Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Centre for Information Policy Leadership

Ten Recommendations for Global AI Regulation

October 2023

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Ten Recommendations for Global AI Regulation

Contents

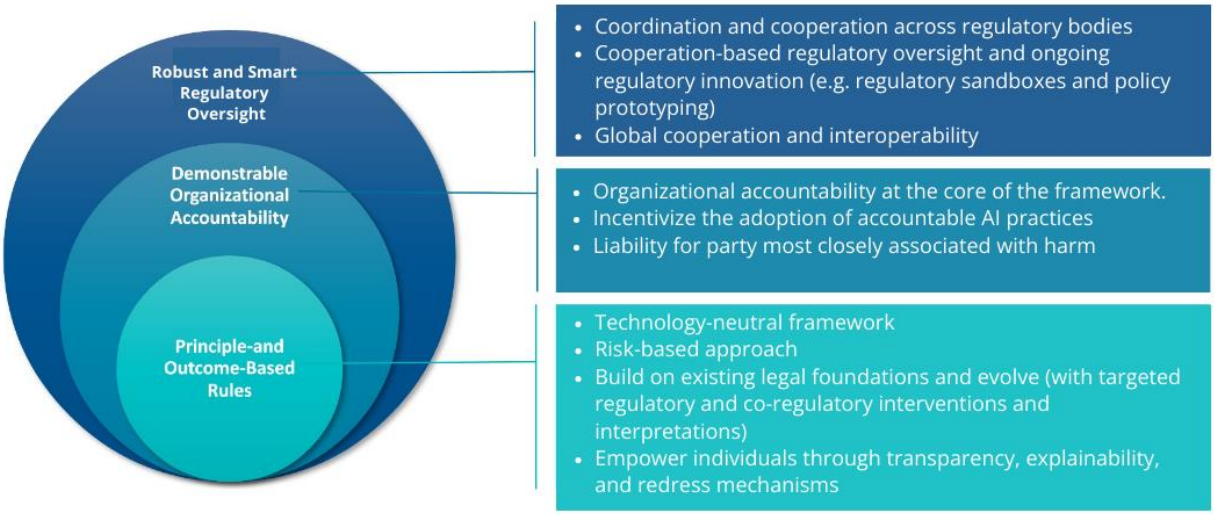
INTRODUCTION

Artificial Intelligence (AI)[i] is generating wide and growing societal benefits, including powering medical research, addressing climate change, transforming industries, and modernizing governments. At the same time, the rapid rollout and adoption of new applications, such as generative AI chatbots and image generators, have intensified longstanding concerns and raised new questions related to privacy and data protection, transparency and explainability, human rights, intellectual property, security, bias, workforce impacts, generation and dissemination of misinformation and disinformation; and other societal effects. In response, organizations are developing operational controls and governance frameworks to ensure responsible development and deployment of AI; industry experts are working to develop standards; policymakers are writing new laws; and regulators are testing the limits of existing authorities and proposing new ones.[ii] However, there is no consensus among countries on the best approach to regulating AI: should the focus be hard regulation, co-regulatory models, certifications and assurances, industry standards, or some combination?[iii]

CIPL has been a thought leader on organizational accountability and a risk-based approach to data policy and practices for over 20 years, and was an early contributor toward scoping challenges and defining solutions for AI governance and industry practices. Key CIPL contributions in this space include *Artificial Intelligence and Data Protection in Tension* (October 2018), *Hard Issues and Practical Solutions* (February 2020) and *Artificial Intelligence and Data Protection: How the GDPR Regulates AI* (March 2020).[iv] CIPL has also prepared detailed responses to public consultations on AI policy in Brazil, the European Union, the United Kingdom, and the United States.[v]

Drawing on this experience and our extensive engagement with private sector leaders developing and deploying AI technologies, policymakers, and regulators, CIPL offers in this paper ten recommendations to guide AI policymaking and regulation to enable accountable, responsible, and trustworthy AI. These ten recommendations encapsulate CIPL's view on a layered or three-tiered approach to AI regulation:

a) principle- and outcome -based rules,
b) demonstrable organizational accountability, and
c) robust and smart regulatory oversight.



Such an approach provides future-proof rules grounded in core principles that can guide ethical development and deployment of AI even as technology and use cases evolve. We describe this approach below.

Recommendations for Regulating AI

CIPL recommends a risk-based and tiered approach to regulating AI that builds on existing laws and standards and on accountable practices of organizations. This approach should be backed by innovative regulatory oversight and co-regulatory instruments.

Any legislative or regulatory approach to AI should follow these overarching recommendations:

A. Principle-and Outcome-Based Rules
   1. Create a flexible and adaptable framework that defines the outcomes to be achieved, rather than prescribing details of how to achieve them
   2. Adopt a risk-based approach that considers risks and benefits holistically
   3. Build on existing hard and soft law foundations
   4. Empower individuals through transparency, explainability, and mechanisms for redress

B. Demonstrable Organizational Accountability
   5. Make demonstrable organizational accountability a central element of AI regulations
   6. Advance adoption of accountable AI governance practices
   7. Apportion liability carefully, with a focus on the party most closely associated with generating harm

C. Smart Regulatory Oversight
   8. Create mechanisms for coordination and cooperation across regulatory bodies
   9. Institute cooperation-based regulatory oversight and enable ongoing regulatory innovation
   10. Strive for global interoperability

I. PRINCIPLE- AND OUTCOME-BASED RULES

1. Create a flexible and adaptable framework that defines the outcomes to be achieved, rather than prescribing details of how to achieve them

To be effective, AI regulations must be able to remain relevant as technology and use-cases continue to advance. Any rules should be *technology neutral*: a framework that is overly prescriptive and specific to individual technologies or current business models and practices risks becoming quickly outdated and inhibiting beneficial innovations. Indeed, an approach based on lists of specific technologies will require frequent amendments to keep up with technological change. If rules include lists of presumptively high-risk technologies and applications, they should enable those presumptions to be rebuttable and to evolve over time.

Rules should also be *principle- and outcome-based*. They should enable organizations to ensure the required outcomes (e.g., fairness, non-bias, transparency, accuracy, security, human oversight)

through risk-based, verifiable internal policies, procedures and controls that are appropriate in their specific contexts, without prescribing how to achieve these outcomes. Such an approach provides developers the flexibility to innovate, including the ability to innovate in the actual controls, technical tools and safeguards, while maintaining consistency with core principles and outcomes.[vi]

Similarly, a regulatory framework should not be overly prescriptive as to methodologies of AI impact- and risk assessments, but describe criteria that should be considered when assessing risks and benefits of an AI application and leave it to the competent regulators to provide further tailored, realistic and practical guidance in collaboration with those developing and deploying AI technologies.

At the same time, the rules should provide as much certainty as possible regarding their scope of application. For example, a regulatory framework for AI must define AI so that stakeholders can clearly understand what systems are covered by the rules. Absent such clarity and outcome-based focus, regulatory ambiguity and regulatory over-prescription will risk inhibiting investment and innovation—especially for small- and medium-sized enterprises (SMEs) and start-ups that are powerful engines of AI innovation and investment.

2. Adopt a risk-based approach that considers risks and benefits holistically

Any regulatory approach to AI should seek to protect fundamental human rights and minimize risks to individuals and society, while enabling development and use of AI for the benefit of both. Indeed, risks to individuals and society can also materialize from *not* using effective AI technologies, such as those that have the capacity to predict and prevent disease, or to reduce online harm, cybersecurity threats and fraud. A holistic risk-based approach promotes this goal by facilitating practical protective measures that are proportional to the risks and benefits of a particular AI system. Its focus is on potential impacts of AI technology in the context of specific use cases.

A risk-based regulatory framework for AI would provide non-exhaustive criteria to assist organizations to determine the likelihood and severity of any harm resulting and the measures required to mitigate it. Assessing and understanding the potential impact of their AI applications allows organizations to tailor their mitigations to the actual risks and avoid the implementation of unnecessary measures. For example, k-nearest neighbors algorithm (KNN) is a machine learning algorithm used in a variety of applications[vii]: in retail to recommend products, in healthcare to predict the risk of heart attacks and prostate cancer, in finance to detect fraudulent activity, in agriculture to predict crop yield and in transportation to predict and optimize traffic. These different uses of the same KNN algorithm have different levels of risk – the likelihood and severity of harm in recommending songs or clothing differ from the same in emergency medicine.

A risk-based framework should also assess the potential benefits of an AI system for individuals, organizations and society. These can then be weighed against the identified risks of deploying (or not deploying) AI. For example, the risks from autonomous vehicles (AVs) depend on the different surroundings in which they are deployed. There is arguably a lower risk of harm to people in deploying autonomous vehicles in mining and farming as opposed to urban or residential areas. At the same time, autonomous vehicles in the former settings offer different benefits, such as easing labor supply, supporting sustainable farming and improving productivity[viii]. The same example is helpful for highlighting the importance of weighing carefully the metrics by which to measure risk: one could assess risk of AVs versus the status quo (a world where most cars are driven by people) and/or a designated, optimal standard for AVs.

In short, the outcomes of holistic risk assessments for the use of a specific AI technology may vary significantly across use cases. From a policymaking perspective, this means that it may be hard to definitively identify high or low risk uses in advance, as the context of deployment is key. A risk-based approach is preferable to a categorical approach of defining AI systems that are automatically deemed to be high-risk. For example, considering all AI systems that are making inferences based on biometric data as high risk could encompass relatively low-risk ancillary uses, such as where AI is used to apply filters or improve video quality in video calls.

A more suitable approach would be:

A) to describe factors, criteria and potential harms that risk assessments should consider;
B) to provide, at most, an illustrative list of potentially higher vs. lower risk uses that can be rebutted in each particular case;
C) to provide ongoing guidance on how to assess risks and benefits based on learnings over time.

3. Build on existing hard and soft law foundations

A flexible and adaptable AI regime should build on existing legal frameworks, including regulations and legislation ("hard laws") and "soft law" (e.g., the OECD AI Principles). Many sectors where AI finds application are already highly regulated (e.g., healthcare, finance) and existing laws and regulations already provide requirements, compliance structures, and remedies that apply to the use of AI. However, relevant existing laws and regulations may also have to be interpreted in a new way and adapted to the realities of AI. Where there are regulatory gaps concerning AI-related risks, they should be closed with targeted regulatory and co-regulatory intervention, prioritizing sectors where existing regulations do not apply.

Relying on existing hard law frameworks to the extent possible reduces the risk of creating overlapping or conflicting rules that could lead to legal uncertainty and inconsistent protections. Existing anti-discrimination, consumer protection, intellectual property, and importantly, data protection and privacy rules are relevant to address many of the most important risks associated with AI. For example, in March 2020, CIPL produced a comprehensive analysis describing how the EU General Data Protection Regulation already regulates AI in relation to use of personal data.[ix] Where there are gaps in legislative frameworks—as in the United States, which currently lacks a comprehensive federal privacy law—filling them is an important foundation for sound AI regulation.[x] Where existing frameworks are relevant, regulatory agencies can foster compliance by issuing guidance on how those rules apply to AI. By consulting with a range of stakeholders, regulators can identify circumstances where such guidance will be most useful.

Further, it is important to recognise that existing rules may require some adaptation and evolved regulatory interpretation to align them with developments in AI technology. For example, certain principles found in many data protection law, such as lawful basis for processing, purpose specification and use limitation, may be in tension with the needs of AI systems and the way they operate.

As to the concept of lawful basis, there may not be sufficient lawful bases in current data protection laws to enable AI developers to use sensitive categories of personal data, such as health, gender, and ethnicity to ensure the AI model is trained and operates in a way that does not result in biased or discriminatory outcomes. Also, if personal data are processed based on a specific legal basis for specified purposes, and used only for those purposes or for "compatible" purposes, often with

narrow interpretations as to what constitutes "compatible", this may contradict the nature of how AI algorithms operate and learn. Given AI's potential for discovering new and unforeseen uses of data, these principles may unnecessarily thwart beneficial applications of AI unless they are given a broader interpretation for the AI context. An example for such a broader interpretation would be to apply a broader definition of "compatible" that includes any purposes that do not negate or conflict with the initial purpose and do not increase the risk of harm to individuals. Another solution would be to consider algorithmic training as a purpose on its own, separate from the purpose of deploying the algorithm to a particular use-case. This would allow broader collection and use of data in the training phase in order to ensure proper training and functioning of the algorithm. Finally, similar tensions arise with respect to the principles of data minimization and retention limitation, which may limit algorithms' opportunities to "learn" through recognition of correlations. In short, if certain traditional data protection principles are interpreted too rigidly, they may block development and deployment of beneficial AI applications—or have unintended consequences such as introducing unwanted bias, by limiting access to diverse training data.[xi] Regulators need to be able to evolve the interpretation of existing data protection principles through regulatory guidance developed in consultation with AI developers and deployers.

Finally, existing rules must be augmented with soft law frameworks, industry standards and co-regulatory tools developed in partnership with stakeholders, such as codes of conduct, certifications, and assurance models. International standards can help establish baseline requirements for development and deployment of AI that reflect shared understandings and values arrived at through multistakeholder development processes. The G7 Digital and Tech Ministers reaffirmed the key role of standards at their Hiroshima Summit in April 2023[xii] and agreed in September to develop a Code of Conduct for organizations developing advanced AI systems[xiii], while the multi-stakeholder Certification Working Group is leading promising work on AI Certification .[xiv] Leveraging soft law frameworks such as the OECD AI Principles can foster international alignment on AI regulations: for example, the Parliament's version of the EU AI Act derives its definition of "Artificial Intelligence" from those principles.

4. Empower individuals through transparency, explainability, and mechanisms for redress

CIPL has advocated for individual empowerment as a core principle of sound privacy regulation, and the same holds true for AI. For AI to be trustworthy and beneficial to all, regulations, co-regulatory frameworks, and industry practices must empower individuals through:

- **Transparency.** Developers and deployers of AI should provide context-appropriate and meaningful transparency about the inputs and operations of AI systems, while preserving privacy and data protection, security, safety, and trade secrets. Such contextualized transparency should extend to business users of AI systems, auditors, regulators, and the general public.

    - **High-risk AI systems** should document how the system is intended to be used, known inappropriate uses and risks, and recommendations for deployers on how to manage those risks.
    - **Generative AI** requires steps to ensure that users understand models' data practices and limitations. Ideally, developers and deployers should provide transparency through multiple mechanisms, including policies, terms of service, in-product notifications, and centralized resource hubs.

- **Explainability.** Explainability is an aspect of transparency and a means of boosting accountability and trust. It requires that developers and deployers meaningfully explain how AI systems affect decisions and outcomes that impact individuals, while bearing in mind trade-offs, such as between explainability and security/safety, and explainability and accuracy. The more complex and accurate the algorithm is, the harder it may be to explain how it actually works. There may also be technical constraints on explainability in some circumstances. For example, it may not always be possible to explain how large language models (LLMs) generate specific results based on individual inputs or model parameters. Organizations will have to document the relevant trade-offs to demonstrate how and why they prioritised accuracy over explainability. A case in point may be AI algorithms used in healthcare and medicine, where AI may enable certain health benefits not achievable by non-AI tools, yet may not be explainable. Under those circumstances, accuracy may take precedence over explainabilty. In short, depending on the context, risks, and potential benefits of a specific use case, requiring full explainability as a condition for use may not be appropriate in all instances.

- **User Feedback and Redress.** Where individuals do not understand an AI-made decision, or believe they have been harmed by AI, there should be clear options for user feedback, inquiries, complaints, further transparency, the right to contest the decision, a requirement for human review and, ultimately, redress, as well as action by enforcement authorities, where appropriate and necessary. Developers and business users should consider how to enable further transparency, human review in case of a contested use of AI, as well as opportunities for complaint capture and redress as part of the design of end-to-end solutions that leverage AI.

II. DEMONSTRABLE ORGANIZATIONAL ACCOUNTABILITY

5. Make demonstrable organizational accountability a central element of AI regulations

To ensure accountability within the broader ecosystem, regulations should facilitate organizations' demonstrable use of accountability frameworks and governance programs that provide the tools and processes for organizations to implement all relevant legal requirements and other standards. As in other areas of traditional corporate compliance and business ethics—and more recently in data, security, and digital spheres—accountability must be built into and implemented across all stages of the AI lifecycle and the AI "technology stack", including AI datacenter infrastructure, models, and applications.[xv]

There are a variety of accountability frameworks that provide useful models for devising organizational accountability and AI governance programs, including the U.S. NIST AI Risk Management Framework, Singapore's Model AI Governance Framework, and CIPL's own Accountability Framework described in Annexes 1-3 of this report.[xvi]

Organizations also need to be able to *demonstrate* accountability internally – to their C-suite and corporate Boards, as well as externally - to shareholders, investors, regulators and the general public. Certifications, audits, codes of conduct, and assessments are helpful tools for demonstrating accountability. Indeed, these accountability mechanisms are essential in digital policy and regulation, including for developers and deployers of artificial intelligence, for the following reasons:

- They demonstrate to all actors across the organization a commitment and the ability to ensuring that products and services meet specific criteria.
- They enable organizations to translate principle- and outcome-based legal requirements into demonstrable and risk-based controls, ensuring more effective regulation and better compliance in practice.
- They play an important role in providing legal certainty and strengthening trust, including in business-to-business contexts.

Any AI regulation should explicitly include demonstrable accountability as a core element, as well as enable development and use of co-regulatory frameworks, such as certification schemes and codes of conduct, that facilitate and demonstrate such accountability.

6. Advance adoption of accountable AI governance practices

While a core set of accountability practices should be required for organizations developing and deploying AI, policymakers and regulators should also proactively encourage and incentivize adoption of broader accountability practices, frameworks, tools and technologies. They should work with stakeholders to co-develop tools and frameworks for building and demonstrating AI accountability. The goal should be to create an environment wherein organizations see adoption of well-developed accountability frameworks as differentiators for creating value and deepening trust in their data practices, beyond fulfilling baseline legal and regulatory obligations.

Policymakers and regulators should also understand drivers and challenges with respect to accountable technology practices and technology solutions, such a Privacy Enhancing Technologies (PETs), and take steps to incentivise their further development and wider adoption.[xvii]

Consideration should be given to a broad range of potential incentives for accountability[xviii], including:

- Formally recognising demonstrated or certified accountability as a mitigating factor in enforcement actions and in assessing sanctions and/or levels of fines;
- Using demonstrated accountability as a form of "licence to operate", by giving accountable organizations greater freedom to develop and deploy AI models responsibly;
- Allowing broader use of data in AI projects for socially beneficial research that has been validated by relevant risk assessments, mitigations, oversight, and controls in accountability programs;
- Enabling parties acquiring AI systems to fulfil due diligence requirements by procuring systems that have been certified to recognized standards for responsible AI.
- Using demonstrated AI accountability as an eligibility criterion for public procurement projects, to incentivize contractors to obtain responsible AI certification.

7. Apportion liability carefully, with a focus on the party most closely associated with generating harm

Adoption of organizational accountability mechanisms by all actors in the AI ecosystem will lead to better compliance and outcomes on the ground, and likely result in less need to resort to questions around liability. Yet, where liability does raise concerns, active debates are underway regarding the appropriate apportionment across parties in the AI ecosystem.

In principle, liability should be assigned chiefly to the party most closely associated with generating the harm in question, but assigning liability may be complex in practice. The analysis will be shaped by existing legal standards and precedent, as well as the extent to which parties disclose relevant information through transparency and reporting requirements.

Depending on the circumstances, liability might be assigned to the developer, the deployer, end users, or some combination. Developers might be the appropriate focus of liability for systems that have been insufficiently tested for potential harms, or provided to users with misleading indications on capabilities. On the other hand, users share responsibility for how they use AI systems, as they determine whether to use a system for a higher risk use or in ways that are expressly contraindicated by guidance provided by developers.

As in other areas of commerce, contracts—including specific, emerging AI contracting practices—will play an important role in apportioning the responsibilities and liabilities of parties in the AI development and deployment life cycle. For example, if a developer contractually prohibits a high-risk use case of their product, the risk of misuse should shift to the user who has wilfully breached the terms of that contract. In scenarios where third parties provide AI models or AI-enabled solutions, accountability between model developers and deployers should be specified in contracts.

III. SMART REGULATORY OVERSIGHT

8. Create mechanisms for coordination and cooperation across regulatory bodies

AI is used across sectors governed by different regulations and regulators. For example, data protection authorities (DPAs) will have general competence over the processing of personal data using AI. Other regulators have more sector-specific remits over AI applications—as in in healthcare, housing, financial services, telecommunications, pharmaceuticals and in cross-cutting disciplines, such as intellectual property. In most circumstances, there should not be a need for a new, overarching AI regulator, as that would likely result in regulatory overstep, overlap, inconsistency, and lack of legal certainty. Rather, it is more appropriate to

a) enhance the competencies and capabilities of existing regulators to be ready for AI oversight and supervision; and
b) enable high-level AI policy coordination and collaboration across existing authorities.

While each regulator should maintain competence over its own remit (e.g., for purposes of legal certainty, DPAs should retain general competence over AI applications involving the processing of personal data and/or impacting individuals' privacy), a standing central governmental coordination body should be created to set high-level AI policies and goals applicable across all sectors and industries, and facilitate alignment, regulatory coordination, and joint action between different regulatory bodies, where necessary and appropriate. The coordination body can provide regulators a space in which to discuss trade-offs between different policy objectives including efficiency, productivity, fairness, privacy, security and resilience. It can also provide clarity on where parties should turn for guidance in specific circumstances of AI development and deployment.

This approach would be beneficial to both organizations and regulators by fostering consistency in regulatory approaches as well as holistic and inter-disciplinary policy and guidance that is easier to implement and monitor by specialized regulators and industry over time. Such an approach may also be helpful for harmonizing new laws and regulations with existing ones.

An example of cross-regulatory coordination is the UK Digital Regulation Cooperation Forum (DRCF). It includes a permanent CEO and staff, joint activities, joint guidance, and other regulatory action, as well as formal collaboration projects and staff secondments. AI has been a focus of the DRCF's work, as evidenced by its multi-year workstream on algorithmic transparency. [xix] Other countries, such as Australia, France, Ireland, and the Netherlands, have also established cooperation mechanisms for regulators.[xx]

9. Institute cooperation-based regulatory oversight and enable ongoing regulatory innovation

As technology continues to evolve, regulators, regulatory techniques, and tools need to evolve as well.

a) Regulators need to enhance their capabilities, capacities and how they operate in a world where there are competing and multiple interests at stake. For example, the task of data protection authorities should not be limited to protecting fundamental rights of individuals, but also to enable responsible and accountable use of data and development of AI technology for the benefit of society and economy in a way that protects fundamental rights. This requires a shift in regulatory mindset, regulatory priorities and regulatory action. This shift is essential if current regulators are to remain relevant and effective in a new digital world.

b) Regulators should take a risk-based approach in order to be strategic and effective. This requires understanding the risks and benefits of AI systems and focusing on those areas that present the highest risks to individuals and society, while preserving the benefits of AI technology and its advancement. It also requires regulators to prioritize all of their work – regulatory strategy, guidance, supervision and enforcement and to focus on areas that create the highest risks for individuals and society.

c) Traditional oversight mechanisms based exclusively or primarily on *ex post* enforcement may no longer be sufficient in a digital- and AI enabled society. Reliance on fixing market failures by enforcement alone will not result in desired outcomes. Given the pace of advancement of AI technology and the need to understand its risks and benefits, there is a pressing need for a more co-operative approach based on ongoing constructive engagement between regulators and regulated entities, sharing of experiences and information on technological developments, and working together to develop realistic compliance targets and interpretations of applicable rules. This requires both regulators and regulated entities to be transparent and ready to engage in constructive information sharing in real time as technologies and business practices change.[xxi] Investing in *ex ante* measures such as incentivizing proactive and demonstrable accountability is likely to achieve better outcomes than expensive *ex post* enforcement. Of course, enforcement should still remain a regulatory option and an important lever for repeated, serious and negligent breaches that cause real harm to individuals and society.

d) Innovative regulatory tools, such as sandboxes and policy prototyping, can be effective for regulatory oversight of new technology such as AI. They provide regulators with deeper understanding and first-hand experience of AI applications and developments, aimed at the general market. They also provide for safe harbor for industry to test the risks and benefits

of responsible innovation with a direct link to the competent regulator. Governments must provide appropriate funding and resources for regulators to develop and engage in regulatory sandboxes and be able to scale these activities for larger group of participants, including on a sectoral basis.

Regulatory Sandboxes: 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 real-life settings under the supervision of a regulator.[xxii] They can be used to help address and resolve some of the more challenging aspects of deploying AI against the backdrop of prevailing legal requirements, particularly those that appear inconsistent or in tension with new technologies. Examples include:

- The UK information Commissioner Office has run an established sandbox program since 2020, with a particular focus on emerging technologies and biometrics;[xxiii]
- The Singapore Infocomm Media Development Authority (IMDA)'s Data Regulatory Sandbox enables businesses to obtain regulatory guidance for innovative, data-intensive technologies. IMDA also operates a specific sandbox to foster development and adoption of Privacy-enhancing Technologies (PETs).[xxiv]
- The Norwegian Data Protection Authority (Datatilsynet) launched a special regulatory sandbox for AI applications;[xxv]
- The Colombian government has developed a regulatory sandbox to promote Privacy by Design and Default in AI projects;[xxvi]
- France's CNIL operates a sandbox that has completed projects in digital health and educational technology. In 2023, it announced a new initiative focused on AI[xxvii];
- The EU's draft AI Act would enable, and may ultimately require, member states to establish regulatory sandboxes for AI. Spain was the first member state to pilot an AI sandbox.[xxviii]

Regulatory sandboxes should be designed in a manner that encourages innovation, information sharing, and other modes of cooperation. Any AI regulatory framework should provide an explicit statutory basis for regulators to set up sandboxes, including cross-regulatory sandboxes with appropriate and relevant regulators including data protection, competition, media, consumer, health/pharma, telecom, and financial regulators. Regulations should be mindful of how legal authorities or enforcement priorities may affect company participation. At the same time, to ensure public trust in the results, sandboxes should include assurances that individuals will continue to be protected from harms even as policy experimentation takes place.

Policy prototyping: These are pilot projects that mobilize public and private actors to jointly explore, assess, and develop different legislative models of governance prior to their actual enactment. The process typically involves selecting a group of participants, such as early-stage technology companies, to develop and apply policy prototypes in partnership with government, industry, and academic experts. Meta's OpenLoop program has been a leading practitioner of policy prototyping, including for the EU's proposed AI Act.[xxix] Singapore's IMDA has a policy prototyping program within its Data Regulatory Sandbox that has focused on notice, consent, and disclosure; AI transparency and explainability; and transparency and consent in the metaverse, including contexts for application of legitimate interest as a basis for processing.[xxx]

10. Strive for global interoperability

Given the global nature of AI technology – from the data it uses for training, to research and development, computing infrastructure, and applications that cross borders – it is clear that no government can satisfactorily address AI policy and regulation in isolation. Cooperation at the international level is essential to ensure that individuals and societies globally can rely on the benefits of trustworthy and accountable AI and that new risks are assessed and mitigated on an ongoing basis. This work would benefit from a dedicated international forum that enables governmental and other stakeholders to cooperate on AI policy.

Furthermore, international cooperation must foster interoperability of AI policies and regulations. As CIPL has noted in the context of data protection, global interoperability enables responsible provision of services across borders, broadens access, reduces compliance costs, increases legal certainty, and ensures consistent protection of the rights and interests of individuals.[xxxi] Different jurisdictions will have their own priorities, legal traditions, and body of existing regulation, but may be able to coalesce around core principles and approaches in considering AI policy and regulation – similar to those CIPL has advanced in this paper. They can also take steps to codify interoperability through recognition and certification mechanisms, including through participation in the Global Cross-Border Privacy Rules (CBPR) system in the context of data protection and trusted cross-border data flows[xxxii]. There have been encouraging efforts toward AI interoperability through the aforementioned G7 initiative, the OECD AI Principles,[xxxiii] trade and economic agreements like the Digital Economic Partnership Agreement (DEPA)[xxxiv], and the Global Partnership on AI.[xxxv]

ANNEX I – CIPL ACCOUNTABILITY FRAMEWORK

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

ANNEX II – 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 mapped 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

[i] For this report, CIPL uses the term "artificial intelligence" in a manner consistent with the definition of "AI systems" developed by the U.S. National Institute of Standards and Technology (NIST) in its Risk Management Framework 1.0, adapted from a comparable definition developed by the Organisation for Economic Co-operation and Development (OECD): "*An AI system [is referred to as] as an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy*." See https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[ii] For example: Italian Data Protection Authority "Garante" ban on ChatGPT on March 30, 2023 available at https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870832 ; European Data Protection Board (EDPB) creates task force on Chat GPT, available at https://edpb.europa.eu/news/news/2023/edpb-resolves-dispute-transfers-meta-and-creates-task-force-chat-gpt_en; UK Competition and Markets Authority launches initial review of artificial intelligence models, available at https://www.gov.uk/cma-cases/ai-foundation-models-initial-review; The Office of the Privacy Commissioner of Canada has launched an investigation into ChatGPT, available here https://www.priv.gc.ca/en/opc-news/news-and-announcements/2023/an_230404/; Joint statement on enforcement efforts against discrimination and bias in automated systems USA Consumer Financial Protection Bureau, Department of Justice's Civil Rights Division, Equal Employment Opportunity Commission and Federal Trade Commission available at https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint-Statement%28final%29.pdf.

[iii]According to the OECD more than 800 AI policy initiatives and strategies have been designed across 69 countries, territories, and the European Union https://oecd.ai/en/dashboards/overview.

[iv] CIPL, "First Report: Artificial Intelligence and Data Protection in Tension," October 2018, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_first_ai_report_-_ai_and_data_protection_in_tension__2_.pdf; "Second Report: Hard Issues and Practical Solutions," February 2020, 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; "Artificial Intelligence and Data Protection: How the GDPR Regulates AI," March 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-hunton_andrews_kurth_legal_note_-_how_gdpr_regulates_ai__12_march_2020_.pdf.

[v] CIPL, "Response to NTIA Request for Comment on AI Accountability Policy," June 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_ntia_ai_accountability_policy_june2023.pdf; "CIPL's Top Ten Recommendations for Regulating AI in Brazil," October 2022, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/[en]_cipls_top_ten_recommendations_for_regulating_ai_in_brazil__4_october_2022_.pdf; "Response to UK DCMS Proposed Approach to Regulating AI," September 2022, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_uk_dcms_proposed_approach_to_regulating_ai_23_09_22.pdf; "CIPL Response to the EU Commission's Consultation on the Draft AI Act," July 2021, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_the_consultation_on_the_draft_ai_act__29_july_2021_.pdf.

[vi] There are no universally accepted definitions for AI developers, deployers, and users; indeed, a 2023 taxonomy jointly developed by the EU and U.S. described the definitions of deployment, developer, and user as "pending" (See "EU-U.S. Terminology and Taxonomy for Artificial Intelligence: First Edition", https://www.nist.gov/system/files/documents/noindex/2023/05/31/WG1%20AI%20Taxonomy%20and%20Terminology%20Subgroup%20List%20of%20Terms.pdf). In this paper, we use the term "developers" to refer to parties that design and build AI systems, "deployers" as parties that make such systems available for use, and "users" as the end-users that operate those systems on an ongoing basis. A single entity could play each of these roles at different points or simultaneously.

[vii] *K-Nearest Neighbors Algorithm*, IBM, available at https://www.ibm.com/uk-en/topics/knn#:~:text=Next%20steps-,K%2DNearest%20Neighbors%20Algorithm,of%20an%20individual%20data%20point.

[viii] *3 ways autonomous farming is driving a new era of agriculture*, World Economic Forum, 2022 Available at: https://www.weforum.org/agenda/2022/01/autonomous-farming-tractors-agriculture/

[ix] CIPL, "Artificial Intelligence and Data Protection: How the GDPR Regulates AI," March 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-hunton_andrews_kurth_legal_note_-_how_gdpr_regulates_ai__12_march_2020_.pdf.

[x] For more on the intersection between AI and data protection regulation, see CIPL AI First Report - *Artificial Intelligence and Data Protection in Tension*; CIPL AI Second Report - *Hard Issues and Practical Solutions*; and CIPL/Hunton Andrews Kurth White Paper - *How the GDPR Regulates AI* – all available here https://www.informationpolicycentre.com/ai-project.html.

[xi] For more on this topic, see CIPL, "First Report: Artificial Intelligence and Data Protection in Tension," October 2018, cipl_first_ai_report_-_ai_and_data_protection_in_tension__2_.pdf (informationpolicycentre.com), and CIPL, "Second Report: Hard Issues and Practical Solutions," February 2020, AI Project - Centre for Information Policy Leadership (informationpolicycentre.com).

[xii] Ministerial Declaration The G7 Digital and Tech Ministers' Meeting 30 April 2023 (g7digital-tech-2023.go.jp)

[xiii] G7 Hiroshima AI Process: G7 Digital & Tech Ministers Statement", September 2023, accessed at 3e39b82d-464d-403a-b6cb-dc0e1bdec642-230906_Ministerial-clean-Draft-Hiroshima-Ministers-Statement68.pdf (politico.eu).

[xiv] Certification Working Group, "Unlocking the Power of AI – Steps for Effective Certification to Help Drive Innovation and Trust," June 2023, https://www.responsible.ai/post/white-paper-draft-from-the-certification-working-group.

[xv] For further discussion of governance across these layers of the AI Technology Stack, see Microsoft, "Governing AI: A Blueprint for the Future," May 2023, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw.

[xvi] National Institute of Standards and Technology (NIST), "AI Risk Management Framework," https://www.nist.gov/itl/ai-risk-management-framework; Personal Data Protection Commission (PDPC), "Singapore's Approach to AI Governance," https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework, CIPL, "Organizational Accountability," https://www.informationpolicycentre.com/organizational-accountability.html.

[xvii] See CIPL's report on Privacy-enhancing Technologies (forthcoming, Fall 2023).

[xviii] Incentivizing Accountability: How Data Protection Authorities and Law Makers Can Encourage Accountability, CIPL, 23 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..

[xix] The Competition and Markets Authority (CMA), the Information Commissioner's Office (ICO), the Office of Communications (Ofcom), and the Financial Conduct Authority (FCA) take part in the DRCF. See https://www.gov.uk/government/collections/the-digital-regulation-cooperation-forum. For more on the DRCF's algorithmic transparency workstream, please see https://www.gov.uk/government/publications/transparency-in-the-procurement-of-algorithmic-systems-findings-from-our-workshops.

[xx] In the Netherlands, the Authority for Consumers and Markets (ACM), the Dutch Data Protection Authority (AP), the Dutch Authority for the Financial Markets (AFM), and the Dutch Media Authority (CvdM) launched the Digital Regulation Cooperation Forum (SDT). See https://www.acm.nl/en/about-acm/cooperation/national-cooperation/digital-regulation-cooperation-platform-sdt. France's Center of Expertise for Digital Platform Regulation (PEReN) was formed under the authority of the Ministries of Economy, Culture, and Digital Technology. Grounded in expertise in data science, it is a source of technical expertise and support to France's digital regulators. See https://www.peren.gouv.fr/en/. Australia's Digital Platform Regulators Forum brings together the Australian Competition and Consumer Commission (ACCC), Australian Communications and Media Authority (ACMA), eSafety Commissioner (eSafety) and Office of the Australian Information Commissioner (OAIC). See https://www.accc.gov.au/about-us/media/media-

updates/communique-digital-platforms-regulators-forum. Ireland created the Economic Regulators Network, which brings together seven regulators. See https://www.econreg.ie/about/our-members/.

[xxi] See Christopher Hodges and CIPL, "Organizational Accountability in Data Protection Enforcement," October 2021, 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_.pdf

[xxii] See CIPL Paper "Regulatory Sandboxes in Data Protection – Constructive Engagement and Innovative Regulation in Practice - March 8, 2019 https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_regulatory_sandboxes_in_data_protectionconstructive_engagement_and_innovative_regulation_in_practice__8_march_2019_.pdf.

[xxiii] ICO Regulatory Sandbox https://ico.org.uk/for-organisations/regulatory-sandbox/.

[xxiv] IMDA Data Regulatory Sandbox, https://www.imda.gov.sg/how-we-can-help/data-innovation/data-regulatory-sandbox.

[xxv] Datatilsynet AI Regulatory Sandbox, available at https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/.

[xxvi] Sandbox on privacy by design and by default in Artificial Intelligence projects, Columbian Superintendence of Industry and Commerce, available at https://globalprivacyassembly.org/wp-content/uploads/2021/07/B6.-SIC-Colombia-Sandbox-on-privacy-by-design-and-by-default-in-AI-projects.pdf

[xxvii] CNIL, "Digital health and EdTech: the CNIL publishes the results of its first 'sandboxes'", July 2023, https://www.cnil.fr/en/digital-health-and-edtech-cnil-publishes-results-its-first-sandboxes.

[xxviii] At the time of writing, EU policymakers had not yet resolved whether to make member states' creation of AI sandboxes voluntary or mandatory. See Luca Bertuzzi, "EU Council sets path for innovation measures in AI Act's Negotiations", *Euractiv,* July 10, 2023. https://www.euractiv.com/section/artificial-intelligence/news/eu-council-sets-path-for-innovation-measures-in-ai-acts-negotiations/. See also "First Regulatory Sandbox on Artificial Intelligence Presented," *Digibyte,* June 27, 2022, https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented.

[xxix] See "Introducing Open Loop, a global program bridging tech and policy innovation", available at https://ai.facebook.com/blog/introducing-open-loop-a-global-program-bridging-tech-and-policy-innovation /; AI Impact Assessment: A Policy Prototyping Experiment, https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3772500_code715910.pdf?abstractid=3772500&mirid=1; and https://openloop.org/programs/open-loop-eu-ai-act-program/.

[xxx] IMDA, "Policy Prototyping," Policy Prototyping | IMDA - Infocomm Media Development Authority. Meta's TTC Labs is a partner for IMDA's program.

[xxxi] CIPL, "Ten Principles for a Revised US Privacy Framework," March 2019, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_principles_for_a_revised_us_privacy_framework.pdf.

[xxxii] CIPL, "International Data Flows: Cross Border Privacy Rules, Privacy Recognition for Processors, and Global CBPR and PRP: Frequently Asked Questions," June 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_cpbr_and_prp_faq_jun23.pdf.

[xxxiii] OECD, "Recommendation of the Council on Artificial Intelligence," May 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[xxxiv] The Digital Economic Partnership Agreement (DEPA) is an agreement among New Zealand, Singapore, and Chile. See https://www.mfat.govt.nz/en/trade/free-trade-agreements/free-trade-agreements-in-force/digital-economy-partnership-agreement-depa/.

[xxxv] The Global Partnership on Artificial Intelligence, https://gpai.ai/.