document/d/12HXv7Kb4dZLnA0BkL7DiccBxoq-Slg2meB-sUBq\\_QQOI/edit](https://docs.google.com/document/d/12HXv7Kb4dZLnA0BkL7DiccBxoq-Slg2meB-sUBq_QQOI/edit) (last visited Aug. 1, 2022).



Civil society and governments have also made progress in developing and implementing AI impact assessments. In 2018, AI Now, a research institute focused on the implications of AI in social domains, published its findings on AIAs including opportunities and challenges for public agencies.<sup>34</sup> Their assessment emphasizes transparency and due process. Data & Society, a non-profit organization, has identified and outlined ten “constructive components” to guide impact assessments.<sup>35</sup>

In the U.S., the bipartisan American Data Privacy and Protection Act (ADPPA) would require users of AI systems posing a “consequential risk of harm” to consumers to carry out an impact assessment. These assessments would then be submitted to the Federal Trade Commission, who would be charged with enforcing the law. The ADPPA was passed out of the House Energy & Commerce Committee in a historic 53-2 vote in July 2022.<sup>36</sup> The National Institute of Standards & Technology (NIST) has also highlighted the importance of impact assessments as an accountability tool in its draft “AI Risk Management Framework,” which the agency is developing at the direction of Congress.<sup>37</sup>

In 2020, the Canadian government developed an online AIA questionnaire for agencies that consists of approximately 60 questions related to “business process, data, and system designed decisions.”<sup>38</sup> Some of the questions address the early stages of project development, and others relate to the affected users and communities, the level of human involvement, and the specific sector. Once an impact level is determined for an AI system, agencies must follow certain actions to mitigate risks and harms.<sup>39</sup> In June 2022, the Canadian federal government introduced comprehensive AI legislation, the Artificial Intelligence and Data Act (AIDA).<sup>40</sup> The AIDA is intended to focus on AI systems that have the greatest impact on people and requires an impact assessment to determine a system’s risk level.<sup>41</sup>

---

<sup>34</sup> See generally Reisman et al., *supra* note 2, at 15-20.

<sup>35</sup> Moss et al., *supra* note 12, at 14.

<sup>36</sup> Press Release, Bipartisan E&C Leaders Hail Committee Passage of the American Data Privacy and Protection Act (Jul. 20, 2022), <https://energycommerce.house.gov/newsroom/press-releases/bipartisan-ec-leaders-hail-committeepassage-of-the-american-data-privacy>.

<sup>37</sup> AI Risk Management Framework, NAT'L INST. OF SCIENCE & TECH, <https://www.nist.gov/itl/ai-risk-management-framework> (last updated Mar. 14, 2022); see also AI Risk Management Framework, Second Draft NAT'L INST. OF SCIENCE & TECH. at 9, 28 (Aug. 18, 2022), [https://www.nist.gov/system/files/documents/2022/08/18/AI\\_RMF\\_2nd\\_draft.pdf](https://www.nist.gov/system/files/documents/2022/08/18/AI_RMF_2nd_draft.pdf).

<sup>38</sup> See generally *Algorithmic Impact Assessment (AIA) tool*, Government of Canada, <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html> (last updated Apr. 19, 2022).

<sup>39</sup> *Id.*

<sup>40</sup> Digital Charter Implementation Act, 2022, Part 3, Bill C-27, First Session, Forty-fourth Parliament, 70-71 Elizabeth II, 2021-2022 (introduced June 16, 2022).

<sup>41</sup> *Id.*; see also Maya Medeiros and Jesse Beatson, *Bill C-27: Canada's first artificial intelligence legislation has arrived*, Norton Rose Fulbright (June 23, 2022), <https://www.nortonrose-fulbright.com/en/knowledge/publications/55b9a0bd/bill-c-27-canadas-first-artificial-intelligence-legislation-has-arrived>.



## 2. Third-Party Auditing

*Third-party auditing* represents another approach to AI accountability, albeit more difficult to implement soundly in the near-term. Third-party audits generally take one of two forms. In the first, the auditor accesses and reviews an algorithm's source code, the data it ingests, and outcomes it produces.<sup>42</sup> In this scenario, the auditor designs mock exercises to test outcomes on multiple types of individuals. In the second form, the auditor conducts interviews and/or oversees workshops to review an algorithm's outputs.<sup>43</sup>

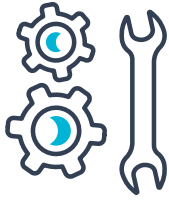
While third-party audits are effective tools in industries with mature technical standards and established professional norms such as privacy and cybersecurity,<sup>44</sup> third-party auditing for AI systems is not yet a well-developed practice. Unlike privacy and cybersecurity fields and as discussed in more detail below, technical and professional standards related to AI are in the early stages. Consequently, mandating audits before these prerequisites are in place could unintentionally harm consumer trust and transparency in AI. *Only once critical building blocks are in place can third-party AI audits serve as an important accountability tool to complement impact assessments.*

---

<sup>42</sup> Alfred Ng, *Can Auditing Eliminate Bias from Algorithms?*, The Markup (Feb. 23, 2021), <https://themarkup.org/ask-the-markup/2021/02/23/can-auditing-eliminate-bias-from-algorithms>.

<sup>43</sup> *Id.*

<sup>44</sup> HIPAA Privacy, Security, and Breach Notification Audit Program, Dept. of Health and Human Serv., <https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/audit/index.html> (last updated Dec. 17, 2020) (outlining the objectives from the HHS Office of Civil Rights when conducting external privacy audits).



## Lack of standards.

Many issues with third-party auditing arise from the current lack of technical and professional standards that are critical for cultivating trust in, and accountability through, audits. AI technical standards are still in development, with leading AI experts agreeing that there are no consensus standards for auditing AI systems.<sup>45</sup> NIST recently described testing standards for AI as “underdeveloped.”<sup>46</sup> The preeminent global standards development body, the International Standards Organization (ISO), is in the early stages of its AI work program.<sup>47</sup> In academia, Stanford University’s Center for Human-Centered AI (HAI) launched a public competition in July 2022 to solicit new ideas for AI audits, underscoring that AI audits are viewed as nascent tools among leading experts in the field.<sup>48</sup> As one Stanford HAI paper concluded, AI audits may hold great promise, but are “not always easy to conduct” and “do not always yield discrete conclusions.”<sup>49</sup>

The situation regarding privacy and cybersecurity audits is quite different. For example, third-party privacy and cybersecurity audits rely upon established standards from bodies like the ISO.<sup>50</sup> Technical standards are also incorporated into the NIST Cybersecurity and Privacy Frameworks, which serve as common reference points against which companies can audit their practices.<sup>51</sup> Mature technical standards also underpin regulator-approved privacy and data protection certifications such as the European Cloud Code of Conduct.<sup>52</sup>

45 See Katharine Miller, *Radical Proposal: Third-Party Auditor Access for AI Accountability*, Stan. Inst. for Human-Centered Artificial Intelligence (Oct. 20, 2021), <https://hai.stanford.edu/news/radical-proposal-third-party-auditor-access-ai-accountability> (underscoring scholar Deb Raji’s assertion that “algorithmic auditing is a nascent field with no professional codes of conduct or standards for what constitutes a thorough audit”).

46 See AI Risk Management Framework, Second Draft, *supra* note 37, at 30.

47 ISO/IEC JTC 1/SC 42, *Artificial Intelligence*, <https://www.iso.org/committee/6794475.html> (last visited July 19, 2022) (outlining the structure of current ISO AI standards).

48 *AI Audit Challenge*, Stanford University Center for Human-Centered AI, <https://hai.stanford.edu/policy/ai-audit-challenge> (last visited July 19, 2022); see also Mozilla Open Source Audit Tooling (OAT) Project, Mozilla Found., <https://foundation.mozilla.org/en/what-we-fund/fellowships/oat/> (last visited Dec. 10, 2022) (establishing a project which supports algorithmic auditors with resources and to develop standards).

49 Danaë Metaxa and Jeff Hancock, *Using Algorithm Audits to Understand AI*, Stan. Univ Human-Centered Artificial Intelligence (Oct. 2022), <https://hai.stanford.edu/sites/default/files/2022-10/HAI%20Policy%20Brief%20-%20Using%20Algorithm%20Audits%20to%20Understand%20AI.pdf>.

50 ISO/IEC 27001, *Information Security Management Standards*, <https://www.iso.org/isoiec-27001-information-security.html> (last visited July 19, 2022).

51 *Framework for Improving Critical Infrastructure Cybersecurity*, Nat’l Inst. of Science & Tech. (Apr. 16, 2018), <https://www.nist.gov/cyber-framework/framework>; *The NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management*, Nat’l Inst. of Science & Tech. (Jan. 2020), <https://www.nist.gov/privacy-framework>.

52 *The EU Cloud Code of Conduct (CoC)*, <https://eucoc.cloud/en/home> (last visited Aug. 1, 2022).

Importantly, there is no consensus around the professional standards for AI auditors. Auditors typically maintain professional bodies to institute baseline criteria, maintain professional ethics, and educate staff to meet the market demand for audits. For example, System and Organization Control (SOC) audits are conducted by Certified Public Accountants and governed by the American Institute of Certified Public Accountants. Likewise, educational bodies such as the International Association of Privacy Professionals and the International Information System Security Certification Consortium offer professional certifications and training to meet the demand for qualified privacy and cybersecurity experts. While there is considerable attention on AI governance, no comparable educational or governing body currently exists for third-party AI auditing.

*Practical problems and potential unintended consequences.* The lack of widely accepted standards for third-party AI audits creates practical problems and potential unintended consequences.

*First*, without technical standards, the quality of AI audits varies wildly between competing third-party consulting firms and may reduce trust in AI. Not all audits offered by third-party firms are equal, and the lack of standards means that companies can seek out an AI auditor offering more favorable methods, criteria, and scope. This may create financial and ethical incentives that companies can game to seek favorable results. A group of Stanford University and Berkeley University AI researchers recently concluded that “clear and well-scoped audit standards,” among other considerations, are needed to “ensure that audits promote, rather than degrade overall [AI] accountability.”<sup>53</sup> It is inevitable that AI-related standards will one day mature, but in the meantime, steps can be taken to adopt tools like AIAs that can drive accountability today.

*Second*, even if a quality audit is conducted, the lack of accepted technical standards can impede accountability and weaken public trust in both audits and the technology.<sup>54</sup> Some groups have suggested that, given the lack of robust standards, organizations may use third-party audits for reputational purposes, often after controversy, or to implement piecemeal solutions. For example, a software vendor faced criticism that its video recruiting technology was biased because it used

---

<sup>53</sup> Inioluwa Deborah Raji, Colleen Honigsberg, Peggy Xu and Daniel Ho, *Outsider Insight, Designing a Third Party Audit Ecosystem for AI Governance*, 2022 AAAI/ACM Conf. on AI, Ethics, and Soc’y (AIES’22), <https://arxiv.org/pdf/2206.04737.pdf>.

<sup>54</sup> *Id.*



predictive algorithms to evaluate a candidate’s “employability.”<sup>55</sup> Follow-on news coverage suggested that a subsequent third-party audit focused only on a narrow use case and failed to verify remediation practices.<sup>56</sup> A tool that, due to lack of clarity, is open to such mischaracterization or potential misuse by organizations will do little to build trust in AI technology or ensure accountability.<sup>57</sup>

*Third*, even if widely accepted technical and professional standards for AI audits are established, third-party auditing may continue to present some challenges. For example, third-party audits could create information security and intellectual property rights concerns by necessitating intimate auditor access to a company’s most sensitive technology, such as source code and training data. Conversely, the use of nondisclosure agreements may prevent auditing firms or the community at large from learning and improving through comparing methodologies or developing best practices—an important practice in the absence of settled audit standards.<sup>58</sup>

*Fourth*, qualifications in the AI auditing field are still being developed and defined.<sup>59</sup> One auditing firm may conduct thorough, comprehensive investigations with adherence to strict standards, while another may stamp a seal of approval on an algorithm according to a spurious methodology.<sup>60</sup> Rumman Chowdhury, founder of the algorithmic auditing company Parity, notes that the lack of professional standards across different auditing firms is a key issue deserving scrutiny: “[t]here are plenty of people out there who are willing to call something an audit, make a nice-looking website and call it a day and rake in cash with no standards.”<sup>61</sup> Thus, premature mandatory third-party audit requirements may have the unintended consequence of harming AI accountability and public trust while existing tools like AIAs can build trust in AI technology.

---

55 Electronic Privacy Information Center, *Complaint and Request for Investigation, Injunction, and Other Relief (in re: HireVue, Inc.)*, *supra* note 5, at 5, 11.

56 Alex C. Engler, *Independent auditors are struggling to hold AI companies accountable*, Fast Company (Jan. 26, 2021), <https://www.fastcompany.com/90597594/ai-algorithm-auditing-hirevue>.

57 *Id.*

58 *Id.*

59 The organization ForHumanity is seeking to fill this gap with research and frameworks, although it does not seem to have engaged organizations developing and deploying AI systems at scale. See <https://forhumanity.center/> (last visited Sept. 7, 2022).

60 See Ng, *supra* note 42.

61 *Id.*



## [New York City's AI auditing law.](#)

New York City enacted a law on “automated employment decision tools” that will take effect on April 15, 2023.<sup>62</sup> The law requires employers to conduct annual bias audits of “software-driven tools that substantially assist in making decisions to hire or promote.”<sup>63</sup> As written, however, the law leaves open essential questions about how the bias audits will be conducted, what a bias audit entails, and who may conduct them.<sup>64</sup> The law also requires a summary of the audit to be published on the employer’s website. News reports have found that the lack of clear guidelines about AI audits has caused confusion among New York City’s employers, as even advocates of the law recognize the lack of standards for AI bias.<sup>65</sup> Local officials have issued draft implementing rules in an effort to address this uncertainty.<sup>66</sup> Nonetheless, New York City’s experience underscores that while lawmakers are motivated to address legitimate concerns around bias, making an AI auditing requirement work in practice can be fraught with complications and unintended consequences.

---

62 *A Local Law to amend the administrative code of the city of New York, in relation to automated employment decision tools*, enacted Dec. 11, 2021, NYC law no. 2021/144, <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9> (visited July 25, 2022); see also *New York City Enacts Law Restricting Use of Artificial Intelligence in Employment Decisions*, Gibson Dunn, Dec. 27, 2021, <https://www.gibsondunn.com/new-york-city-enacts-law-restricting-use-of-artificial-intelligence-in-employment-decisions/>.

63 *A Local Law to amend the administrative code of the city of New York, in relation to automated employment decision tools* § 20-871; see also Matthew Jedreski, Erik Mass, and K.C. Halm, *New York City’s Groundbreaking New Law Will Require Audits of AI and Algorithmic Systems That Drive Employment Decisions*, Davis Wright Tremaine LLP, Dec. 13, 2021, <https://www.dwt.com/blogs/artificial-intelligence-law-advisor/2021/12/nyc-employment-ai-bias-audit-law>.

64 *Id.* § 20-870; see also Roy Maurer, *New York City to Require Bias Audits of AI-Type HR Technology*, Soc’y for Human Res. Mgmt. (Dec. 20, 2021), <https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/new-york-city-require-bias-audits-ai-hr-technology.aspx>.

65 Richard Vanderford, *New York’s Landmark AI Bias Law Prompts Uncertainty*, Wall St. J. (Sept. 21, 2022), <https://www.wsj.com/articles/new-yorks-landmark-ai-bias-law-prompts-uncertainty-11663752602>

66 *Notice of Public Hearing and Opportunity to Comment on Proposed Rules*, New York City Dep’t of Consumer and Worker Prot. (Dec. 2022), <https://rules.cityofnewyork.us/wp-content/uploads/2022/12/DCWP-NOH-AEDTs-1.pdf> (clarifying originally proposed rules and deferring enforcement of the law until April 2023).

### 3. Conformity Assessment

While not yet finalized, the EU’s proposed Artificial Intelligence Act (AI Act) is a landmark proposal for a comprehensive legal framework for AI. The AI Act would adopt a risk-based regulatory approach and leverage the EU’s New Legislative Framework to require conformity assessments for defined “high-risk” categories of AI systems.<sup>67</sup> As a general matter, a “conformity assessment” is a set of processes whereby a company demonstrates that its product, service, or system meets the requirements of a particular standard.<sup>68</sup> In the EU, conformity assessments originated in product safety legislation, as a tool to regulate physical products posing specific safety risks. The AI Act is the first regulatory framework of its kind to apply conformity assessments to non-physical products like software.

The AI Act’s mandates would differ depending on the level of risk posed by an AI system. The bill aims to prevent unacceptable risks to “health, safety, and fundamental rights” through extensive obligations on “high-risk” AI providers.<sup>69</sup> Certain technologies may be banned outright,<sup>70</sup> while AI applications that are not banned or deemed “high-risk” will face minimal regulation.<sup>71</sup>

67 Proposal for a Regulation of the European Parliament and of the Council, (EU) 2021/206 (Apr. 4, 2021); see also Mauritz Kop, *EU Artificial Intelligence Act: The European Approach to AI*, Transatlantic Antitrust and IPR Dev. (2021), <https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/>.

68 See, e.g., *Certification & Conformity*, Int’l Standards Org. (ISO), <https://www.iso.org/conformity-assessment.html> (last visited Jul. 13, 2021).

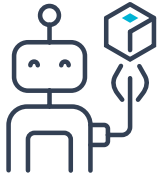
69 See Proposal for a Regulation of the European Parliament and of the Council, *supra* note 66 § 3(3.5).

70 Luciano Floridi, Matthias Holweg, Mariarosaria Taddeo, Javier Amaya Silva, Jakob Mökander, and Yuni Wen, *capAI, A Procedure for Conducting Conformity Assessment of AI systems in line with the EU Artificial Intelligence Act* (March 23, 2022), <https://ssrn.com/abstract=4064091>.

71 *What is the EU AI Act?*, Future of Life Inst., <https://artificialintelligenceact.eu/> (visited Jul. 13, 2021)



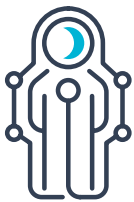




## [Implementation for high-risk systems.](#)

Under the proposed AI Act, the type of conformity assessment required would depend on the specific nature of the high-risk AI system. High-risk AI systems used as safety components of consumer products, such as AI systems within physical objects like medical devices or toys, must undergo third-party pre-deployment conformity assessments under EU product safety law.<sup>72</sup> Another category involves “stand-alone” high-risk systems, which are software-based AI systems used in contexts like employment, law enforcement, and education. *Two options exist for providers of such systems: 1) a pre-deployment, self-assessed conformity assessment based on the company’s internal controls, or 2) a pre-deployment conformity assessment conducted with the involvement of a government body.*<sup>73</sup>

After completing the conformity assessment process verifying that the AI Act’s obligations have been met, the system receives a CE mark (an obligatory compliance marking) allowing the tool to be placed on the EU market.<sup>74</sup> Providers must also comply with post-market monitoring obligations analyzing the performance of high-risk AI tools and their compliance with the law throughout their lifecycle.<sup>75</sup>



## [Questions raised by the conformity assessment model.](#)

The proposed AI Act is made more efficient by adopting a risk-based approach. Unfortunately, compliance with many of its requirements for conformity – for example, that AI systems are “accurate,” and that data sets are “representative” – will require standards to enable compliance. The European Commission has only recently requested their development by European standards organizations, an approach to standards policy that is not replicable in the United States.<sup>76</sup>

---

<sup>72</sup> Jakob Mökander, Maria Axente, Federico Casolari and Luciano Floridi, *Conformity Assessments and Post-market Monitoring: A Guide to the Role of Auditing in the Proposed European AI Regulation*, *Minds & Machines* 32, 241-268 (2022), [link.springer.com/article/10.1007/s11023-021-09577-4#Sec4](https://link.springer.com/article/10.1007/s11023-021-09577-4#Sec4).

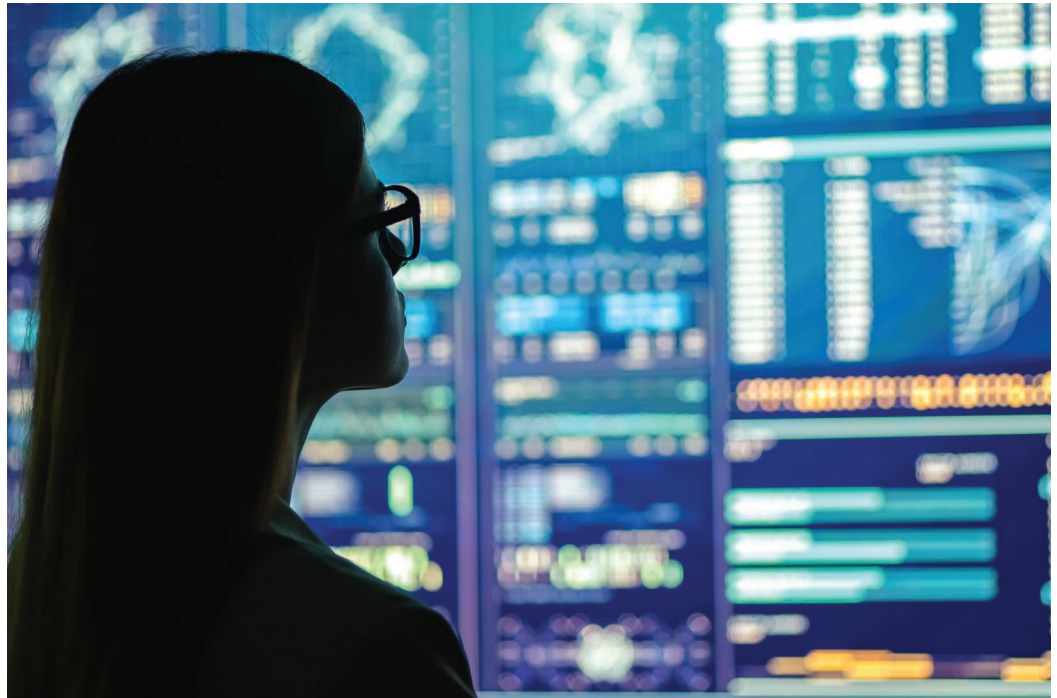
<sup>73</sup> Lilian Edwards, *Expert explainer: The EU AI Act Proposal*, Ada Lovelace Inst. (Apr. 2022), <https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/>.

<sup>74</sup> *CE Marking*, Eur. Union, [https://europa.eu/youreurope/business/product-requirements/labels-markings/ce-marking/index\\_en.htm](https://europa.eu/youreurope/business/product-requirements/labels-markings/ce-marking/index_en.htm) (last visited 23 Aug. 2022).

<sup>75</sup> See Mökander et al., *supra* note 72.

<sup>76</sup> *Artificial Intelligence*, CEN/CENELEC, <https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/> (outlining the new joint technical committee 21 ‘Artificial Intelligence’ and why standardization is needed) (last visited Aug 2, 2022).





In addition, the AI Act would require that new assessments must be completed whenever high-risk AI systems are subject to “substantial modification.” Yet the Act does not establish a clear threshold – e.g., whether changes in end-users, routine software updates, or even a machine learning model’s continuous changes would warrant a new assessment.<sup>77</sup> These ambiguities are expected to persist into the near future.

---

<sup>77</sup> Katerina Demetzou, *Introduction to the Conformity Assessment Under the Draft EU AI Act*, Future of Privacy Forum (Aug. 12, 2022), <https://fpf.org/blog/introduction-to-the-conformity-assessment-under-the-draft-eu-ai-act-and-how-it-compares-to-dpias/>.

# Recommendations

AIAs have three important advantages over other AI accountability tools: (1) they are *familiar* to organizations already conducting impact assessments for privacy and data protection; (2) they are *practical* because they do not rely on technical standards, which are currently nascent; (3) they are *future-proof* because they can adapt as AI systems and AI governance evolve.

In the near term, we recommend the following:

- **Policymakers should encourage algorithmic impact assessments.** Governments, industry organizations, and companies have taken the first steps, with some governments already adopting AIAs as the preferred AI accountability tool. Meanwhile, governance frameworks, including NIST's AI Risk Management Framework, explicitly incorporate impact assessments or enable compatibility with them.
- **Policymakers should support harmonization.** With policymakers at various levels of government and around the globe looking to develop regulatory frameworks for AI technology, harmonization is essential. Policymakers should avoid creating a patchwork of inconsistent requirements by looking to best practices in the field and engaging in robust international cooperation.
- **Policymakers should support standards work related to AI systems.** Standards development is typically the culmination of years of consensus-building among experts from academia, the private sector, government, and civil society. Because standards development can be a lengthy process, policy support should include securing funding and resources for standards development, bolstering knowledge-sharing, and promoting strategic engagement with international bodies.
- **Policymakers may wish to consider regulatory backstops to undergird companies' self-conducted algorithmic impact assessments.** This could help promote oversight, accountability, and public trust, while retaining the adaptability and technology-appropriate advantages of AIAs. However, any regulatory requirements should be carefully crafted to focus on ensuring that internal controls and processes, based on industry standards, are working effectively, rather than imposing retroactive audit requirements that ultimately do not yield more benefits than their costs.

# About The Authors



## Dileep Srihari, Senior Policy Counsel

Dileep Srihari is Senior Policy Counsel in Access Partnership's Washington, DC office. He focuses on data policy, cybersecurity, and telecommunications infrastructure issues. Dileep joined Access Partnership from the Computing Technology Industry Association (CompTIA), where he engaged in policy development and advocacy before Congress and federal agencies including the Federal Communications Commission (FCC). He also spent eight years at the Telecommunications Industry Association (TIA) where he worked with ICT vendors on spectrum, broadband infrastructure deployment, and supply chain security policy.

Dileep was earlier an attorney at WilmerHale with a practice that included appellate litigation and regulatory advocacy on topics including wireless interference protections, television program access, and process safety management. He previously worked on Capitol Hill for Senator Hillary Rodham Clinton of New York. Dileep holds an undergraduate degree in computer science and electrical engineering from Cornell University and a law degree from the Georgetown University Law Center.



## Meghan Chilappa, Policy Counsel

Meghan Chilappa is Policy Counsel in Access Partnership's DC office. She focuses on global policy and regulatory issues encompassing AI, cybersecurity, government access to data, and digital governance. Prior to Access Partnership, she was a Privacy Consultant at Deloitte. In law school, she held positions at the U.S. Department of Defense and the U.S. Department of Justice. She is published in Slate about the long-term impacts of internet shutdowns and served as a 2020 LENS (Law, Ethics, and National Security) Scholar.

Prior to law school, Meghan worked in public diplomacy and politics in Washington, DC. At Meridian International Center, she oversaw international exchange programs with the U.S. State Department. She led Foreign Policy for America's Political Action Committee (PAC) for the 2018 midterm elections in the U.S. She holds a BA in International Relations and Spanish from Syracuse University and a law degree from American University Washington College of Law.



Founded in 1999, Access Partnership shapes policy on behalf of the world's leading technology companies introducing fairness and stability for services and products entering new markets. We create policy, regulatory and legal routes to markets being adopted worldwide, remaining fair to all parties.

**Access Partnership**

1730 Rhode Island Ave. NW  
Suite 512  
Washington, DC 20036 USA

t: +1 202 503 1570

e: [washingtondc@accesspartnership.com](mailto:washingtondc@accesspartnership.com)