SUBMITTED: June 12, 2023

CIPL Response to NTIA Request for Comment on AI Accountability Policy

Docket Number NTIA–2023–0005
Submitted June 12, 2023

I. EXECUTIVE SUMMARY

The Centre for Information Policy Leadership (CIPL)[1] welcomes the opportunity to respond to the request for comments by the National Telecommunications and Information Administration (NTIA) on Artificial Intelligence (AI) system accountability measures and policies.[2]

CIPL has been on the forefront of promoting organizational accountability and a risk-based approach as cornerstones of effective data protection law, policy, and oversight for more than 20 years. Organizational accountability principles together with a risk-based approach play an increasingly important role in all areas of digital policy, law, and regulation, especially regarding the use of artificial intelligence (AI), because:

- They are critical to building and delivering trust in the modern digital age;
- They deliver a future-proof and outcomes-based approach to regulation that enables private and public sector organizations to develop and adopt new technologies and harness their benefits in a responsible and human-centric way;
- They enable organizations to translate and implement legal and policy requirements into operational controls and best practices; and
- They enable organizations to demonstrate compliance with accountability requirements internally (to management and boards) and externally (to customers, regulators, and wider public).

CIPL supports federal data privacy legislation in the United States and considers the passage of such a law to be a foundational layer to effective and responsible AI regulation. This recognition is informed by CIPL's research and practical experiences at the intersection of AI, data protection, and organizational accountability. Robust frameworks for organizational accountability and a risk-based approach to laws and regulations (with respect to enforcement and compliance) will enable the achievement of diverse policy goals, including protecting individuals from harms and fostering the beneficial use of technology and innovation.

[1] CIPL is a global privacy and data policy think tank in the law firm of Hunton Andrews Kurth LLP and is financially supported by the law firm and 85+ member companies that are leaders in key sectors of the global economy. CIPL's mission is to engage in thought leadership and develop best practices that ensure both effective privacy protections and the responsible use of personal information in the modern information age. CIPL's work facilitates constructive engagement between business leaders, privacy and security professionals, regulators, and policymakers around the world. For more information, please see CIPL's website at http://www.informationpolicycentre.com/. Nothing in this submission should be construed as representing the views of any individual CIPL member company or of the law firm of Hunton Andrews Kurth.

[2] AI Accountability Policy Request for Comment, 88 FR 22433, April 13, 2023, available at https://www.federalregister.gov/documents/2023/04/13/2023-07776/ai-accountability-policy-request-for-comment.

II. THE CIPL ACCOUNTABILITY FRAMEWORK

CIPL has been a proponent of organizational accountability for responsible data use and has researched and written extensively on this topic.[3]

A central element of CIPL's work on organizational accountability is the CIPL Accountability Framework. The core elements in CIPL's Accountability Framework are: leadership and oversight; risk assessment; policies and procedures (including fairness and ethics); transparency; training and awareness; monitoring and verification; and response and enforcement. (Figure 1). By encouraging businesses to implement comprehensive privacy and data governance programs based on CIPL's Accountability Framework or similar frameworks, CIPL has sought to ensure that businesses have processes in place that help them comply with applicable legal requirements and good practices and that enable them to demonstrate accountable data use practices.

Figure 1: CIPL Accountability Framework – Universal Elements of Accountability

The CIPL Accountability Framework has been used to promote organizational accountability in the context of building, implementing, and demonstrating comprehensive privacy programs.[4] This framework can also be used to help organizations develop, deploy, and organize robust and comprehensive governance and compliance programs in the AI context and demonstrate

[3] See CIPL Organizational Accountability Project, available at https://www.informationpolicycentre.com/organizational-accountability.html.

[4] See CIPL Report, *What Good and Effective Data Privacy Accountability Looks Like: Mapping Organisations' Practices to the CIPL Accountability Framework*, May 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_accountability_mapping_report_27_may_2020__v2.0.pdf.

accountability in AI.[5] Beyond the CIPL Accountability Framework, a wide range of tools are available for organizations looking to improve processes around AI development and deployment.

Globally, accountability is recognized as a key building block for effective regulation and ensuring corporate compliance. It has become a central focus of many regulatory regimes, including data protection and privacy, anti-corruption, anti-money laundering, white collar crime and corporate fraud, export controls and sanctions, and healthcare. Equally, in the broader digital regulation and AI space, accountability obligations can enable responsible AI innovation, promote trust in the AI ecosystem and facilitate the responsible collection and use of data for AI training, development, and deployment.[6]

CIPL believes that any AI regime should be based on accountability as a bedrock principle and should proactively encourage and incentivize accountability. Accountability requires organizations to operationalize and translate principles- and outcomes-based rules through appropriate and demonstrable policies, procedures, controls and governance to deliver compliance. Regulators should have the ability to provide relevant guidance for specific sectors or applications as needed to assist organizations in operationalizing principles- and outcomes-based rules.

Since CIPL started its work on accountable AI in 2018, we have seen a growing trend among many responsible organizations have been building coherent and comprehensive accountable AI frameworks, with an enhanced focus on data stewardship and organizational accountability considering the specific challenges of continuously evolving AI technology.

In 2023, CIPL launched a project to research organizations' experiences using CIPL and other accountability frameworks to guide their AI accountability programs, to collect best practices in order to promote their wider adoption, and to support future co-regulatory mechanisms. CIPL intends to publish this research later in 2023.

III. SPECIFIC QUESTIONS

1. *What is the purpose of AI accountability mechanisms such as certifications, audits, and assessments? Responses could address* the *following:*

Accountability mechanisms such as certifications, audits, and assessments are essential in digital policy and regulation. They allow organizations to demonstrate to their board, customers, the public, and regulators that a product or service meets specific criteria and is trustworthy. They also enable organizations to implement principle- and outcome-based legal requirements into measurable and demonstrable concrete steps and controls. This ensures effective regulation and compliance in practice, as opposed to laws on books and in theory. There is also evidence that certifications and assurance models play an important role in providing legal certainty and business confidence, including in business-to-business contexts.

---

[5] See CIPL Report, *Artificial Intelligence and Data Protection: Delivering Sustainable AI Accountability in Practice Second Report: Hard Issues and Practical Solutions*, February 2020, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_second_report_-_artificial_intelligence_and_data_protection_-_hard_issues_and_practical_solutions__27_february_2020_.pdf.

[6] See CIPL White Paper, *Top Ten Recommendations for Regulating AI in Brazil*, October 4, 2022, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/%5Ben%5D_cipls_top_ten_recommendations_for_regulating_ai_in_brazil__4_october_2022_.pdf.

Stakeholders continue to have important discussions about the extent to which certifications and audits should be voluntary versus required. Apart from the additional costs in terms of time and money, some organizations may be reticent to disclose model details, assessments, and training data to their auditors or third-party certifiers absent legal requirements, and there may be valid reasons for organizations to wish to limit such disclosures in some circumstances (e.g., for privacy and intellectual property reasons, or malicious actors gaining access to the information and using information to circumvent the intended purpose of the AI). However, measures such as nondisclosure agreements may be helpful for addressing such concerns. Discussions about the approach to voluntary vs. mandatory certifications and audits as well has how to ensure the protection of trade secrets and intellectual property should continue before firm rules are set and should consider appropriate risk-based parameters to limit unnecessary audit and certification mandates. It may be advisable to tread carefully initially and evolve any transparency requirements over time, once the market has had time to understand and implement the rules and it becomes more clear how they work in practice.

Certifications—Certification schemes and codes of conduct involve the use of independent third-party certifiers or monitoring bodies, as well as dispute resolution providers that are associated with such schemes. These entities can play important front-line enforcement and oversight roles and remediate many issues before a regulator needs to step in, provided that protections are put in place to honor intellectual property and trade secret concerns. These entities review and certify organizations' compliance and accountability programs and ensure that they comply with the relevant standard to which they were certified. When necessary, they can suspend certifications and take other remedial actions against non-compliant organizations. The dispute resolution functions of these schemes relieve regulators from the burden of dealing with large numbers of "easy" cases, allowing them to focus their enforcement attention on more important and strategic matters.[7] Certifications can be especially useful as a source of guideposts for SMEs seeking to implement comprehensive and consistent accountability programs.

Audits—An audit of an AI system should include two components: 1) a concrete measurement and 2) organizational accountability. Concrete measurement considers whether the performance or behavior of the AI model aligns with the stated expectation of the model. Organizational accountability considers the AI model within the context of its development or deployment and includes an evaluation of the effectiveness of the organization's comprehensive accountability framework as it is applied to the development or deployment of the AI. Monitoring and verification is an important element of the CIPL accountability wheel and enables organizations to verify their policies and procedures both internally and externally. Audits can be conducted by independent, third-party entities or by a second-party contracted by the organization to audit the organization.

Assessments— Organizations conduct internal risk assessments to identify potential risks and harms associated with the use of the AI and to help identify the necessary and appropriate mitigation measures. Such assessments are also a key component for organizational accountability, documenting characteristics of AI models for future review by independent auditors, or regulatory bodies. Internal assessments can be conducted by employees of the organization developing or deploying AI or by a second-party contracted by the organization to assess AI development or deployment.[8]

[7] See *id.*
[8] See Article 35 GDPR, Data protection impact assessment; see also European Union AI Act proposal, requiring high-risk AI systems to undergo an approved conformity assessment.

a. What kinds of topics should AI accountability mechanisms cover? How should they be scoped?

To ensure accountability within the broader ecosystem, it is vital that AI accountability mechanisms cover the entire life cycle from design and development to application and use and include all the core elements of organizational accountability discussed above in Section II. Important topics to consider within the entire life cycle include the sources and accuracy of data for AI models, the integrity and reliability of AI models, and downstream use and liability.

AI accountability mechanisms should address the full range of potential harms, including but not limited to:

- Harms to individuals, such as those related to erosion of privacy, unfairness, bias, discrimination, or violation of intellectual property.
- Harms that may be both individual and societal, including threats such as misinformation and disinformation, cybersecurity risks, and risks to critical infrastructure.

Accountability mechanisms should ensure that AI-based products or services abide by all relevant and applicable laws and regulations, including Section 5 of the Federal Trade Commission Act; housing and employment regulations; and intellectual property laws.

Absent a structured organizational accountability framework, development teams risk conducting assessments using ad hoc procedures that may result in insufficient care in and documentation of decision-making.

As we discuss further below, approaches such as datasheets for datasets, model cards for AI models, and system cards for AI systems can provide researchers, auditors, regulators, and AI users or customers summaries of key information about how an AI model was developed and how it is expected to perform. Absent clear standards for such documentation efforts, organizations may take inconsistent approaches that result in the omission of key information.

b. What are assessments or internal audits most useful for? What are external assessments or audits most useful for?

Internal audits/assessments— Internal assessments or audits are most useful for identifying, classifying, and scrutinizing risk and creating mitigations where necessary and possible. They also help organizations achieve compliance with legal requirements and meet internal organizational quality control and compliance goals; this is an integral element of CIPL's accountability wheel (monitoring and verification, see Figure 1). Internal assessments are attractive because they are agile and organizations do not have to be concerned with disclosing intellectual property or trade secrets. In organizations with robust accountability policies in place, internal risk assessments and audits will help the organization determine whether to move to market, verify that internal policies are working, and adjust where necessary. Importantly, internal assessments and audits provide an organization with the opportunity to document decisions that may later become part of liability litigation or regulator investigations. Documenting decisions and other metrics is especially important to AI accountability.

Organizations can demonstrate their commitment to responsible and accountable AI by publishing assessments externally in a manner that addresses the organization's intellectual property and trade secret concerns. Ideally, publishing internal assessments publicly in some form can cultivate trust among customers, clients, and other stakeholders.

External audits— External, third-party audits can play an important role in any accountability framework, including AI accountability. Ideally, external audits are conducted by certified entities with a duty to protect public interests and ensure that legal criteria are met.[9] Additionally, public researchers often conduct third-party, external audits to better understand the impact of products and services on certain groups.[10] Again, these audits can help increase trust by demonstrating that an AI application possesses necessary characteristics. Additionally, external audits can provide more robust and neutral views on particularly difficult ethical and compliance issues and may therefore be viewed as more credible by the public. External audit requirements should be designed through consultation with stakeholders and updated regularly based on technological developments and new practices.

There are specific challenges associated with external audits of AI models and systems, for compliance with non-binding trustworthy AI goals as well as laws. These include:

- Ensuring that deployers provide transparency and notice when deploying AI systems;
- Accessing the data that models are trained on; and
- Access to pre-deployment internal assessments.

In addition, before an AI model is tested for bias, it should be tested for functionality. However, external auditors and researchers often face a "black box" problem and cannot recreate the models to test for functionality because they lack access to the actual datasets that were used to train the model.

Whether assessments are mandatory or voluntary, organizations should be given flexibility in how they conduct them so long as they meet certain standards and are producible upon request by regulators.

c. An audit or assessment may be used to verify a claim, verify compliance with legal standards, or assure compliance with non-binding trustworthy AI goals. Do these differences impact how audits or assessments are structured, credentialed, or communicated?

Assessments and audits can be useful tools for verifying claims, compliance with legal requirements, adherence to non-binding trustworthy AI goals or standard. Fundamentally, any audit and assessment must be sufficiently rigorous and thorough to enable a reliable representation to regulators or the public that the standards the review purports to have verified have in fact been met by the organization, regardless of whether the standards are legally required or voluntary. In other words, the nature of the standard or claim to be verified should not materially change the way in which an audit or assessment of that standard is "structured, credentialed, or communicated." Of course, the level of risk associated with an AI application or model should inform whether external audits might be made mandatory as well as the level of detail of the audit or assessment, and the audience for the verification may inform the way in which it should be communicated.

[9] See, for example, 12 CFR 363.2(a), requiring independent audits by public accountants of insured depository institution's financial statements.
[10] See, for example, Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 2018, available at http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

d. Should AI audits or assessments be folded into other accountability mechanisms that focus on such goals as human rights, privacy protection, security, and diversity, equity, inclusion, and access? Are there benchmarks for these other accountability mechanisms that should inform AI accountability measures?

The risk associated with developing or deploying an AI product or service is extremely dependent on the context of the AI development and deployment. Regulations should provide organizations flexibility regarding whether to fold AI audits and assessments into broader assessments where appropriate, while not requiring this approach in all instances. As noted, an effective organizational accountability framework can be applied to any area of law and subject matter, and, as a general matter, it would be appropriate for organizations to apply their accountability frameworks broadly and to make their audits or assessments as comprehensive as appropriate under the circumstances, covering all relevant compliance and ethics issues facing them.

There is an increasing trend among organizations to leverage security and privacy impact assessments for AI and even human rights assessments. Certainly, there are benefits in using the same methodology and not burdening teams with performing several assessments in parallel. However, there is no consensus on how to identify and assess human rights risks and harms and how to do this in an integrated way for all disciplines—AI, privacy security, safety, children's protection, etc.

CIPL has mapped examples of AI accountability measures undertaken by organizations from different sectors, geographies, and sizes based on the CIPL Accountability Framework and included this table in Annex A of this response.

e. Can AI accountability practices have meaningful impact in the absence of legal standards and enforceable risk thresholds? What is the role for courts, legislatures, and rulemaking bodies?

Self-regulatory AI accountability practices can be quite impactful, particularly where organizations work together to establish realistic standards and procedures to demonstrate compliance with those standards. Organizations such as the International Organization for Standardization (ISO),[11] the Institute of Electrical and Electronics Engineers (IEEE), Partnership on AI, and the International Association of Privacy Professionals (IAPP) are all undertaking efforts to advance artificial intelligence standards and certifications. Together, these efforts demonstrate the breadth of support among stakeholders for the development of standards for AI accountability best practices. Indeed, organizations are already self-imposing ethical AI frameworks, which include multi-disciplinary teams overseeing AI development. AI data scientists and technologists are also developing and exchanging best practices in a thriving academic and commercial field of AI research and development (R&D). These efforts should be encouraged and leveraged for the development of bottom-up standards, as opposed to only having government-imposed, top-down rules and legislation. Non-governmental entities can move more quickly than national governments and can therefore adapt their standards to changing technologies. Industry-driven certifications also encourage accountability because they make it easier for organizations to select a technology partner that has been independently vetted and has successfully demonstrated AI accountability practices. Having said that, high-level and principles-based legislation coupled with regulatory guidance on implementation can provide important guardrails for industry within which to develop, deploy, and use AI technologies.

[11] ISO/IEC 42001 AI Management System Standard is currently under development. More information is available at, https://www.iso.org/standard/81230.html (last accessed June 9, 2023). .

The role of courts, legislatures, and rulemaking bodies—An important element of AI accountability within any organization is the risk assessment phase. To encourage meaningful impacts of these assessments, organizations would benefit from guidance on risk taxonomy from regulators and relevant rulemaking bodies.

Additionally, any illustrations of high-risk AI applications provided in regulatory guidance should be treated as rebuttable presumptions. This would enable organizations to take account of the highly contextual nature of AI applications and give them the opportunity to demonstrate that the use of an AI application in a specific context does not present a high risk. For example, using AI to assess diabetic retinopathy, as part of a triage process for initial screenings that reduces the risk of high priority patients having to wait weeks for ophthalmologists to review imagery, may not necessarily involve a high-risk, as the AI application is intended to perform a triage and not to provide a final diagnosis. Conversely, relying solely on AI to diagnose diabetic retinopathy and instigate treatment, without any additional medical review, may be high-risk AI.

2. ***Is the value of certifications, audits, and assessments mostly to promote trust for external stakeholders or is it to change internal processes? How might the answer influence policy design?***

Certifications, audits, and assessments do both—they help promote trust in external stakeholders *and* drive changes to internal accountability controls, procedures, and culture. For example, accredited certification programs should overlap with and address all key organizational accountability elements (see discussion of key elements of accountability in Section II), which in turn will incentivize organizations to continuously do better to maintain certification. The Better Business Bureau's accreditation program requires a participating business to provide the BBB with otherwise non-public information to certify compliance with its standards and cooperate with BBB to improve internal practices and procedures where necessary.[12] Verifying AI accountability requirements via third-party or self-certification schemes can ease the burden placed on regulatory bodies and promote verifiable and demonstrable internal accountability practices.

Given the rich and active development of and collaboration on best practices among data scientists in the commercial and academic AI research space, it is important that key lessons and takeaways are fed back into the policy and law making process. For example, for some years, AI and data technologists have been raising the issue of needing to use and retain sensitive personal data, to ensure appropriate AI training and prevent bias and discrimination. These AI requirements were in direct tension with data protection legal requirements that sought to limit the use of sensitive personal data, or subjected it to explicit consent. Significantly, these calls from the research community have resulted in policy changes in the EU and Dubai. For example, the proposed EU AI Act now allows for use of sensitive personal data for AI and algorithmic training.[13]

[12] See BBB Accreditation Standards, available at https://www.bbb.org/bbb-accreditation-standards/ (last accessed June 7, 2023).
[13] See Article 10(5), AI Act (Version 1.1).

3. AI accountability measures have been proposed in connection with many different goals, including those listed below. To what extent are there trade-offs among these goals? To what extent can these inquiries be conducted by a single team or instrument?

Due to the lack of standards to address AI considerations holistically, it is important that policymakers work collaboratively with a range of experts to ensure an appropriately balanced approach in connection with different policy goals. Many AI systems at issue are complex at a foundational level, and important discussions continue about how to provide meaningful transparency and explainability about these complex systems while preserving privacy and trade secrets. In addition, organizations may have to make trade-offs between considerations such as transparency of the AI system and security or accuracy. It is important for organizations to document and evidence their decision-making with respect to trade-offs. Also, some of the leading companies developing AI models and systems are using a range of emerging approaches to document machine learning systems, such as "model cards" and "data set nutrition labels". Although, standards for such documentation are yet to emerge.

Within an organization, AI development and deployment, depending on the context, can benefit from a multidisciplinary workforce tasked with overseeing the accountable development, deployment, and monitoring of an AI system. The 2023 IAPP-FTI Consulting Privacy and AI Governance Report found that more than 50% of respondents are building AI governance programs within existing privacy programs.[14] This may be because many AI accountability principles, such as transparency, fairness, security, and accountability, also exist in comprehensive privacy programs. Additionally, robust privacy programs often engage various teams within an organization, so leaders may find privacy professionals to be well-suited to execute AI accountability goals. A recent CIPL member roundtable explored the changing role of the Chief Privacy Officer within the context of AI adoption.[15] This discussion highlighted that responsible AI development and deployment is an organization-wide responsibility and cannot be tasked to a single team. Rather, it must be distributed across a broader cross-functional workstream with all relevant stakeholders involved, including privacy, compliance, ethics, intellectual property, security, legal, and more.

a. The AI system does not substantially contribute to harmful discrimination against people.

While bias and discrimination are key concerns surrounding AI models and systems, it is important to recall that humans themselves are not consistently rational, unbiased or always capable of explaining why they reach certain decisions. AI has the potential to mitigate some of the biases affecting human decision-making, and any national AI strategy should recognize this and aspire for AI to do so. At the same time, there are legitimate concerns over placing the outcome of certain decisions solely in the hands of an automated system which may produce a discriminatory result, for instance due to unidentified bias in the underlying data.

There are many measures that can be taken today to address questions related to such bias and discrimination:

---

[14] See IAPP & FTI, *Privacy and AI Governance Report*, January 2023, 3, available at https://iapp.org/resources/article/ai-governance-report/.

[15] See CIPL Roundtable, *Quo Vadis, CPO? An Evolving Role in Changing Times*, June 2, 2023, available at https://www.linkedin.com/pulse/quo-vadis-cpo-evolving-role-changing%3FtrackingId=HPsbBeaCQEqtsrSDovvGsA%253D%253D/?trackingId=HPsbBeaCQEqtsrSDovvGsA%3D%3D.

- **Facilitate access to data, including sensitive personal data:** AI technologists have confirmed that, to avoid bias, AI systems must be tested by reference to potentially sensitive categories of data, such as gender, race and health. Denying access to or preventing the retention or use of such potentially sensitive data makes it more difficult to detect, remedy, and prevent bias and may further limit the ability to explain why the AI application is arriving at discriminatory conclusions. It is important to note, however, that where sensitive data is processed to prevent bias from occurring, appropriate protections, including masking, anonymization and pseudonymization, and accountability safeguards will be of increased importance to mitigate potential privacy harms. There is also a need to better understand and mitigate against "proxy bias" in AI systems, by which a "facially neutral practice . . . disproportionately harms members of a protected class."[16]

- **Develop techniques to identify and address the risk of bias:** Many organizations today are developing techniques to specifically address the issue of discrimination in AI applications. For example, some organizations rely on counterfactual fairness testing. This technique checks for fairness in outcomes by determining whether the same result is achieved when a specific variable, such as race or gender, changes.[17] A 2019 report on "Perspectives on Issues in AI Governance" released by Google details other algorithmic fairness techniques designed to "surface bias, analyze data sets, and test and understand complex models in order to help make AI systems more fair".[18] These include Facets,[19] the What-If Tool,[20] Model and Data Cards[21] and training with algorithmic fairness constraints. Accenture has developed its own tool to "[assess] fairness and actions needed to mitigate bias", which it uses both internally with respect to its own AI projects, as well as externally, on client projects involving the deployment of AI applications to help clients address the fairness standard.[22] IBM has also created several tools to address issues of ethics and fairness in AI, including AI Fairness 360, a comprehensive open-source toolkit of metrics to check for unwanted bias in datasets and

[16] See Anya E.R. Prince & Daniel Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data*, 105 Iowa L. Rev. 1257 (2020), available at https://ilr.law.uiowa.edu/print/volume-105-issue-3/proxy-discrimination-in-the-age-of-artificial-intelligence-and-big-data#:~:text=AIs%20armed%20with%20big%20data%20are%20inherently%20structured,directly%20by%20non-suspect%20data%20available%20to%20the%20AI.

[17] See Singapore Personal Data Protection Commission, Model Artificial Intelligence Governance Framework, First Edition, January 2020, available at https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organization/AI/SGModelAIGovFramework2.pdf at page 46.

[18] To read more about these techniques, please see Perspectives on Issues in AI Governance, Google, 2019, available at https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf.

[19] See Google Research, Facets, available at https://pair-code.github.io/facets/ (last accessed June 9, 2023).

[20] See Google Research, What-If Tool, available at https://pair-code.github.io/what-if-tool/ (last accessed June 9, 2023).

[21] See Google Research Blog, Mahima Pushkarna & Andrew Zaldivar, *The Data Cards Playbook: A Toolkit for Transparency in Dataset Documentation*, November 17, 2022, available at https://ai.googleblog.com/2022/11/the-data-cards-playbook-toolkit-for.html; see also Google Research Blog, Huanming Fang & Hui Miao, Introducing the Model Card Toolkit for Easier Model Transparency Reporting, July 29, 2020, available at https://ai.googleblog.com/2020/07/introducing-model-card-toolkit-for.html.

[22] See Accenture, Fairness you can bank on, available at https://www.accenture.com/us-en/case-studies/applied-intelligence/banking-aib (last accessed May 24, 2023).

machine learning models,[23] and state-of-the-art algorithms to mitigate such bias,[24] as well as IBM Watson OpenScale, a tool for tracking and measuring outcomes of AI to help intelligently detect and correct bias, as well as explain AI decisions.[25]

- **Data scientist training:** Organizations developing AI applications and tools have invested heavily in training their data scientists that are engineering these systems and more training is needed. Part of this training includes raising awareness about different sources and types of bias, instruction on how to avoid and address bias when developing or deploying AI, and how to detect and test AI models and systems for bias prior to deployment.
- **Ethics review processes:** Many organizations have established internal and/or external AI ethics committees, data review boards or similar bodies as additional oversight mechanisms to drive organizational accountability, foster responsible decision-making, and ensure that new data uses uphold corporate and societal values, including combating harmful discrimination.

b. The AI system does not substantially contribute to harmful misinformation, disinformation, and other forms of distortion and content-related harms.

[No response provided.]

c. The AI system protects privacy.

CIPL has published several research reports since 2018 exploring the tensions between AI and data protection principles and legislation.[26] CIPL's 2020 report on "Hard Issues and Practical Solutions" explores ways to navigate challenges that advancements in AI pose with respect to concepts central to numerous data protection laws around the world, such as fairness, transparency, purpose limitation, data minimization.[27]

Of course, it is well documented that there are not only tensions between data privacy principles and AI technology, but also between more granular principles and objectives of data protection and AI, such as between data privacy and accuracy, accuracy and fairness, privacy and fairness, privacy and accuracy, accuracy and explainability, and explainability and security.[28]

Our research found that the use of robust accountability frameworks by organizations developing and deploying AI is especially critical for helping organizations navigate these tensions while preserving AI's benefits and for regulators examining their practices. Our ongoing research on the operations of AI Ethics Councils points to the important role that these bodies can have in ensuring that AI accountability frameworks address privacy in concert with other policy goals, like reducing bias, preventing discrimination, and promoting safety. While accountability mechanisms do not eliminate

---

[23] IBM moved the AI Fairness 360 tool to non-profit technology consortium Linux Foundation in July 2020, available at https://ai-fairness-360.org/ (last accessed June 9, 2023).
[24] See IBM Policy Lab, Anjelica Dortch & Dr. Stacy Hobson, *Mitigating Bias in Artificial Intelligence*, May 2021, available at https://www.ibm.com/policy/wp-content/uploads/2021/05/AI_Bias_IBMPolicyLab.pdf.
[25] See IBM, *Watson OpenScale on Cloud Pak for Data*, available at https://www.ibm.com/docs/en/cloud-paks/cp-data/3.5.0?topic=services-watson-openscale (last updated Oct. 2022).
[26] Please find a list of these papers at https://www.informationpolicycentre.com/ai-project.html.
[27] See supra at note 5.
[28] See UK Information Commissioner's Office, *Our work on Artificial Intelligence*, available at https://ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/.

the need for trade-offs, overseeing such mechanisms through dedicated, cross-cutting AI ethics bodies enables decision-making that minimizes those trade-offs to the greatest extent possible.

d. The AI system is legal, safe, and effective.

Accountability requires organizations to be thoughtful about the risks and impacts of their processing activities on individuals and establish processes (such as risk assessments) and controls to anticipate and address these in compliance with the law or other standards. Typically, accountability is implemented through comprehensive compliance and management programs. The concept of organizational accountability has become a common feature of privacy and data protection regulation globally. It has also been deployed in many other compliance areas, such as anti-bribery, anti-money laundering, export control, medicine and food regulation.[29] It can also be deployed in the AI context. Accountability-based compliance and governance programs enable organizations to operationalize principles-based laws and standards into risk-based, verifiable, demonstrable and enforceable corporate practices and controls, supported by technology tools. This enables organizations to be responsible data stewards and developers, deployers and users of technology, including in the AI context, by assessing the potential impacts of a given application, implementing policies and procedures to ensure accountability and to continuously improve and adapt to change.

Organizational practices rooted in such accountability-based compliance programs benefit individuals by delivering real, relevant and effective protections based on legal requirements and ethical standards. They also help organizations demonstrate legal compliance to regulators, business partners and individuals. This results in increased trust by these constituencies in organizations' development, deployment and use of AI.

Finally, es explained above, the need to ensure the safety, security, and effectiveness of AI systems may also result in tensions with other rights and interests, such as data privacy. Accountability provides organizations with a roadmap to navigate these tensions and trade-offs in a way that is repeatable and consistent and provides as much legal certainty as possible.

e. There has been adequate transparency and explanation to affected people about the uses, capabilities, and limitations of the AI system.

Explainability falls under the broader concept of transparency in the context of AI. It is a way of providing transparency about the outcome of an AI decision, output, or process. As noted in the Singapore Personal Data Protection Commission's Model AI Governance Framework, "[a]n algorithm deployed in an AI solution is said to be explainable if how it functions and arrives at a particular prediction can be explained".[30]

Explainability can be realized in AI systems in many ways. For example:

---

[29] See CIPL, *The Concept of "Organizational Accountability" Existence in US Regulatory Compliance and its Relevance for a Federal Data Privacy Law*, July 2019, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_organizational_accountability_-_existence_in_us_regulatory_compliance_and_its_relevance_for_a_federal_data_privacy_law_3_july_2019_.pdf.

[30] See Singapore Personal Data Protection Commission, *Model Artificial Intelligence Governance Framework: Second Edition*, January 2020, available at https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf.

- **Implementing best practices to facilitate traceability:** Traceability requires documenting data inputs and other processes that produce the AI system's decision. Organizations can improve traceability, and in turn foster explainability, by creating and implementing best practices such as model cards and datasheets or codes of conduct regarding the collection, deployment and use of data.
- **Developing and employing explainability tools and techniques:** We mentioned counterfactual fairness testing in response to a previous question (3(a)) as it relates to preventing and detecting bias. This technique is also useful to facilitate explainability, as it checks for fairness in outcomes by determining whether the same result is achieved when a specific variable, such as race or gender, changes. If a different result is reached by an AI system when such a variable changes, then data scientists and AI engineers will be able to look more closely at why that is, not only to remedy the bias but also to be able to explain, at least in part, what is driving the algorithm to arrive at such a result.
- **Considering the human decision-making alternative:** In the offline world, humans are often unable to consistently explain their preferences for one option over another. While we may be able to subsequently ask for an explanation, this explanation at best will be logical, and almost certainly not technical or mathematical. Considering approaches to transparency in an offline world can provide perspective and inform the level and type of transparency to strive for when building AI systems.
- **Consider the audience and use case involved:** What explainability involves and looks like may differ depending on the audience involved. For example, an organization may need to explain an AI outcome to an individual who is directly and negatively impacted by a decision, a regulator in cases of investigation and enforcement or participation in regulatory AI sandboxes, business partners who are interested in utilizing the AI solution, or for purposes of internal explainability and transparency to an oversight board or senior leaders. These different audiences imply different types and requirements of explainability which should be fulfilled appropriately. Similarly, explainability may differ depending on the use case involved. For instance, full technical explainability in the context of fraud detection and prevention would not be appropriate as it could allow fraudsters to circumvent the AI solution for fraud detection and prevention. However, technical details could be provided to regulators, in the event of an investigation, and to individual-facing entities, so they are able to provide information to individuals. Still, AI transparency must be balanced with an organization's need to protect intellectual property.
- **Consider alternatives where explainability is not feasible:** The concept of explainability is often challenged by the "black box" phenomenon. As noted by the Norwegian DPA, "the black box makes it practically impossible to explain how information is correlated and weighted in a specific process".[31] The black box can create surprising or unanticipated results with invisible or unintelligible reasoning, even for developers. In addition, AI systems develop and change because of additional inputs, so decisions may not be easily repeatable. Where it is not possible to explain an AI outcome to individuals because of the black box problem (which can arise in systems that involve deep learning), other options to deliver meaningful information

[31] See Artificial Intelligence and Privacy, Datatilsynet (Norwegian Data Protection Authority), January 2018, available at https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf at page 19.

and empowerment of the individual should be considered, including human review of AI decisions where appropriate, redress mechanisms, and feedback tools.

- **Consider different transparency tools based on the context of the AI application:** Project ExplAIn, a collaborative project between the United Kingdom Information Commissioner's Office (ICO) and the Alan Turing Institute to create practical guidance for AI explainability, surveyed citizen juries and empirically demonstrated that individuals facing AI healthcare scenarios cared more about accuracy than transparency, while transparency expectations were heightened for the use of AI in job recruitment and criminal justice scenarios.[32] This suggests that transparency, and the tools used to achieve it, may differ based on what the AI application is used for, what the consequences are, and what rights individuals have going forward. To illustrate these different considerations for transparency, consider the use of facial recognition technologies by airlines to check boarding passes or by customs officials to allow individuals into a country. The decision made by the organization deploying AI in these cases is very significant, but transparency regarding the code itself is unlikely to be of concern to the impacted individual, particularly if a faulty decision can be reversed quickly. Instead, the concern in such cases may primarily be with how to quickly contest or correct the faulty decision.[33]

While explainability in AI can be achieved through many different avenues, it may not always be appropriate. Explainability can be in tension with accuracy—the more data and complex modelling an AI system uses in order to be as accurate as possible, the more difficult it may be to explain it. Also, it is important to recognize that disclosing too much information about an AI process may not only result in confusion and information overload for some individuals while helping others game the system but may also threaten commercial intellectual property interests by disclosing trade secrets. While respect for the rule of law and individual rights is of the utmost importance, this must be balanced with a company's ability to innovate and protect its intellectual property rights associated with its AI applications and inventions.

As evidenced by the above measures, there are many ways in which explainability can be implemented in AI systems. Risk-based flexibility, within appropriate standards or regulatory parameters, for organizations to decide which methods are most appropriate to implement explainability is crucial given that AI applications differ widely from one context to another. This approach should provide the necessary combination of flexibility as well as appropriate certainty for organizations to effectively implement explainability.

CIPL recommends that NTIA:

1) mention some of the various ways that explainability can be implemented in AI systems and encourage organizations to develop further ways to facilitate explainability and transparency as it relates to AI outcomes.

2) highlight the importance of innovation and robust competition and how disclosure of AI algorithms and decision-making processes must be balanced against commercial IP rights and business interests.

---

[32] See Information Commissioner's Office, *Explaining decisions made with AI*, available at https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/ (last accessed June 9, 2023).
[33] See supra at note 5 at page 14.

3) highlight that transparency requirements must be risk-based within appropriate standards or regulatory parameters. For example, low-risk AI applications such as those used for automated decisions that do not have legal effects or similarly significant effects may require lower levels of transparency than higher risk AI uses.

4) recognize that there are different levels of AI transparency, based on the intended audience. For example, transparency to auditors or regulators can be used as a compensation or a proxy for the lack of or inability to provide granular transparency to individuals.

5) recognise that transparency is contextual, and will evolve over time. Hence, transparency requirements should be meaningful, user-centric, and not prescriptive. It is all about the outcomes, as opposed to prescriptive compliance with a given transparency requirement.

f. There are adequate human alternatives, consideration, and fallbacks in place throughout the AI system lifecycle.

[No response provided.]

g. There has been adequate consultation with, and there are adequate means of contestation and redress for, individuals affected by AI system outputs.

Transparency, explainability, and redress are intrinsically linked to the assessment of fairness in an AI application. In other words, providing for user-centric and meaningful transparency and explainability of the AI decision-making process and enabling redress to individuals are likely to increase the chances that a specific data processing in AI is fair.

Redress is likely to assume new importance in the effective governance of AI, and therefore should warrant renewed attention. Even with the proper accountability-based controls and constraints designed to minimize or prevent AI risk, we will never achieve zero risks or harms. Thus, it is important, particularly in the context of automated decision-making with a legal or similarly significant impact, that individuals have an effective and efficient avenue for contesting outcomes and appealing decisions.[34] Individuals must understand how a decision has been made and must be able to contest the decision, request human review, and correction.[35] Many of the concerns around fairness to consumers can be addressed in large part by providing rapid and effective redress through organizational accountability. Redress allows individuals to contest and change an outcome they believe is inaccurate, unfair, or otherwise inappropriate.

When organizations are developing new technologies and considering the impact of those technologies, it is unlikely that they will foresee and limit every negative impact. In some cases, an organization may determine that the risks are too high to deploy the technology. However, the trade-offs in other contexts may warrant that the technology be deployed, but also that it must provide visible and effective avenues to correct situations where biased or incorrect decision-making occurs. Organizations should ensure that redress is meaningful—and that it does not merely become a rubber

[34] This sort of human review already exists in US law in certain contexts where automated processing informs decision-making. The Fair Credit Reporting Act and Equal Credit Opportunity Act both provide consumers with some right to explanation and to contest the decision, while Title VII of the Civil Rights Act as well as the Fair Housing Act provide individuals with a right to challenge decisions.

[35] See United Kingdom Data Protection and Digital Information (No.2) Bill, House of Commons, Session 2022-23 (see Articles 22A & 22B, substituting UK GDPPR prohibition on automated decision-making with rules that focus on individual rights and redress).

stamp on an automated decision. If unfairness or inaccuracy is uncovered, the accountability framework of organizations must have processes in place to adjust for this and limit similar situations in the future. Considering and developing these remedies and processes will be an essential part of deploying AI, and regulators evaluating the use of AI and impact on data protection should look for these visible avenues of redress as one way to demonstrate responsible implementation of AI technologies. [36]

h. There is adequate management within the entity deploying the AI system such that there are clear lines of responsibility and appropriate skillsets.

CIPL's accountability framework emphasizes the importance of leadership and oversight as a crucial element of organizational accountability. This is even more true considering the increasing use and adoption of AI in organizations, and the need to have clear oversight and supervision of the benefits and risks of AI development and use.

Also, CIPL's ongoing research on best practices in establishing organizational AI accountability points to the importance of organizations establishing internal or external AI ethics bodies that include individuals with a range of diverse life experiences, multidisciplinary skill sets, and, if the body is internal, different roles within the organization. It is equally important that clear procedures be spelled out for escalating decisions on developments or deployments to the ethics body, including cadences for ongoing review of research and business activities as well as emergency escalations for time-sensitive decisions.

4. Can AI accountability mechanisms effectively deal with systemic and/or collective risks of harm, for example, with respect to worker and workplace health and safety, the health and safety of marginalized communities, the democratic process, human autonomy, or emergent risks?

It is important for AI accountability frameworks to attempt to address such risks – to describe such potential harms with precision and to move organizations toward specific pathways to identify, avoid, minimize, or mitigate them. For example, the NIST AI Risk Management Framework identifies baskets of potential harms to specific social groups, societal harms such as the democratic system, and "harms to ecosystems", including the global financial system, supply chains, and the natural environment.[37] That said, measuring effectiveness of risk management across these areas, especially with respect to more diffuse challenges like harm to ecosystems-- is challenging due to the difficulty of defining measurable outcomes and appropriate timeframes within which to track them. Organizations may have other challenges when it comes to effectively measuring systemic risk, which may include the availability of diverse perspectives and skill sets within the organizations to reasonably assess all risks, including evolving risks.

[36] See supra at note 5 at page 32.
[37] NIST AI Risk Management Framework 1.0, available at https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf (last accessed June 9, 2023).

5. ***Given the likely integration of generative AI tools such as large language models (e.g., ChatGPT) or other general-purpose AI or foundational models into downstream products, how can AI accountability mechanisms inform people about how such tools are operating and/or whether the tools comply with standards for trustworthy AI?***

Transparency vis-à-vis individuals and the public must be achieved in a manner that is realistic and effective in practice. For example, transparency for generative AI tools should focus on the capabilities and challenges of the generative AI systems involved and not the detailed workings of the system. In addition to publishing system cards, generative AI developers can clearly communicate the limitations of the tools to users, and organizations that deploy these systems should be required to disclose their use. In the case of generative AI it is crucial to ensure that the end user is aware at any point in time that they are interacting with a generative tool, for instance, and that the output was generated by generative AI. Finally, it is essential that users of generative AI tools understand the limitations, features, and use-cases of generative AI.

6. ***The application of accountability measures (whether voluntary or regulatory) is more straightforward for some trustworthy AI goals than for others. With respect to which trustworthy AI goals are there existing requirements or standards? Are there any trustworthy AI goals that are not amenable to requirements or standards? How should accountability policies, whether governmental or non-governmental, treat these differences?***

[No response provided.]

7. ***Are there ways in which accountability mechanisms are unlikely to further, and might even frustrate, the development of trustworthy AI? Are there accountability mechanisms that unduly impact AI innovation and the competitiveness of U.S. developers?***

If properly designed and implemented, accountability frameworks will foster rather than frustrate the development of trustworthy AI. Accountability frameworks are only likely to hamper the development of trustworthy AI if they incorporate rigid and prescriptive compliance measures that are not risk-based or scalable and that focus on the technology itself, rather than the actual risk posed by the AI application or use. Rules must also be outcomes-based, directing the desired outcomes, as opposed to prescribing precisely how to achieve such outcomes. It is more likely that industry standards and certification schemes will help promote best practices in the development and use of trustworthy AI.

8. ***What are the best definitions of and relationships between AI accountability, assurance, assessments, audits, and other relevant terms?***

Assurance, assessments, audits, and similar concepts are all features of any mature organizational accountability program. CIPL published a white paper titled, "Organisational Accountability—Past, Present and Future" in October 2019, in which we expand on organizational accountability as a powerful tool for effective data regulation and innovation.[38]

[38] See CIPL, *Organisational Accountability – Past, Present, and Future,* October 2019, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_-_organisational_accountability_%E2%80%93_past_present_and_future__30_october_2019_.pdf.

Accountability is a mainstay of privacy and data protection regulation globally. CIPL also believes that accountability concepts are well-suited for responsible AI development and deployment. In the United States, the accountability concept can be traced back to the 1977 Foreign Corrupt Practices Act (FCPA) and the 2002 Sarbanes-Oxley Act (SOX Act). From 1987 it appeared in the United States' US Sentencing Commission Federal Sentencing Guidelines, and from 2019 it was adopted as part of the Department of Justice guidance for white-collar prosecutors. Key elements of accountability are well established in Anti-Money Laundering regulations and in regulatory guidance for various segments of the US healthcare industry including hospitals, nursing homes, third party billing services and medical equipment suppliers.[39]

In the business setting, accountability can be referred to as corporate responsibility, governance, stewardship, or duty and is associated with co-regulation and voluntary codes of practice. Regardless of these associations or the specific contexts in which it is implemented, an accountable organization is one that can demonstrate that it has effective internal processes in place to comply with its legal and regulatory obligations. Thus, accountability can be described as a framework that operationalizes and translates principles-based laws or standards into effective and enforceable internal policies, procedures, controls, and governance programs, with external guidance from regulators and advisers. This requires an organization to be thoughtful about risks to its business and the individuals it affects, to establish controls and incentives that drive responsible and ethical behaviour, and to demonstrate that this is the case. Accountability requires organizations to show that they are fully cognizant and in control of their impact on people and the environments in which they operate.

As detailed by the CIPL Accountability Framework (Figure 1), effective organizational accountability includes assurance, assessments, internal audits, and, in some cases, external audits. This Framework is applicable to all regulatory contexts, including data protection, privacy, and AI governance.

EXISTING RESOURCES AND MODELS

9. What AI accountability mechanisms are currently being used? Are the accountability frameworks of certain sectors, industries, or market participants especially mature as compared to others? Which industry, civil society, or governmental accountability instruments, guidelines, or policies are most appropriate for implementation and operationalization at scale in the United States? Who are the people currently doing AI accountability work?

Organizations that are developing, deploying, and/or using AI technologies have been working rapidly to put in place governance mechanisms for the responsible and ethical development, deployment, use and sale of AI technologies and associated data collection, sharing, and use. This effort often begins with organizations establishing core principles to guide the development and use of AI and other new technologies, consistent with their broader corporate or organizational principles. For example, IBM established its Principles for Trust and Transparency to guide company-wide approaches to AI in 2018.[40] Since then, IBM has also released open-source tools like AI Explainability 360 and AI Fairness 360.

Some organizations, including IBM, have gone further to translate such principles into practice through internal and/or external AI ethics councils that stand alone or as part of a multi-tiered

[39] See *id.*

[40] See IBM, *Principles for Trust and Transparency*, May 2018, available at https://www.ibm.com/policy/trust-principles/.

governance framework. Microsoft takes a multi-pronged approach: An Office of Responsible AI (ORA) articulates overall Responsible AI Principles and leads external engagement, an AI and Ethics in Engineering and Research (AETHER) Committee composed of representatives from across the company advises on specific decisions about AI development and deployment, and a Responsible AI Strategy in Engineering (RAISE) team focuses on building Microsoft's Responsible AI Principles and Responsible AI Standard into engineering practice.[41]

Currently, there is not a single global framework to guide responsible AI programs but a number of AI frameworks and standards have gained attention among some organizations, including the U.S. NIST AI Risk Management Framework, U.S. Blueprint for an AI Bill of Rights, UK Guidance on AI and Data Protection, Singapore PPDC's Model AI Governance Framework, and CIPL's Accountability Framework discussed elsewhere in this submission (see above and Annex A). The UK government established the Centre for Data Ethics and Innovation (CDEI) to address and produce useful tools for responsible AI; CDEI has been working on developing AI assurance tools.[42]

"Policy prototyping" has a useful role to play in advancing approaches to AI accountability. Policy prototypes are collaborative pilot projects that mobilize a coalition of public and private actors. These programs are also regulatory innovation labs intended to enable the development and testing of a policy idea in the field of new and emerging technologies, including AI. The policy idea can be inspired by a law that is being considered, a self-regulatory instrument, a code of conduct, a set of industry guidelines, etc. Policy prototyping programs are also empirical programs that provide evidence-based policy input to policymakers either to improve existing governance frameworks or to inform new ones. A successful example of policy prototyping is Meta's Open Loop project.[43] Open Loop projects have been deployed in Europe in the context of AI risk assessments and the envisioned policy approach of the proposed EU AI Act and in Singapore and Mexico on transparency and explainability. The EU project was very useful to start-ups in the AI space who valued the process of conducting an AI risk assessment along with the guidance that was prepared for them to conduct such an assessment.[44] Engaging in this exercise helped the start-ups to develop better AI applications early in the product development process.[45]

10. ***What are the best definitions of terms frequently used in accountability policies, such as fair, safe, effective, transparent, and trustworthy? Where can terms have the same meanings across sectors and jurisdictions? Where do terms necessarily have different meanings depending on the jurisdiction, sector, or use case?***

Fairness: Defining fairness has been an ongoing challenge both in the context of AI and elsewhere in privacy and data protection. The longstanding test for what is an "unfair" business practice employed

[41] See Putting Principles Into Practice at Microsoft, available at https://www.microsoft.com/en-us/ai/our-approach?activetab=pivot1:primaryr5 (last accessed June 2, 2023).
[42] See UK Government, Centre for Data Ethics and Innovation, *From principles to practice: Launching the Portfolio of AI Assurance techniques*, June 7, 2023, available at https://cdei.blog.gov.uk/2023/06/07/from-principles-to-practice-launching-the-portfolio-of-ai-assurance-techniques/.
[43] See Meta Open Loop, available at https://openloop.org/ (last accessed June 7, 2023).
[44] See Artificial Intelligence Act, Open Loop, Meta, available at https://openloop.org/programs/open-loop-eu-ai-act-program/ (last accessed June 7, 2023).
[45] See CIPL's Response to UK Department for Digital, Culture, Media and Sport (DCMS) Policy Paper on Establishing a Pro-innovation Approach to Regulating AI, September 2022, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_uk_dcms_proposed_approach_to_regulating_ai_23_09_22.pdf.

by the Federal Trade Commission is whether the practice causes a substantial injury that is not outweighed by any countervailing benefits to consumers or competition that the practice produces and that causes an injury that consumers themselves could not reasonably have avoided.[46] Under the GDPR, the processing of personal data must be fair (Art. 5 (1) (a)), meaning not in secrecy or through deceptive means. [47] In practice, fairness appears to be an amorphous concept that is subjective, contextual, and influenced by a variety of social, cultural, and legal factors. The same data used in different contexts may raise entirely different reactions to fairness questions. For example, if universities use prospective student data to train an algorithm that tailors advertising to "non-traditional prospects" such as first-generation university students, the assessment of fairness may be different than if the same data is used to identify students most able to pay for university and direct advertising toward those more well-resourced populations.[48]

The contextual nature of fairness creates significant challenges for regulators charged with interpreting and enforcing the law, for organizations charged with implementing it, and for individuals whose rights are to be protected by it. The difficulty and importance of defining and ensuring fairness are only magnified in AI contexts.

Fairness should be addressed from two dimensions: fair process (meaning processes that consider the impact on individuals' interests) and fair outcome (meaning the appropriate distribution of benefits). Both dimensions need to be addressed if we are to maximize the value of data and its applications for all those with an interest in it. Fairness is not absolute and may require continual and iterative reassessment in the relevant context. Equally important as these technical tools are the variety of procedural and accountability mechanisms to ensure fairness. Organizations can create internal governance structures and accountability frameworks, and then utilize tools such as AI data protection impact assessments (AI DPIAs) or data review boards to implement AI accountability. These mechanisms are particularly useful in the development phase of AI applications, but also in the review and monitoring phases. Of course, providing transparency and mechanisms for redress will be essential to ensuring fairness throughout the deployment of AI technologies. All of this exemplifies the point that fairness has to be ensured throughout the lifecycle of an AI application—from evaluation of the AI use case and input data, algorithmic modelling, development, and training to deployment, ongoing monitoring, verification, and oversight. [49]

Transparency: Like fairness, transparency is a concern exacerbated by the complexity of AI, but it is also a potential solution for many of the fears around AI technologies. The goals of transparency are to inform individuals and regulators about how AI systems are used to make decisions, hold organizations accountable for their practices, policies and procedures concerning AI, help detect and correct bias, and generally foster trust in the use and outcome of proliferating AI. Transparency has been a difficult challenge in AI, as it is often unclear what we mean by transparency. The challenge of transparency in AI is made more difficult due to the complex and changing nature of AI algorithms.

---

[46] See Federal Trade Commission, Policy Statement on Unfairness, Dec. 17, 1980, available at https://www.ftc.gov/legal-library/browse/ftc-policy-statement-unfairness.

[47] See Kuner, Christopher and Bygrave, Lee A. and Docksey, Christopher and Docksey, Christopher and Drechsler, Laura and Tosoni, Luca, *The EU General Data Protection Regulation: A Commentary/Update of Selected Articles*, May 4, 2021, available at https://ssrn.com/abstract=3839645.

[48] For background information, see Douglas MacMillan & Nick Anderson, *Student Tracking, Secret Scores: How College Admissions Offices Rank Prospects Before They Apply*, Washington Post, October 14, 2019, available at https://www.washingtonpost.com/business/2019/10/14/colleges-quietly-rank-prospective-students-based-their-personal-data/?fbclid=IwAR24p1HKEaHfN0K7kH4H5XBeDw4qgRib_v-o48afJ5bF5z10dIe9vCtiVac.

[49] See supra at note 5 at page 11.

One of AI's strengths is spotting complex patterns that had previously been missed,[50] but such complexity is inherently hard to explain in terms that are easily understood by most humans. Because these systems are complex and often changing, providing information about an algorithm may only partially fulfil the goals of transparency.

Transparency may differ depending on the audience it is geared toward—the individual or category of individuals impacted by the decision, the regulator, a business partner, or internal stakeholders like an oversight board or organizational leaders. Each of these audiences may require a different specific approach to transparency.

The level and method of transparency should ultimately be tied to the context and the purpose of AI applications. It is also clear that transparency is a broader concept in the context of AI—it includes explainability and understandability, as well as transparency concerning redress options and the ability to contest an AI decision. Finally, transparency also means the ability to articulate benefits of a particular AI technology and tangible benefits to individuals, as well as to broader society. Transparency in this form educates individuals and drive greater trust and acceptance of these new applications. [51]

11. What lessons can be learned from accountability processes and policies in cybersecurity, privacy, finance, or other areas?

CIPL published a white paper in 2019 titled, "The Concept of 'Organizational Accountability': Existence in US Regulatory Compliance and its Relevance for a Federal Data Privacy Law." This paper details how organizational accountability exists in many areas of US law and compliance, including anti-corruption, corporate fraud and white-collar crime, anti-money laundering, and healthcare.

Our research showed that the concept of organizational accountability is deeply engrained in the US legal system. Accountability's key features are also typically consistent across different regulatory areas and are in line with the essential elements of the CIPL Accountability Framework (Figure 1). C-suite leadership, ethics and compliance officers, and boards are familiar with accountability frameworks required under other laws and often apply the same frameworks to foster consistency in reporting across various regulatory areas.

In addition, CIPL analyzed two FTC consent decrees related to privacy and security violations.[52] We found that both decrees imposed significant and consistent accountability requirements aligned with the CIPL Accountability Framework. CIPL argued that the settlements provided a useful model for other organizations considering proactively how to strengthen accountability in their privacy and security programs.

---

[50] See, for example, National Institutes of Health, *Artificial Intelligence Accurately Predicts Protein Folding*, July 27, 2021, available at https://directorsblog.nih.gov/2021/07/27/artificial-intelligence-accurately-predicts-protein-folding/.

[51] See supra at note 5 at page 16.

[52] See CIPL Discussion Paper, *Organizational Accountability in Light of FTC Consent Orders*, November 13, 2019, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_-_organizational_accountability_in_light_of_ftc_consent_orders__13_november_2019_.pdf.

12. What aspects of the United States and global financial assurance systems provide useful and achievable models for AI accountability?

See above response.

In addition, the Financial Stability Board established after the 2008 financial crisis, could provide a useful model for the establishment of a global AI oversight body and mechanism. Such an international AI body could provide an appropriate fora for building consensus, convergence, and the development of model rules and standards that would apply globally. It is essential that there is a global dialogue and convergence in how countries respond to calls for regulation of AI technologies, given the global reach of technologies and companies developing and using these technologies.

13. What aspects of human rights and/or industry Environmental, Social, and Governance (ESG) assurance systems can and should be adopted for AI accountability?

As we note in our responses to Questions 1 and 4 above, it is important for AI Accountability frameworks to address risks of harm to individuals as well as systemic and collective risks of harm. Sound ESG assurance systems also address this range of individual and collective risks. AI Accountability expressly requires organizations to perform **contextual risk assessments** that examine whether a particular instance of development or deployment is likely to cause harm and/or impact the rights of individuals (such as anti-discrimination rights related to employment, housing, and credit opportunities), and broader, systemic outcomes (e.g., to the environment or to democratic systems), while also surfacing mitigation measures that enable legitimate uses.[53]

CIPL's forthcoming report on the adoption of a holistic data strategy will show how some companies have made data uses—including uses of data in the context of AI—a board-level issue, sometimes linked to ESG. Forward-thinking boards have come to recognize their dual obligations to shareholders and to society at large and the importance of situating data governance within their ESG frameworks.

The 'S' and 'G' of ESG are of particular importance with respect to data that may be processed within AI systems. Formalizing processes for complying with individuals' data rights and principles of ethical data use are central to addressing the **social** impact of data.[54] Moreover, **governance** is at the core of a holistic data strategy by requiring coordination and collaboration among traditionally siloed competencies, like legal, compliance, information security, finance, engineering, risk, audit, and ethics. Indeed, Responsible AI is an organization-wide responsibility.[55]

---

[53] See CIPL's Response to NTIA Privacy, Equity, and Civil Rights Request for Comment, Docket No. NTIA-2023-0001, submitted March 6, 2023, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_ntia_privacy_equity_and_civil_rights_request_for_comment__6_march_2023_.pdf.

[54] "Companies now have a social responsibility to be respectful of personal and behavioral data. They must weigh their reputation and investor benefits from prioritizing ESG against profits derived from third-party data collection and use. By positioning privacy as a social value, companies build a level of trust from society's expectation of privacy that had been lost. By being more scrupulous with data collection, consumers will feel comfortable sharing personal and sensitive information that will eventually build brand reputation and convert into investor-friendly profits." Sean Song, *Why Business Leaders Must Incorporate Data Privacy Into ESG Frameworks*, CPO Magazine, June 29, 2022, available at https://www.cpomagazine.com/data-privacy/why-business-leaders-must-incorporate-data-privacy-into-esg-frameworks/.

[55] See supra at note 15.

14. Which non-U.S. or U.S. (federal, state, or local) laws and regulations already requiring an AI audit, assessment, or other accountability mechanism are most useful and why? Which are least useful and why?

CIPL has provided feedback to proposed AI laws, regulations, and policies in jurisdictions including the European Union, Brazil, the UK, and Canada.[56] We have observed promising directions as well as areas for improvement across these proposals. For example, with respect to the EU's proposed AI Act, CIPL welcomed the Act's conception as a risk-based regulation,[57] but urged that the final version include organizational accountability requirements that are proportionate and flexible enough to respond to evolving use cases, with a judicious approach to classification of systems as "high-risk."[58] With respect to the UK,[59] CIPL pointed to the role that expert bodies such as the UK Centre for Data Ethics and Innovation, which has developed a portfolio of AI assurance techniques, can play in providing expert advice to regulators.[60] As a general observation, any requirements for ex ante assessments need to have set deadlines and evaluators with the necessary skillsets, or risk stifling innovation, especially for smaller organizations.

ACCOUNTABILITY SUBJECTS

15. The AI value or supply chain is complex, often involving open source and proprietary products and downstream applications that are quite different from what AI system developers may initially have contemplated. Moreover, training data for AI systems may be acquired from multiple sources, including from the customer using the technology. Problems in AI systems may arise downstream at the deployment or customization stage or upstream during model development and data training.

a. Where in the value chain should accountability efforts focus?

A comprehensive framework for AI accountability requires focus on all parts of the value chain. All organizations/parties along the value chain, from developers to deployers of AI systems, should implement their own accountability framework or program that is relevant to their activities related to the AI. Such accountability frameworks should address all key elements of organizational accountability as described above, including proper risk assessments relevant to the activities of the participant in the value chain. One recent white paper proposes that AI regulation be fine-tuned to focus on specific layers of the AI "technology stack" (e.g., data center infrastructure, foundational

---

[56] See CIPL Public Consultations, available at https://www.informationpolicycentre.com/public-consultations.html.
[57] See CIPL's Response to the EU Commission's Consultation on the Draft AI Act, July 29, 2021, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_the_consultation_on_the_draft_ai_act__29_july_2021_.pdf.
[58] Natascha Gerlach, "The case of the EU AI Act: Why we need to return to a risk-based approach," IAPP, March 23, 2023, available at https://iapp.org/news/a/the-case-of-the-eu-ai-act-why-we-need-to-return-to-a-risk-based-approach/.
[59] See CIPL's Response to UK Department for Digital, Culture, Media and Sport (DCMS) Policy Paper on Establishing a Pro-innovation Approach to Regulating AI, September 23, 2022, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_uk_dcms_proposed_approach_to_regulating_ai_23_09_22.pdf.
[60] See supra at note 42.

models, and applications).[61] The same principle applies to accountability: managing risk for data centers that hold large and sensitive data sets used to train models means something different than for models and consumer-facing applications running on them.

The initial proposal of the EU AI Act, had room for improvement in the balance between the developers and deployers of AI technologies. The nature of general purpose AI, in particular, requires care in apportioning accountability across these two groups, in recognition of the fact that certain and especially nefarious uses may be beyond the reasonable ability of the developer to anticipate or control.

b. How can accountability efforts at different points in the value chain best be coordinated and communicated?

Accountability can be communicated through a comprehensive and coherent regulatory framework or approach covering the entire value chain, including requirements for accountability across the specific elements of the AI technology stack. Transparency is a key ingredient of coordination: customers along that stack can require vendors to share their approaches to AI accountability and key data, where appropriate (e.g., model weights and "human in the loop" steps to monitor for bias).

c. How should vendors work with customers to perform AI audits and/or assessments? What is the role of audits or assessments in the commercial and/or public procurement process? Are there specific practices that would facilitate credible audits (e.g., liability waivers)?

One practice that would ensure credible audits is to have third-party certifications of AI applications for vendors that would streamline their customers' due diligence with respect to the AI application. Such certification would involve assessments and audits by the third-party certifier with respect to the relevant criteria for the certification.

Cloud computing and information technology (IT) services offer some interesting parallels and lessons for AI-related policy. Whilst cloud and IT service providers recognize the need for independent audits, they are reluctant to allow each client the separate right to audit and inspect their systems and premisses. Instead, third party audit reports and certification schemes can play a useful role by providing assurances without unduly burdening providers.

d. Since the effects and performance of an AI system will depend on the context in which it is deployed, how can accountability measures accommodate unknowns about ultimate downstream implementation?

Organizations can do well-informed, forward-thinking analysis of the broadest range of potential outcomes if their teams overseeing AI accountability include individuals with diverse life experiences (e.g., age, ethnicity, geography, religion, gender, sexual orientation), expertise (e.g., engineering, the humanities, data science, social sciences, political science), and corporate functions (e.g., development, deployment, sales, human resources, government relations). This analysis is an integral part of the purpose and function of any risk assessment within the context of an organization's accountability framework or program. Such risk assessments must attempt to identify all reasonably

[61] See Microsoft, *Governing AI: A Blueprint for the Future,* May 2023, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw.

conceivable risks that might be associated with the AI system and assess their likelihood and severity to enable appropriate mitigations and controls to address them. Such risk assessments might miss total "unknowns", but those might be "knowable" and identified and addressed by parties who implement the technology downstream in their own risk assessments.

16. ***The lifecycle of any given AI system or component also presents distinct junctures for assessment, audit, and other measures. For example, in the case of bias, it has been shown that "[b]ias is prevalent in the assumptions about which data should be used, what AI models should be developed, where the AI system should be placed—or if AI is required at all." How should AI accountability mechanisms consider the AI lifecycle? Responses could address the following:***

    a. ***Should AI accountability mechanisms focus narrowly on the technical characteristics of a defined model and relevant data? Or should they feature other aspects of the socio-technical system, including the system in which the AI is embedded? When is the narrower scope better and when is the broader better? How can the scope and limitations of the accountability mechanism be effectively communicated to outside stakeholders?***

[No response provided.]

    b. ***How should AI audits or assessments be timed? At what stage of design, development, and deployment should they take place to provide meaningful accountability?***

[No response provided.]

    c. ***How often should audits or assessments be conducted, and what are the factors that should inform this decision? How can entities operationalize the notion of continuous auditing and communicate the results?***

To ensure comprehensive internal assessments, it is important to allow organizations some level of flexibility to determine when to complete risk assessments for AI products and services. As discussed above, a reasonable requirement would ensure that both AI developers and deployers conduct a risk assessment once and then again in the event of material changes to the underlying data sets and models, which can include changes in business models, risk awareness, law, technology, and other external and internal factors. Additionally, the rate of audits and assessments should depend on the level of risk associated with the AI application or use; the higher the risk, the more frequently the system should be assessed.

    d. ***What specific language should be incorporated into governmental or non-governmental policies to secure the appropriate timing of audits or assessments?***

Please see our answer to the previous question.

17. How should AI accountability measures be scoped (whether voluntary or mandatory) depending on the risk of the technology and/or of the deployment context? If so, how should risk be calculated and by whom?

Implementing accountability and compliance frameworks, including risk assessments, should be mandatory. The required mitigations should depend on the outcome of contextual risk assessments, which are a core feature of any accountability framework or program, so there needs to be flexibility on the types of accountability measures. Organizations should assess the risks associated with their AI technology and uses. However, CIPL strongly believes that there must be a wider debate about what risks/harms must be consider as well as on appropriate methodologies, such as the NIST AI Risk Management Framework. There is a need for consensus-building regarding what constitutes risks and harms and how to assess these in light of the potentially important benefits of AI technologies for people, economies, and societies. Organizations should have flexibility to develop and fine-tune their risk assessment methodologies, so long as they are effective and outcomes and decisions are demonstrable and explainable to regulators.

Finally and importantly, risk assessments should consider the risks and potential harms resulting from NOT using the AI application at hand as well as weigh the overall benefits to individuals and society of the application against the identified harms.

18. Should AI systems be released with quality assurance certifications, especially if they are higher risk?

CIPL supports the use of certifications. It is important for any national regulatory approach to recognize that AI systems perform a range of tasks and, depending on the context of the AI system, can range from low- to high-risk use. One might consider reserving certification requirements for high-risk AI systems, while keeping in mind that the risk associated with a particular AI system is highly contextual.

19. As governments at all levels increase their use of AI systems, what should the public expect in terms of audits and assessments of AI systems deployed as part of public programs? Should the accountability practices for AI systems deployed in the public sector differ from those used for private sector AI? How can government procurement practices help create a productive AI accountability ecosystem?

As a general matter, organizational accountability requirements should apply to both the private and public sector organizations and the public should be able to expect consistent rigor of accountability in the public and private sectors. The accountability frameworks and programs in both sectors should address the same key elements of organizational accountability, including oversight, risk assessment, transparency, policies and procedures, training and awareness, monitoring and verification, and response and enforcement.[62] However, each such accountability framework or program should be adapted to the specific context and the subject matter and purpose of the program.

[62] See Figure 1, page 2.

ACCOUNTABILITY INPUTS AND TRANSPARENCY

20. What sorts of records (e.g., logs, versions, model selection, data selection) and other documentation should developers and deployers of AI systems keep in order to support AI accountability? How long should this documentation be retained? Are there design principles (including technical design) for AI systems that would foster accountability-by-design?

All accountability frameworks and practices and all risk assessments must be recorded and demonstrable on request. The ability to demonstrate accountability is a key component of organizational accountability. In particular, this is important regarding AI models and training, given the nature of the technology. Documentation about the approach taken, inputs and outputs, and trade-offs must be maintained and shared downstream, as well as with regulators on request, in case of an investigation.

21. What are the obstacles to the flow of information necessary for AI accountability either within an organization or to outside examiners? What policies might ease researcher and other third-party access to inputs necessary to conduct AI audits or assessments?

[No response provided.]

22. How should the accountability process address data quality and data voids of different kinds? For example, in the context of automated employment decision tools, there may be no historical data available for assessing the performance of a newly deployed, custom-built tool. For a tool deployed by other firms, there may be data a vendor has access to, but the audited firm itself lacks. In some cases, the vendor itself may have intentionally limited its own data collection and access for privacy and security purposes. How should AI accountability requirements or practices deal with these data issues? What should be the roles of government, civil society, and academia in providing useful data sets (synthetic or otherwise) to fill gaps and create equitable access to data?

Governments, industry, and non-governmental stakeholders collectively have much work to do to create appropriate structures for the sharing of data to meet the needs of accountable AI. A useful point of departure is the OECD Recommendation on Enhancing Access to and Sharing of Data, adopted in 2021.[63] It is equally important to ensure that data protection regulations requiring purpose limitation and data minimization, for example, are flexible enough to enable appropriate uses of data for accountable AI.[64] Finally, data localization policies, stemming from data protection laws and other laws, impede organizations' and governments' ability to use data required for AI training and modelling. This negative impact on the proper development and use of AI should be specifically addressed as part of the G7 Data Free Flow with Trust initiative.[65]

23. How should AI accountability "products" (e.g., audit results) be communicated to different stakeholders? Should there be standardized reporting within a sector and/or

[63] See OECD Recommendation of the Council on Enhancing Access to and Sharing of Data, May 2021, available at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463.
[64] See supra at note 5 at page 18.
[65] See The G7 Digital and Tech Ministers' Meeting Ministerial Declaration, April 30, 2023, Available at https://www.soumu.go.jp/main_content/000879099.pdf.

across sectors? How should the translational work of communicating AI accountability results to affected people and communities be done and supported?

While standardization can be very helpful, it is important that any requirements include sufficient flexibility for results to be customized appropriately, according to context. For some applications, especially consumer-facing ones, encouraging publication of "plain language" summaries may be helpful. This also relates to the issue providing transparency that is tailored to the specific audience or purpose (e.g. consumer, regulator and enforcer, business partner, etc.). See discussion above at 3(e).

BARRIERS TO EFFECTIVE ACCOUNTABILITY

24. What are the most significant barriers to effective AI accountability in the private sector, including barriers to independent AI audits, whether cooperative or adversarial? What are the best strategies and interventions to overcome these barriers?

To overcome barriers to effective AI accountability practices, regulators should provide timely and responsive guidance, incentivize third-party certifications, and treat organizational accountability practices as mitigating factors in enforcement actions. Importantly, accountability practices, including assessments and audits, should be incentivized beyond the threat of enforcement. CIPL detailed what factors could specifically be considered as mitigating factors in enforcement actions in our white paper, "Organizational Accountability in Data Protection Enforcement: How Regulators Consider Accountability in their Enforcement Decisions", published in October 2021 in collaboration with Professor Christopher Hodges of Oxford University.[66]

These factors include:

- the existence of any frameworks, systems, programs, processes, practices, policies and procedures, measures or tools that organizations have put in place to comply with legal requirements or other external standards, or to implement their own internal behavioral objectives, corporate ethics requirements, goals and public promises;
- the organization's participating in relevant certifications, labels, seals, or codes of conduct;
- the effectiveness of an organization's current accountability mechanism(s) and instruments set forth in a) and b) above, and how they are operated;
- an organization's transparency around the existence of the mechanisms or instruments described in a) and b) above, and the organization's ability to demonstrate their existence and effectiveness;

[66] See CIPL White Paper, *Organizational Accountability in Data Protection Enforcement How Regulators Consider Accountability in their Enforcement Decisions*, October 2021, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_organizational_accountability_in_data_protection_enforcement_-_how_regulators_consider_accountability_in_their_enforcement_decisions__6_oct_2021__3_.pdf; see also CIPL Discussion Paper, *Incentivising Accountability: How Data Protection Authorities and Law Makers Can Encourage Accountability*, July 2018, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_accountability_paper_2_-_incentivising_accountability_-_how_data_protection_authorities_and_law_makers_can_encourage_accountability.pdf.

- an organization's current or historic cooperation with the DPA in an investigation or enforcement context, including in connection with questions around the existence or effectiveness of any mechanisms or instruments described in a) and b) above; and
- other relevant factors.

25. Is the lack of a general federal data protection or privacy law a barrier to effective AI accountability?

Yes, the US needs a comprehensive, risk-based federal privacy law to create baseline protections and consistency across industry and sectors. Insofar as many of the AI accountability issues relate to the use of personal data and the impact on individuals, a comprehensive privacy law would help address these issues and foster trust. For reference, CIPL examined and explained how data protection laws regulate AI systems in white paper titled, "Artificial Intelligence and Data Protection: How the GDPR Regulates AI."[67]

As CIPL wrote in a March 2023 blog post, US federal privacy legislation should require organizations to adopt and implement comprehensive accountability frameworks through which they assess and mitigate risks to individuals, provide transparency on their practices to stakeholders, and monitor and verify for effectiveness. Organizations should be required to assess risks associated with their uses of personal data, while enabling them to calibrate measures of protection in accordance with the level of risk.[68]

The proliferation of AI tools, their expanding impact on individuals and societies, and their reliance on large volumes of granular, often personal data, requires effective data protection by both the private sector and governments.[69] Clarifying the application of existing data protection law on AI will be essential to ensuring that protections apply to data most essential to individuals' privacy or that otherwise pose risk of harm from misuse.

26. Is the lack of a federal law focused on AI systems a barrier to effective AI accountability?

A risk-based and outcome-based federal AI law that builds upon and is consistent with a federal privacy law could be helpful for fostering effective AI accountability.

In a forthcoming paper, CIPL will describe its recommended elements of AI laws and regulations. They include:

- **A flexible and adaptable framework.** An effective approach to regulating AI should be able to evolve and adapt to changes in the AI ecosystem. It should be *technology agnostic*–rules should apply to any systems that meet definitions provided by the rules. These rules should also be *principle*- and *outcome*-based to enable organizations to progress towards the achievement of specified outcomes (e.g., fairness, transparency, accuracy, human oversight) through risk-based, concrete, demonstrable, and verifiable internal measures. Regulators

[67] See CIPL White Paper, *Artificial Intelligence and Data Protection: How the GDPR Regulates AI*, March 2020, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-hunton_andrews_kurth_legal_note_-_how_gdpr_regulates_ai__12_march_2020_.pdf.
[68] See *Congress: Keep Moving Forward on Federal Privacy Legislation, with Organizational Accountability Front and Center*, CIPL Blog, LinkedIn, March 3, 2023, available at https://www.linkedin.com/pulse/congress-keep-moving-forward-federal%3FtrackingId=9iVTRIaXShS00LT0TcYJww%253D%253D/?trackingId=9iVTRIaXShS00LT0TcYJww%3D%3D
[69] See supra at note 5 at page 4.

should bear in mind that global companies may be unable to comply fully with a very prescriptive framework due to complexities and potential conflicting rules posed by such different foreign regulations, making a principle-based approach more suitable.

- **A risk-based approach that considers risks and benefits holistically.** Any regulatory approach to AI should ensure a high level of protection for individuals' fundamental rights while encouraging innovative and responsible use and development of AI. A risk-based approach enables achievement of this goal, by encouraging practical and proportionate interventions. Its focus is on potential impacts of AI technology in the context of specific uses. In this framework, an effective AI regime could provide non-exhaustive criteria to assist organizations in their assessments of (a) the benefits of the AI, (b) the likelihood and severity of harm, (c) the probability of such harm occurring, and (d) the measures taken to mitigate it.
- **Building on existing legal foundations.** A flexible and adaptable AI regime should build on and interoperate with existing legal frameworks while considering the impact of AI applications in particular use contexts. Any new regulation should avoid creating duplicative obligations or conflicting requirements, which could lead to fragmentation and inconsistent protections for individuals and uncertainty regarding their rights.
- **Organizational accountability as a central element of AI regulations.** Organizational accountability should be a core element of AI regulations. Organizations should be required to have a comprehensive internal compliance program that they can demonstrate on request. Accountability is also an ongoing internal change management process, requiring regular updates from organizations to keep pace with evolving laws, regulations, technology, and business practices.
- **Incentivizing development and implementation of accountable AI practices.** While a core set of accountability practices should be required for organizations developing and deploying AI, regulations should incentivize adoption of broader accountability practices.
- **Creating mechanisms for coordination across existing and any new regulatory bodies**. As noted above, AI crosses sectors and disciplines which may be governed by different regulations and overseen by separate regulators, and coordination across them will be essential.
- **Opportunities for continued regulatory innovation**, through approaches such as policy prototyping (discussed above) and "sandboxes" to test the application of laws to innovative products and services.

27. What is the role of intellectual property rights, terms of service, contractual obligations, or other legal entitlements in fostering or impeding a robust AI accountability ecosystem? For example, do nondisclosure agreements or trade secret protections impede the assessment or audit of AI systems and processes? If so, what legal or policy developments are needed to ensure an effective accountability framework?

Intellectual property rights pose important questions, and challenges, for AI accountability frameworks to address. On the one hand, intellectual property enjoys robust protections in many legal regimes across the globe, and protection of those rights is important to incorporate in AI accountability framework from an ethical and compliance perspective. At the same time, there is concern within the external auditing community, which includes academics and public researchers, among others, that organizations developing or deploying AI will use terms of service to limit the abilities of third-party auditors. Specifically, language in terms of service agreements may be

construed to limit the access that researchers can have when auditing AI products and services.[70] In the United States, key questions about the applicability of IP rights in contexts such as generative AI remained under consideration by courts at the time of writing of this submission.[71]

28. What do AI audits and assessments cost? Which entities should be expected to bear these costs? What are the possible consequences of AI accountability requirements that might impose significant costs on regulated entities? Are there ways to reduce these costs? What are the best ways to consider costs in relation to benefits?

The cost of AI assessments and audits depends on the particular AI product and the context of its use. It is important to create effective audit and certification schemes that are scalable and affordable for organizations of all sizes. Costs can further be reduced by requiring all organizations to implement their own internal accountability frameworks and programs, as discussed above. In the context of such programs, all AI developers and deployers should be required to conduct context-based risk assessments. These risk assessments in turn can support external audits because they require organizations to document risks, risk mitigations, and decisions. Making them available to auditors or third-party certification bodies will also reduce costs. They also enable effective and streamlined, and hence less costly, regulatory investigations and enforcement.

29. How does the dearth of measurable standards or benchmarks impact the uptake of audits and assessments?

[No response provided.]

AI ACCOUNTABILITY POLICIES

30. What role should government policy have, if any, in the AI accountability ecosystem? For example:

Regulators have an important role to play in ensuring proper application of principle-based rules and co-regulatory frameworks. They also need to stay on top of AI technology developments and latest applications. To enable responsible AI innovation and experimentation, any US AI regulatory regime should encourage new and agile approaches to regulatory oversight. Regulators need to be ready and equipped with appropriate resources and skills to engage constructively on the topic of AI with industry and government bodies developing and using the technology. In addition, they will need

[70] See Shannon Bond, *NYU Researchers Were Studying Disinformation On Facebook. The Company Cut Them Off*, NPR, August 4, 2021, available at https://www.npr.org/2021/08/04/1024791053/facebook-boots-nyu-disinformation-researchers-off-its-platform-and-critics-cry-f; see also Nandita Sampath, *Opening Black Boxes: Addressing Legal Barriers to Public Interest Algorithmic Auditing*, Consumer Reports, available at https://innovation.consumerreports.org/new-paper-opening-black-boxes-addressing-legal-barriers-to-public-interest-algorithmic-auditing/.

[71] See, for example, Getty Images v. Stability AI, available at https://fingfx.thomsonreuters.com/gfx/legaldocs/byvrlkmwnve/GETTY%20IMAGES%20AI%20LAWSUIT%20complaint.pdf.

modern regulatory oversight tools such as regulatory sandboxes, policy prototyping projects and AI ethics boards, all of which play an important role in the AI regulatory toolbox.[72]

Regulatory sandboxes are important mechanisms for regulatory exploration and experimentation as they provide a test bed for applying laws to innovative products and services in the AI field. At the same time, given that there are so many unanswered questions surrounding AI governance, it is quite challenging to design and assess the most appropriate, feasible and balanced legislative instruments. Collaborative, multi-stakeholder policy prototyping can provide a safe space to explore, assess and develop different legislative models of governance prior to their actual enactment. Such tools may help to inform legislative and policy choices that are more suited to the quickly developing technology industry.

As noted above, regulators should also consider how to promote, incentivise and reward industry best practices and responsible approaches to AI development and use. Such incentives could include, for instance, recognising self-regulatory (and enforceable) commitments of organizations that publicly define the AI values and principles they implement along with progress against benchmarks, using demonstrated accountability as a "licence to operate" by allowing accountable and/or certified organizations greater opportunities deploy AI systems responsibly or using demonstrated AI accountability as a criterion for public procurement projects.[73]

a. Should AI accountability policies and/or regulation be sectoral or horizontal, or some combination of the two?

Some combination of the two may be optimal. Innovative technologies require agile regulatory oversight. As noted in response to questions above, AI accountability policies should be coordinated across any new regulation and the pre-existing ecosystem of sectoral regulations as much as possible. Smart and risk-based regulatory oversight will require streamlined collaboration across regulators and stakeholders. Thus, a high-level and principles-based AI accountability framework applicable across the board could be augmented through more specific regulatory guidance on a sectoral level, where necessary and appropriate.

CIPL believes all players in the AI ecosystem should be accountable for their roles in the development and use of AI systems. There should be accountability requirements for industries that have not historically been regulated, including the technology industry that is developing the AI systems and FinTech organizations that are using the systems. For more regulated industries, like financial services or healthcare, existing regulatory requirements and risk management frameworks should be equally applicable to the use of AI and regulators in those industries should determine whether any additional enhancements to existing guidance is needed. For example, financial institutions need to ensure that their credit decisions are non-discriminatory so if the financial institution is using AI or generative AI in making such decisions it must do so in a way that complies with existing regulatory requirements.

---

[72] See CIPL White Paper, *Regulatory Sandboxes in Data Protection: Constructive Engagement and Innovative Regulation in Practice*, March 2019, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_regulatory_sandboxes_in_data_protection_-_constructive_engagement_and_innovative_regulation_in_practice__8_march_2019_.pdf.

[73] See CIPL Response to the EU Commission's Consultation on the Draft AI Act 2021, submitted July 29, 2021, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_the_consultation_on_the_draft_ai_act__29_july_2021_.pdf.

In some instances, regulatory guidance may need to be updated to account for regulatory expectations around accountability in the use of AI products for these types of decisions.

b. Should AI accountability regulation, if any, focus on inputs to audits or assessments (e.g., documentation, data management, testing and validation), on increasing access to AI systems for auditors and researchers, on mandating accountability measures, and/or on some other aspect of the accountability ecosystem?

These are all important elements of effective AI accountability regulation. Organizations should be required to demonstrate accountability, with flexibilities that enable consideration of the context of the AI deployment.

c. If a federal law focused on AI systems is desirable, what provisions would be particularly important to include? Which agency or agencies should be responsible for enforcing such a law, and what resources would they need to be successful?

As noted in response to Question 26 above, CIPL recommends a risk-based, outcomes-based, and layered approach to regulating AI that builds on existing laws and standards and accountable practices of organizations. This approach should be backed by innovative regulatory oversight and co-regulatory instruments.

(1) Rely on impact assessments performed by organizations to trigger the application of the law that would consider the context and impact of a proposed use of AI, rather than the sector it is utilised in or its type. The regulatory framework would provide illustrations of rebuttable presumptions of high-risk, rather than rigid pre-defined classifications. Organizations would assess the overall output and impact of the AI application, including its benefits and potential reticence risk, rather than focusing on risk only.

(2) Foster innovation through accountable practices of organizations. Rather than imposing prescriptive and indiscriminate requirements, the regulatory approach should set forth a general risk-based accountability requirements and outcomes that organizations should achieve through concrete, demonstrable and verifiable risk-based accountability measures.

(3) Enable consistent and modern approaches to regulatory oversight based on the current ecosystem of regulators. This approach should be complemented by a consistent scheme of voluntary, but enforceable, codes of conduct, certification and labelling, which should be designed through consultation with stakeholders. As discussed above, regulatory sandboxes and policy prototyping can be useful for enabling regulatory iteration in response to technological innovations.[74]

d. What accountability practices should government (at any level) itself mandate for the AI systems the government uses?

Government will have unique considerations compared to non-governmental organizations considering its responsibilities for operationalizing governance, advancing the public good, and

[74] See CIPL's Response to the EU Commission White Paper "On Artificial Intelligence – A European approach to excellence and trust", June 2020, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_eu_consultation_on_ai_white_paper__11_june_2020_.pdf.

protecting the rights and well-being of individuals. Given the impact of government decisions on people, it is essential that any use of AI by government be accountable, fair and proportionate, safe, and legal. There must be appropriate oversight over government use of AI, including external AI ethics review boards or committees.

31. What specific activities should government fund to advance a strong AI accountability ecosystem?

A risk-based approach to AI oversight should be complimented by federal-level schemes of voluntary but enforceable codes of conduct, standards, certifications, and labelling offered by independent bodies. These tools must be scalable to all sizes or organizations, adaptable to specific contexts, and affordable. These co-regulatory tools will enable organizations to foster trust with the broader public by demonstrating that an AI application meets legal criteria. The US government should design co-regulatory tools through consultation with diverse stakeholders and will need to update them regularly based on technological developments and new practices.

Also, the government should incentivise and fund projects that promote Privacy Preserving Technologies (PPTs) and Privacy Enhancing Technologies (PETs), such as the recent agreement between the US and UK governments to fund a prize for the most effective and useful PETs.[75]

32. What kinds of incentives should government explore to promote the use of AI accountability measures?

As noted above, any regime regulating AI in the US should incentivize the development and implementation of accountable AI practices and recognize organizational accountability practices as mitigating factors in the enforcement context. Organizations should be encouraged to adopt such practices to enable AI innovations in a responsible manner while also ensuring compliance with any AI regime and appropriate protections for individuals.

Regulators can illustrate best practices in the form of regulatory guidance to assist organizations in making appropriate assessments and choices.

As discussed above, CIPL has published two white papers describing ways for law and policy makers, as well as enforcement authorities, to incentivize organizational accountability.[76]

33. How can government work with the private sector to incentivize the best documentation practices?

Regulatory guidance, especially with regard to risk assessments, can promote and incentivize comprehensive and appropriate documentation practices. It should be developed in cooperation with the private sector to ensure practicability and wide adoption. Given the fast nature of AI development, it is essential that government and private sector organizations engage constructively, including with

[75] See White House Press Release, *US and UK to Partner on Prize Challenges to Advance Privacy-Enhancing Technologies*, December 8, 2021, available at https://www.whitehouse.gov/ostp/news-updates/2021/12/08/us-and-uk-to-partner-on-a-prize-challenges-to-advance-privacy-enhancing-technologies/#:~:text=Building%20on%20decades%20of%20investment%20in%20privacy-enhancing%20technologies%2C,team%20of%20specialists%20from%20across%20the%20UK%20Government..

[76] See supra at note 66.

individual regulators. There must be more exchange of views, previews of key new developments, capacity building by private sector among public sector agencies and regulators, and open dialogue.

34. Is it important that there be uniformity of AI accountability requirements and/or practices across the United States? Across global jurisdictions? If so, is it important only within a sector or across sectors? What is the best way to achieve it? Alternatively, is harmonization or interoperability sufficient and what is the best way to achieve that?

Harmonization and interoperability, both on a national and global level, should be a goal as much as possible, keeping in mind that any level of consistency and uniformity should continue to ensure the context- and risk-based flexibility that is required for effective AI governance and innovation. In general, consistent approaches to AI regulation will improve and facilitate effective compliance and thus the ultimate outcomes for individuals and society. Effective AI policy should not be confined to national borders and there should be international cooperation to develop dialogue, build consensus on how to regulate AI, create common models of regulation and co-regulatory tools, and address barriers to adoption of trustworthy and accountable AI, while at the same time ensuring countries and their people reap benefits of these new technologies. This may require setting up of an international oversight body to deal with these topics.

ANNEX A

Mapping Best Practices in AI Governance to the CIPL Accountability Framework

The following table outlines examples of accountable AI activities undertaken by select organizations of different sectors, geographies, and sizes, based on the CIPL Accountability Framework and against each accountability element. The practices are not intended to be mandatory industry standards, but serve as specific examples that are calibrated based on risks, industry context, business model, size, and maturity level of organizations.

ACCOUNTABILITY ELEMENT	RELATED PRACTICES
Leadership and Oversight	• Public commitment and tone from the top to respect ethics, values, specific principles in AI development, deployment and use • Institutionalized AI processes and decision-making with escalation criteria • AI/ Ethics/ Oversight Boards, Committees (internal or external) - to review risky AI use cases and to continuously improve AI practices • Appointing a board member for AI oversight • Appointing a responsible AI lead, AI officer or AI champion • Setting up an internal interdisciplinary AI board or AI committee • Ensuring inclusion and diversity in AI model development and AI product teams
Risk Assessment	• Algorithmic impact assessment or fairness assessment tools to monitor and continuously test algorithms to avoid human bias, unfair discrimination and concept drift throughout the entirety of AI lifecycles • Ethics impact assessment / human rights impact assessment / Data protection impact assessment • Developing standardized risk assessment methodologies, which take into account the benefits and the likelihood and severity of risk factors on individuals and/or society, level of human oversight involved in individually automated decisions with legal effects as well as their explainability according to context and auditability • Trade-offs documentation (e.g., accuracy—data minimization, security—transparency, impact on few—benefit to society) for high-risk processing as part of the risk assessment • Data quality assessment via KPIs • Data evaluation against the purpose—quality, provenance, personal or not, synthetic, in-house or external sources • Framework for data preparation and model assessment – including feature engineering, cross-validation, back-testing, validated KPIs by business • Working in close collaboration between business and data experts (data analysts, data engineers, IT and software engineers) to regularly assess the needs and accuracy results to ensure that the model can be properly used
Policies and Procedures	• Adopting specific AI policies and procedures on how to design, use or sell AI • Policies on the application of privacy and security by design in AI life cycle • Rule setting the level of verification of data input and output • Pilot testing of AI models before release • Use of protected data (e.g., encrypted, pseudonymised, tokenised or synthetic data) in some models • Use of high quality but smaller data sets • Use of federated AI learning models, considering trade-off with data security and user responsibilities • Special considerations for organizations creating and selling AI models, software, applications • Due diligence/self-assessment checklists or tools for business partners using AI • Definition of escalation steps with regard to reporting, governance, and risk analysis

	• Ideation phase between all stakeholders (data scientists, business, final user, control functions) where needs, outcomes, validations rules, maintenance, need for explainability, budget, are discussed
Transparency	• Different needs for transparency to individuals, regulators, business partners and internally at the different stages of AI lifecycle based on context • Adequate disclosures communicated in simple, easy to understand manner • Take into account that AI must be inclusive and accessible by those with special needs/disabilities • Set up a transparency trail for explainability of decisions and broad workings of algorithm to make the AI system auditable • Explain that it is an AI/ML decision, if possibility for confusion (Turing test) • Provide counterfactual information • Understand customers' expectations and deploy based on their readiness to embrace AI • Implement tiered transparency • From black box to glass box—looking at the data as well as algorithm/model • Aspiration of explainability helps understand the black box and builds trust • Define criteria of deployment of AI technologies within the organization based on usage scenarios and communicate them to the user • Produce model cards (short documents accompanying AI models to describe context in which model should be used, what is the evaluation procedure) • Data hub for transparency on data governance, data accessibility, data lineage, data modification, data quality, definition, etc. • Tailor transparency to the identified risk: e.g. watermarking for generative AI output
Training and Awareness	• Data scientist training, including how to limit and address bias • Cross functional training – privacy professionals and engineers • Ethics and fairness training to technology teams • Uses cases where problematic AI deployment has been halted • Role of "translators" in organizations, explaining impact and workings of AI
Monitoring and Verification	• Capability for human in the loop in design, in oversight, in redress • Capability for human understanding of the business and processes using AI • Capability for human audit of input and output • Capability for human review of individual decisions with legal effects • Monitoring the eco-system from data flow in, data process and data flow out • Reliance on different audit techniques • Reliance on counterfactual testing techniques • Pre-definition of AI audit controls • Internal audit team specialised on AI and other emerging technologies • Processes must allow human control or intervention in the AI system where both technically possible and reasonably necessary • Model monitoring (back-testing and feedback loop) and maintenance process
Response and Enforcement	• Processes and procedures to receive and address feedback and complaints • Redress mechanisms to remedy an AI decision • Redress to a human, not to a bot • Feedback channel