In collaboration
with Accenture

WORLD ECONOMIC FORUM

Governance in the Age of Generative AI:
A 360° Approach for Resilient Policy and Regulation

WHITE PAPER

OCTOBER 2024

Images: Getty Images, Midjourney

Contents

Foreword

Arnab Chakraborty
Chief Responsible AI Officer, Accenture

Cathy Li
Head, AI, Data and Metaverse; Deputy Head, Centre for the Fourth Industrial Revolution; Member, Executive Committee, World Economic Forum

We are living in a time of rapid innovation and global uncertainty, in which generative artificial intelligence (AI) stands out as a transformative force. This technology impacts various industries, economies and societies worldwide. With the European Union's (EU's) AI Act now in effect, we have a precedent for comprehensive AI regulation. The US, Canada, Brazil, the African Union, Japan and China are also developing their own regulatory approaches. This pivotal moment calls for visionary leadership and a collaborative approach to anticipatory governance.

Over the past year, the AI Governance Alliance has united industry and government with civil society and academia, establishing a global multistakeholder effort to ensure AI serves the greater good while maintaining responsibility, inclusivity and accountability. We have been able to position ourselves as a sounding board for policy-makers who are grappling with the difficulties of developing AI regulatory frameworks, and to convene all players from the AI value chain to create a meaningful dialogue on emerging AI development issues.

With Accenture as its knowledge partner, the Alliance's Resilient Governance and Regulation working group (composed of over 110 members), has contributed to shaping a shared understanding of the global regulatory landscape. The group has worked to establish a comprehensive governance framework that could be used to regulate generative AI use well into the future.

This paper is a culmination of those efforts and equips policy-makers and regulators with a clear roadmap for addressing the complexities of generative AI by examining existing regulatory gaps, the unique governance challenges of various stakeholders and the evolving forms of this technology. The outputs of this paper are designed to be practical and implementable, providing global policy-makers with the tools they need to enhance generative AI governance within their jurisdictions. Through this paper, our *AI Governance Alliance: Briefing Paper Series*, launched in January 2024, and our events and community meetings, we seek to create a tangible impact in AI literacy and knowledge dissemination.

Given the international context in which this technology operates, we advocate for a harmonized approach to generative AI governance that facilitates cooperation and interoperability. Such an approach is essential for addressing the global challenges posed by generative AI and for ensuring that its benefits are shared equitably, particularly with low-resource economies that stand to gain significantly from its responsible deployment.

We invite policy-makers, industry leaders, academics and civil society to join us in this endeavour. Together, we can shape a future where generative AI contributes positively to our world and ensures a prosperous, inclusive and sustainable future for all.

Executive summary

Governments should address regulatory gaps, engage multiple stakeholders in AI governance and prepare for future generative AI risks.

The rapid evolution and swift adoption of generative AI have prompted governments to keep pace and prepare for future developments and impacts. Policy-makers are considering how generative artificial intelligence (AI) can be used in the public interest, balancing economic and social opportunities while mitigating risks. To achieve this purpose, this paper provides a **comprehensive 360° governance framework**:

1. **Harness past**: Use existing regulations and address gaps introduced by generative AI.

   The effectiveness of national strategies for promoting AI innovation and responsible practices depends on the timely assessment of the regulatory levers at hand to tackle the unique challenges and opportunities presented by the technology. Prior to developing new AI regulations or authorities, governments should:

   - **Assess existing regulations** for tensions and gaps caused by generative AI, coordinating across the policy objectives of multiple regulatory instruments
   - **Clarify responsibility allocation** through legal and regulatory precedents and supplement efforts where gaps are found
   - **Evaluate existing regulatory authorities** for capacity to tackle generative AI challenges and consider the trade-offs for centralizing authority within a dedicated agency

2. **Build present**: Cultivate whole-of-society generative AI governance and cross-sector knowledge sharing.

   Government policy-makers and regulators cannot independently ensure the resilient governance of generative AI – additional stakeholder groups from across industry, civil society and academia are also needed. Governments must use a broader set of governance tools, beyond regulations, to:

   - **Address challenges** unique to each stakeholder group in contributing to whole-of-society generative AI governance
   - **Cultivate multistakeholder knowledge-sharing** and encourage interdisciplinary thinking
   - **Lead by example** by adopting responsible AI practices

3. **Plan future**: Incorporate preparedness and agility into generative AI governance and cultivate international cooperation.

   Generative AI's capabilities are evolving alongside other technologies. Governments need to develop national strategies that consider limited resources and global uncertainties, and that feature foresight mechanisms to adapt policies and regulations to technological advancements and emerging risks. This necessitates the following key actions:

   - **Targeted investments** for AI upskilling and recruitment in government
   - **Horizon scanning** of generative AI innovation and foreseeable risks associated with emerging capabilities, convergence with other technologies and interactions with humans
   - **Foresight exercises** to prepare for multiple possible futures
   - **Impact assessment and agile regulations** to prepare for the downstream effects of existing regulation and for future AI developments
   - **International cooperation** to align standards and risk taxonomies and facilitate the sharing of knowledge and infrastructure

Introduction

A 360° framework is needed for resilient generative AI governance, balancing innovation and risk across diverse jurisdictions.

As organizations and individuals consider how best to adopt generative artificial intelligence (AI), new powerful capabilities continue to emerge. For some, humanity's future with generative AI can feel full of promise, and for others, concern. Indeed, across industries and sectors, generative AI presents both opportunities and risks. For example – will generative AI enhance personalized treatment plans improving patients' health outcomes, or will it induce novel biosecurity risks? Will journalism be democratized through new storytelling tools, or will disinformation be scaled?

There is no single guaranteed future for generative AI. Rather, how society adapts to the technology will depend on the decisions humans make in researching, developing, deploying and exploiting its capabilities. Policy-makers, through effective governance, can help to ensure that generative AI facilitates economic opportunity and fair distribution of benefits, protects human rights, promotes greater equity and encourages sustainable practices. Governance decisions made now will shape the lives of present and future generations, how (and whether) this technology benefits society and who is left behind.

In response to the continued growth of the generative AI industry and rapid adoption of its applications across the world, this paper's **360° framework** outlines how to build resilient governance that facilitates AI innovation while mitigating risks, from the development stage to its use. The framework is designed to support policy-makers and regulators in the development of holistic and durable generative AI governance. The specific implementation of the framework, however, will differ between jurisdictions, depending on the national AI strategy, maturity of AI networks, economic and geopolitical contexts, individuals' expectations and social norms.

FIGURE 1 **A 360º approach for resilient policy and regulation**

Make use of existing regulations and address gaps caused by generative AI.

Encourage whole-of-society generative AI governance and cross-sector knowledge sharing.

Incorporate preparedness and agility in generative AI governance and facilitate international cooperation.

Pillar 1

Harness past

Greater clarity and certainty regarding existing regulatory environments is necessary to address emerging generative AI challenges and opportunities.

Successful implementation of national strategies for responsible and trustworthy governance of generative AI requires a timely assessment of existing regulatory capacity – among other governance tools – to tackle the unique opportunities and risks posed by the technology. This includes examination of the adequacy of existing legal instruments, laws and regulations, resolution of regulatory tensions and gaps, clarification of responsibility allocation among generative AI supply chain actors and evaluation of competent regulatory authorities' effectiveness and capacities. Such assessments must respect the fundamental rights and freedoms already codified in international human rights law, such as the protection of particular groups (e.g. minority rights[1] and children's rights[2]) as well as legal instruments that are domain-specific (e.g. to cybercrime[3] and climate change[4]).[5]

1.1 Examine existing regulations complicated by generative AI attributes

With increasing digitalization and a growing trend of monetizing personal and professional data, protection of privacy is both vital and complex. Policy-makers are looking to prioritize privacy-preserving considerations.

While generative AI's emerging properties and capabilities may warrant novel regulations, policy-makers and regulators should first examine their jurisdiction's existing regulations for addressing new challenges. They should also identify where existing regulations may be applied, adapted or foregone to facilitate the objectives of a national AI strategy. Navigating generative AI's interactions with existing regulations requires a nuanced understanding of both the technical aspects and the legal principles underlying the impacted regulations. Table 1 discusses examples of how regulatory instruments can be complicated in the context of generative AI.

Privacy and data protection

Generative AI models amplify privacy, safety and security risks due to their reliance on vast amounts of training data, powerful inference capability and susceptibility to unique adversarial attacks that can undermine digital trust.[6] A number of risks arise from the inclusion of personal, sensitive and confidential information in training datasets and user inputs, lack of transparency over the lawful basis for collecting and processing data, the ability of models to infer personal data and the potential for models to memorize and disclose portions of training data. With increasing digitalization and a growing trend of monetizing personal and professional data, protection of privacy is both vital and complex.

Policy-makers are looking to prioritize privacy-preserving considerations applicable to digital data while also creating affordances for data pooling that could lead to AI-facilitated breakthroughs.[7] Such affordances could be made to promote innovation for public goods in areas such as agriculture, health and education, or within narrowly specified exceptions for data consortia that facilitate the training of AI models to achieve public policy objectives.[8] Another emerging issue for policy-makers is that of ensuring generative AI safety and security, even when it may involve interaction with personal data, as in the case of investigating and responding to severe incidents. This could be addressed through the creation of regulatory exceptions and guardrails to ensure both privacy and responsible AI outcomes.

Copyright and intellectual property

Generative AI raises several issues relating to copyright infringement, plagiarism and intellectual property (IP) ownership (see Issue spotlight 1), some of which are currently being considered by courts in various jurisdictions. Rights related to protecting an individual's likeness, voice and other personal attributes are also implicated by the creation of "deepfakes" using generative AI. A blanket ruling on AI training is uncertain and judges could determine the fairness of certain data uses for specific products based on the product's features or outputs' frequency and similarity to training data.[9] Looking ahead, there is a pressing need for comprehensive examination of regulatory frameworks and for necessary guidance on documenting human creativity in the generation of content as a means of asserting IP protection.

ISSUE SPOTLIGHT 1

Training generative AI systems on copyright-protected data, and tensions with the text and data mining exception

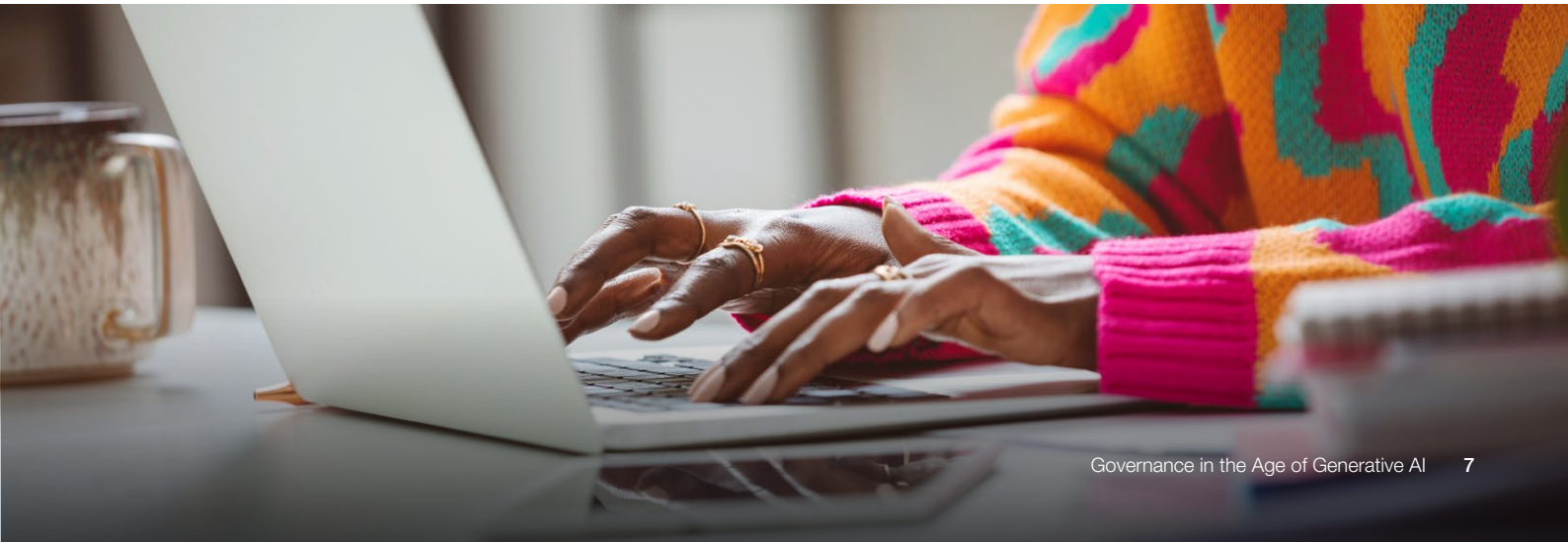
Text and data mining (TDM) is the automated process of digitally reproducing and analysing large quantities of data and information to identify patterns and discover research insights. Various jurisdictions around the world – such as Japan, Singapore, Estonia, Switzerland and the European Union (EU) – have introduced specific exemptions within their copyright laws to enable TDM extraction from copyright-protected content to innovate, advance science and create business value.

Given the vast amounts of data that generative AI systems use to train on and generate new content, jurisdictions should establish regulatory clarity regarding TDM for the purpose of generative AI training. This could be done, for example, by confirming whether AI development constitutes "fair dealing" or "fair use" (a key defence against copyright infringement) or falls within the exemptions recognized in some copyright laws. Countries like the UK are exploring such regulatory exceptions, seeking to promote a pro-innovation AI agenda.[10] Ultimately, there is mounting pressure on governments to resolve the copyright tension definitively.[11]

Licensing and data access on an "opt-in" or "opt-out" basis are also under examination to address TDM concerns, in addition to a range of technologies and standards that attempt to cede control to creators, allowing them to opt out from model trainers.[12] Licensing proponents argue that scraping for generative AI training without paying creators constitutes unlawful copying and is a form of reducing competition.[13] AI developers, however, argue that requirements to pay copyright owners for content used in training would constrain model development, negatively impact venture capital (VC) funding and reduce competition among generative AI models.[14] While they do not eliminate IP law concerns entirely, opt-in/out and licensing efforts could contribute to setting standards that generative AI foundation model providers would be expected to uphold.

Consumer protection and product liability

While AI-specific regulation remains voluntary or pending in jurisdictions outside of the EU, consumer regulation and product liability laws continue to be applicable, regardless of whether they strictly contemplate AI or other technologies. Generative AI has the potential to influence the consumer market by automating various tasks and services. This may, however, also challenge traditional approaches to risk assessment and mitigation (due to the technology's broad applicability and ability to continually learn and generate new and unique content), as well as product safety standards (for example, in health and physical safety). The development of standards should be an iterative, multidisciplinary process that keeps pace with technological advancements.



Competition

Market authorities must ensure that the competitive conditions driving the rapid pace of innovation continue to benefit consumers. Although existing competition laws remain applicable, generative AI raises new concerns related to the concentration of control over critical components of the technology and certain partnership arrangements. For example, generative AI's capabilities are enhanced with access to high-performance compute capacities and certain datasets that may prove critical for model development. The latter can depend on access to a vast number of users, contributing to economies of scale that challenge competition.[15] In response, competition authorities around the globe are starting to provide guidance on competition risks and expectations in generative AI markets.[16] Competition complexities at each layer of the AI stack will need to be evaluated as the technology evolves to enable access and choice across AI models, including general (e.g. ChatGPT), area-specific (e.g. models designed for healthcare) and personal use models. Such evaluations will also need to be considered alongside existing legislation relating to national security, freedom of expression, media and assembly.

TABLE 1 **Selection of complexities introduced by generative AI for existing regulatory areas**

Regulatory area	Emerging complexities (non-exhaustive)	Emerging strategies under consideration by regulators (non-exhaustive)
 Privacy and data protection	Legal basis for user data being used to train generative AI models	Enforcement of data-minimization principles[17] and opt-in/out rights by generative AI providers and deployers[18]
	Incidental collection of personal data by web-crawlers	Clarifying web terms-of-service agreements and encouraging privacy-enhancing technological measures such as the detection and redaction of personally identifiable information[19]
	Specifying purpose limitations for data collection	Guidance for purpose thresholds within domain-specific regulations, e.g. financial services[20]
	Online safety and protection of vulnerable groups, especially minors, from harmful outputs	Position statements highlighting expectations for safety measures and preferences for emerging best practices[21]
 Copyright and IP	Copyright infringement of training data	Clear policy positions and accumulation of legal precedents on the relations between copyright and generative AI[22]
	IP rights and ownership of works generated by AI	Guidance on assessing the protectable elements of AI-generated works[23]
	Attribution and fair compensation for artists and creators	Investments in solutions for attribution and author recognition such as watermarking and content provenance, along with privacy and data protection
	Extension of generative AI model training to additional data modalities (e.g. sensory, biological, motion)	Considerations of new IP challenges and classifications related to emerging data modalities
 Consumer protection and product liability	Liability obligations resulting from scope of multiple applicable regulations	Considerations around whether and in which cases a concern is covered by the existing regulations
	The lack of a specific purpose of the generative AI model before its implementation complicates liability arising from defectiveness and fault	Combining the conventional AI fault and defectiveness criteria with new methods designed for generative AI's technical nuances
	Efficacy of evidential disclosure requirements	Broadening the disclosure requirement to encourage transparency via explainability, traceability and auditability, and include systems that are not just classified as high-risk
 Competition	Business conduct or agreements that enable a dominant firm to exclude rivals	Initiating sectoral studies to develop a baseline understanding of the competitive dynamics of the AI technology stack, reviewing agreements between industry players and examining single firm conduct[24]
	Unfair or deceptive practice	Issuing guidance on unfair or deceptive practice prohibitions if it does not exist[25]
	Impact of downstream applications on competition across several sectors	Stakeholder consultations on how generative AI impacts competition in important markets, e.g. search engines, online advertising, cloud computing and semiconductors[26]

1.2 Resolve tensions between policy objectives of multiple regulatory regimes

Regulators must address emerging tensions and mitigate the risk of undermining legal certainty and respect for legitimate expectations.

The intersectional nature of generative AI technologies and the applicability of multiple regulatory instruments creates a complex environment where regulatory frameworks often overlap and conflict due to competing policy objectives. As technology evolves and becomes more widely adopted, regulators must address emerging tensions and mitigate the risk of undermining legal certainty and respect for legitimate expectations.

Addressing tensions between horizontal regulations

Multiple horizontal regulations, which aim to create broad, industry-agnostic standards, may conflict when they impose requirements that are difficult to reconcile across generative AI contexts or applications. For example, generative AI model developers may have trouble identifying the appropriate lawful basis for data processing and delivery according to data protection rights articulated through the EU's General Data Protection Regulation (GDPR). A similar tension emerges between copyright law – which protects the rights of creators and inventors ensuring that they can control and profit from their creations – and generative AI innovation, which often uses copyrighted material for training.

Addressing tensions between horizontal and vertical regulations

Horizontal regulations may also conflict with vertical regulations tailored to specific sectors. For instance, financial institutions using generative AI may encounter challenges balancing horizontal privacy regulations with financial sector know-your-client (KYC) procedures. Where data protection regulations require organizations to minimize personal data collection linked to a specific purpose, KYC guidelines require financial institutions to conduct thorough due diligence on clients to ensure compliance with anti-money-laundering laws.

1.3 Clarify expectations around responsibility allocation

As defined in the World Economic Forum's Digital Trust Framework,[27] maintaining accountability and oversight for trustworthy digital technologies requires clearly assigned and well-defined legal responsibilities alongside remedy provisions for upholding individual and social expectations. Generative AI introduces complexities into traditional responsibility allocation practices, as examined in Table 2. Policy-makers should consider where supplementary efforts are needed to address gaps and where legal and regulatory precedents can help to clarify generative AI responsibility. The issuance of effective guidance requires consideration of how liability within the generative AI supply chain can vary for different roles and actors as well as consideration of retroactive liabilities and dispute-resolution provisions. Unresolved ambiguity in responsibility allocation can limit investor confidence, create an uneven playing field for various supply chain actors and leave risks unaddressed and harms without redress.

TABLE 2 **Challenges and considerations for generative AI responsibility allocation (non-exhaustive)**

	Example challenges	Considerations for policy-makers
 **Variability**	– Model variations include features (e.g. size), scope (e.g. use purpose), and method of development (e.g. open-to-closed source). – Technical approaches to layering and fine-tuning continuously evolve, enabling general-purpose models to adapt functionality for specific applications. – Entity categorization complexities involve multiple actors from different sectors with overlapping or multiple roles.	– Case-based review: Policy-makers should provide general allocation guidance to cultivate predictability, but include mechanisms that allow case complexities to determine precise allocation. Requiring actors to identify responsibility hand-offs is one approach being examined by jurisdictions. – Terminology: Policy-makers should collaborate to arrive at shared terms for models, applications and roles, e.g. in line with *ISO 420001* from the International Organization for Standardization (ISO). – Regulatory carve-outs: Policy-makers should limit instances when use can lead to unfair advantages, such as where some entities are able to bypass crucial safeguards and accountability measures or engage in regulatory arbitrage.
 **Disparity between actors**	– Single points of failures and power concentration occur as a result of a few foundational models (serving many applications and billions of end users). – Disparities in influence emerge between upstream and downstream actors. – There is limited transparency for downstream actors related to training data and for upstream actors related to end-user activity.	– Proportionality: Policy-makers should consider the control, influence and resources each actor has in the generative AI life cycle, and ability to redress issues resulting in harm. – Third-party certifications: Policy-makers should consider appropriateness and necessity of using third parties for a robust AI certification system (potentially defined through regulation) that enables actors to verify and trust each other's capabilities.
 **Complexity of review**	– Interpretability difficulties relating to outputs arise due to models often operating as "black boxes" to varying degrees. – Traceability difficulties transpire in 1) diversity of data sources, 2) sequence of events that led to a fault, 3) determining whose negligence or malice induced the fault or made the fault more likely. – Physical inspection or verification of changes to generative AI products in the market has limited feasibility.	– Documentation: Policy-makers should incentivize appropriate transparency and vulnerabilities disclosure upstream and downstream to enable responsible decisions. Concerns about trade secrets or data privacy compromise the need to be mitigated. – Traceability mechanisms: Policy-makers should require the ability to trace outputs back to their origins while considering compromise and mitigation measures for IP and data privacy concerns. – Continuous compliance: Policy-makers should integrate standards for market entry and procedures for post-approval changes, and encourage industry review boards and ongoing independent audits.[28]

1.4 Evaluate existing regulatory authority capacity for effective enforcement

Effective regulatory enforcement depends on governments identifying the appropriate authority or authorities and enabling their activity with adequate resources.

> Some argue that a central AI agency is needed to address highly capable AI foundation models. Others consider a central AI agency more prone to regulatory capture and less effective for AI's diverse use cases.

Expansion of existing regulatory authority competencies

While generative AI may elicit consideration of a new AI-focused authority, governments should first assess opportunities to make use of existing regulatory authorities with unique domain knowledge and ensure they can translate high-level AI principles to sector-specific applications. Considerations of how to delegate regulatory authority for AI will depend on a jurisdiction's AI strategy, resources and existing authorities. For example, countries that have a data protection authority (DPA), such as France, tend to rely on the DPA to comprehensively address AI, since data is fundamental to AI models and uses. In the same vein, countries without DPAs, such as the US, may lack a readily apparent existing authority.

Furthermore, the specific mandate and procedural frameworks of existing authorities such as DPAs impact AI governance. For example, Singapore's DPA, the Personal Data Protection Commission (PDPC), sits within a broader authority, the Infocomm Media Development Authority (IMDA), whose mission includes cultivating public trust alongside economic development. Thus, AI governance from Singapore's DPA actively considers both trust and innovation within its regulations. This underscores how generative AI may necessitate the expansion of remits for existing regulators. For example, Singapore's IMDA must now consider issues related to generative AI data ownership and provenance, and the use of data for model training, including potential compensation for creators whose content was trained on.

Coordination of regulatory authorities

Coordination between regulatory authorities can prevent duplication of efforts and enhance operational resilience for overburdened and under-resourced offices. New coordination roles or responsibilities should be considered. For example, the UK has created the Digital Regulation Cooperation Forum (DRCF), encompassing the Competition and Markets Authority (CMA), Financial Conduct Authority (FCA), Information Commissioner's Office (ICO) and Office of Communications (Ofcom) to ensure greater cooperation between regulators on online matters, including within the context of AI. Similarly, Australia's Digital Platform Regulators Forum (DP-REG) – an information-sharing and collaboration initiative between independent regulators – considers how competition, consumer protection, privacy, online safety and data issues intersect.

Dedicated AI agency versus distributed authority between sector-specific regulators

The founding of an AI agency requires careful consideration regarding, for instance, the scope of responsibilities, availability of resources and domain-specific regulatory expertise. For example, would the agency serve to coordinate, advise and upskill sector-specific regulators on AI matters, likely requiring less funding, or would it serve as an AI regulatory authority with enforcement powers, requiring greater funding? Some argue that a central AI agency is needed to address highly capable AI foundation models.[29] Others consider a central AI agency more prone to regulatory capture and less effective for AI's diverse use cases than distributed regulations among existing sector-specific authorities with domain-specific knowledge. Consequently, many would prefer a council-like AI body that coordinates and advises existing sector-specific authorities.[30]

Jurisdictions are finding creative ways to navigate limited funding and political compromise. For example, the EU embedded its new AI Office within the EU Commission,[31] instead of setting it up as a solitary institution, to amplify the effectiveness of the office's limited number of staff. Like the EU, jurisdictions are navigating complex challenges of how to creatively resource a new AI body or authority while ensuring its independence. Still, enforcement of the AI Act, like GDPR, may strain authorities at the member-state level. For instance, while Spain has set up a centralized authority to enforce the act's provisions, France may use existing regulators, such as the DPA, as the authority of record.

Pillar 2

Build present

Governments should address diverse stakeholder challenges to facilitate whole-of-society governance of generative AI and cross-sector knowledge sharing.

> Governments are carefully considering how to avoid over- and under-regulation to cultivate a thriving and responsible AI network, where AI is harnessed to address critical social and environmental challenges.

While regulators play a critical role, they cannot independently ensure the resilient governance of a technology that has simultaneously broad and diversified impacts, and capabilities that continue to evolve. Other stakeholder groups hold key puzzle pieces to assembling resilient governance and a responsible AI system, for example:

- **Industry**: With proximity to the technology, its developers and users, industry is at the front line of ensuring that generative AI is responsibly governed across countless use cases within commercial applications and public services.
- **Civil society organizations (CSOs)**: With expertise on how generative AI uniquely impacts the different communities and issue spaces they represent, CSOs enable informed and holistic policy-making.
- **Academia**: Through rigorous and independent research and educational initiatives, academia is critical to shaping responsible AI development and deployment and ensuring public literacy on responsible use.

Governments must use a broader set of governance tools, beyond regulations, to:

- Address the unique challenges of each stakeholder group in contributing to society-wide generative AI governance
- Facilitate multistakeholder knowledge-sharing and encourage interdisciplinary thinking
- Lead by example by adopting responsible AI practices

2.1 Address challenges of stakeholder groups

Enable responsible AI implementation by industry

Governments are carefully considering how to avoid over- and under-regulation to cultivate a thriving and responsible AI network, where AI developed for economic purposes includes robust risk management, and AI research and development (R&D) is harnessed to address critical social and environmental challenges. Since market-driven objectives may not always align with public interest outcomes, governments can encourage robust and sustained responsible AI practices through a combination of financial mechanisms and resources, clarified policies and regulations, and interventions tailored to industry complexity.

Incentivize proactive, responsible AI adoption by the private sector

Public policy-making processes often lack the private sector's ability to adopt governance protocols for innovative technologies. To address this, governments should assess the applicability of existing AI governance frameworks (e.g. *Presidio AI Framework*,[32] *NIST AI Risk Management Framework*[33]) and encourage proactive industry adoption. In addition to educating industry on frameworks, governments can cultivate an environment where industry is incentivized to proactively invest in responsible AI. Potential strategies include:

- **Financial incentives**: Governments could introduce inducements for responsible AI practices such as tax incentives, grants or subsidies for R&D, talent or training. Policy-makers could consider potential tax rate adjustments to incentivize AI designed to augment (rather than replace) human labour,[34] and carefully consider trade-offs of proposed adjustments.
- **Sustained funding**: Government leaders should ensure investment in both short- and long-term R&D to reach breakthroughs on complex AI innovations and address responsible AI challenges. Jurisdictions with a less advanced AI industry may require greater initial government investment to incentivize VC funding.

> A responsible AI system is of strategic importance to investors for mitigating regulatory and non-regulatory risks (e.g. cyberattacks), and improving top- and bottom-line growth.

- **Procurement power**: Governments should explore preferred procurement measures for AI with demonstrable responsible AI metrics.
- **Access**: Governments should provide opportunities for public-private partnerships and access to public datasets for AI developed with demonstrable responsible AI metrics or that's designed for social or environmental benefit.
- **Responsible AI R&D and training**: Leaders should examine the suitability of requiring a percentage of R&D expenditure for responsible AI governance and/or training for organizations.

Clarify policies and enable measurement

A responsible AI system is of strategic importance to investors for mitigating regulatory and non-regulatory risks (e.g. cyberattacks), and improving top- and bottom-line growth.[35] Over the last decade, investors have helped drive industry investment in environmental issues, and they can play a similar role in incentivizing responsible AI practices, for instance by addressing AI's vast energy use.[36] However, uncertainty in how government AI policies will be implemented and enforced prevents confident investing in responsible AI practices.

Governments should set clear national priorities and policies on responsible AI, reduce ambiguity in existing regulations and provide signals on the trajectory of regulations. Singapore's PDPC, for example, proactively shared advisory guidelines[37] that clarify the application of existing data laws to AI recommendation and decision systems. The guidelines additionally highlight exceptions, with the aim of helping industry navigate regulation.

Encourage businesses to test, evaluate and implement transparency measures, including through:

- **Clear frameworks**: to measure risks as well as social and human rights and environmental impacts
- **Certifications**: to clarify responsible AI practices and testing are satisfactory and draw investors and the public
- **Sandboxes**: to experiment and refine before wider deployment, with incentivized participation
- **Knowledge-sharing**: to promote sharing of benchmarking, e.g. *Stanford AI Index Report*[38]
- **Competitions**: to address complex AI challenges, e.g. National Institute of Standards and Technology (NIST) generative AI challenge[39]
- **Technical standards**: to establish common methodologies and benchmarks for evaluating AI system performance, safety and ethical compliance across different domains and applications. e.g. *ISO 42001*[40]

Tailor interventions to diverse industry needs

Policy-makers need to consider the diversity of AI governance challenges faced by industry stakeholders to identify meaningful points of intervention. Table 3 illustrates how business size can determine the resources available to implement responsible AI governance and the compliance complexities encountered. Other governance challenges can result from industry stakeholder characteristics, such as sector, location, industry maturity, risk sensitivity and role in the AI supply chain.

TABLE 3 | **Governance challenges by business size (non-exhaustive)**

	Challenges	Considerations for policy-makers
Large businesses	**Implementation**: Difficulties may occur for AI governance operationalization and compliance within complex or differently structured organizations.	Policy-makers should provide implementation guidance that builds upon current risk management frameworks, global standards, benchmarks and baselines.
	Competition: Competitors may not invest equally in responsible AI practices.	Policy-makers could review responsible AI practices and regulatory compliance across stakeholders.
	Clarity: Navigating compliance ambiguities or complexities across sectors and between jurisdictions may present challenges.	Where possible, policy-makers can provide guidance on what actions are within or outside regulations, reduce overlap and facilitate interoperability through harmonization.
Small- and medium-sized enterprises (SMEs) and start-ups	**Resources**: Resources to develop and demonstrate robust responsible AI practices to regulators, investors or partners may be limited.	Policy-makers should provide guidance, training and consultation access on AI governance, facilitate insight-sharing between large businesses and SMEs, and use certification mechanisms.
	Applicability: AI governance frameworks and recommendations lack applicability or specificity to the realities of SME operations.	Policy-makers should include input from diverse SMEs in development of national and international governance frameworks.
	Prioritization: Fast pace of start-ups and lack of capital can lead to prioritizing innovation over risk assessment.	Policy-makers can incorporate responsible AI practices and regulatory landscapes into curricula of start-up accelerators and incentivize participation in sandboxes.

Enable leading AI research and education by academia

Through research and education, academia is a critical stakeholder in cultivating a robust AI network. Until the early 2000s, leading AI R&D was primarily conducted within academia. It contributed to providing open-source knowledge that accelerated innovation and optimized development costs. With recognition of the economic potential of AI, investment has since shifted R&D to industry.

Without academia at the forefront of AI R&D, key risks emerge:

- Homogenization of the AI network
- Decline in discoveries that emerge from academia's interdisciplinary research settings
- Decreased independent research around AI ethics, safety and oversight
- Diminished general workforce training
- Barriers to cross-institution collaboration
- Reduced ability to wield academic freedom to challenge prevailing consensus
- Broken AI talent pipeline

Since generative AI has extensive and costly infrastructural needs (e.g. compute capabilities, data), academia's ability to conduct leading research is severely limited.[41] Table 4 outlines the range of challenges facing academic stakeholders that policy-makers should address to cultivate a thriving AI system. These challenges must be considered in the context of the different operating conditions of academic institutions, for example of private, public and community colleges, to ensure equitable access to AI literacy, benefit from the AI economy and a diverse pipeline of responsible AI experts. Similarly, policy-makers should address the unique literacy and access challenges in earlier educational settings, for example in primary and secondary schools.

TABLE 4 **Challenges and considerations for policy-makers to support academic stakeholder groups (non-exhaustive)**

Challenges	Considerations for policy-makers		
	For academic institutions	For researches	For educators
 Appropriate use	Clarify compliance with evolving relevant regulations (e.g. AI, data privacy and copyright) and simplify regulations to enable research in responsible AI.	Provide guidance on responsible generative AI use in research[42] (e.g. data analysis) and training on risks to boost cognizance when conducting research in the age of generative AI (e.g. of potential misuse by respondents to online studies).	Provide guidance on responsible generative AI use by teachers (e.g. essay review and feedback) and students (e.g. critical evaluation of generative AI outputs in essay writing).
 Resources	Ensure access to physical and digital infrastructure needed for faculty, researchers and students to become familiar with AI and use it responsibly.[43]	Provide access to data and compute capabilities to conduct leading AI and generative AI research, and clarify guidelines for accessing public sector data while maintaining privacy.	Provide regularly updated training materials and ensure that educators, regardless of institutional prestige, can keep pace with AI advancements.
 Funding	Close pay gaps between industry and academia to reduce AI brain drain to industry.	Allocate research grants into responsible AI challenges (e.g. hallucinations, bias) that do not require cut-throat competition or complex applications.	Allocate funding for courses on AI and responsible AI.

Ensure access and participation of CSOs

In addition to ensuring technical expertise in governance conversations, there is a critical need for expertise related to the social impacts of AI and generative AI, informed by the lived experiences of those interacting with the technology. CSOs play a key role in representing various citizen groups, individuals and issue spaces and provide related technical and societal expertise. CSOs can also offer independent oversight, holding governments and companies accountable for their AI implementation.

Depending on their missions, CSOs have unique expertise around generative AI implications that policy-makers should make use of, for example:

- Labour protection groups can help inform the skills and training needed to ensure generative AI leads to job growth rather than displacement.
- Environmental groups can provide guidance on ways AI can help address local and global climate challenges, and considerations regarding generative AI's vast energy consumption.
- CSOs focused on creative practice, journalism, mis/disinformation or election monitoring can inform the harnessing of generative AI's creative potential while preserving information integrity and ownership rights.
- CSOs serving marginalized populations or protected classes can help ensure AI policies and technologies holistically consider the varied opportunities and risks posed (see Issue spotlight 2).

ISSUE SPOTLIGHT 2

A child-centric approach for generative AI governance

As the largest digital user group and fastest adopters of technology, children and youth are at the forefront of AI-enabled systems. The effects of using generative AI, both positive and negative, will have wide-ranging and lifelong impacts that will shape the development, safety and worldviews of children.[44] Research agendas are beginning to emerge to aid precise policies around the disproportionate effect of algorithmic bias on minoritized or marginalized children.[45] They can additionally inform policies that address concerns around how generative AI training[46] and use[47] could amplify child sexual abuse material (CSAM),[48] and how generative AI applications, especially the use of chatbots and smart toys, may affect cognitive functioning among children.[49] Existing resources such as UNICEF's *Policy Guidance on AI for Children*,[50] the European Commission's guidance on *Artificial Intelligence and the Rights of the Child*[51] and the World Economic Forum's *AI for Children Toolkit*,[52] provide valuable direction.

Given their limited political agency, economic influence and organizing power, children can often be overlooked in technology governance considerations, even as they are most impacted. Further, the existence of inequalities around the digital divide exacerbates the risks and harmful effects of generative AI for some children more than others, given their inability to participate in shaping generative AI's development or access its benefits. Engaging young users, their guardians and local communities in a meaningful and ongoing way throughout the life cycle of generative AI projects and governance, directly and via CSOs with deep technical or policy expertise in these areas, is vital for children's empowerment and the development of responsible AI innovation. Transparency in how children's rights and input have been considered and implemented is critical to promoting public trust and accountability.[53]

CSOs face significant access and participation challenges preventing them from assessing societal impacts of generative AI technologies, informing governance policies and supply chain accountability, and advocating for the rights of citizen groups and vulnerable populations such as children, as examined in Table 5.

TABLE 5 | **Challenges and considerations for policy-makers to support CSOs (non-exhaustive)**

	Challenges	Considerations for policy-makers
Access	**Under-resourced**: There is a lack of adequate tools and skills to review impacts of generative AI.	– Policy-makers should provide access and training for cutting-edge tools and incentivize industry to share tools. – They should fund R&D to improve tools' abilities, (e.g. detection in minority languages or compressed media). – They should provide funding for CSOs to undertake independent impact assessments.
	Opaque: There are limited metrics on how companies have implemented responsible AI, including principles that have been publicly committed to.[54]	– Policy-makers should standardize and incentivize responsible AI reporting. – They should provide CSOs with easier access to mandated transparency data, (e.g. via EU Digital Services Act and AI Act).
	Limited information: There is a lack of access to training data and weights, and information on how companies moderate public use of AI technologies.	– Policy-makers should incentivize industry to share data with CSOs, while preserving privacy and IP. – They should standardize transparency reporting on how AI companies moderate technology use.
Participation	**Disempowered**: CSO inclusion is often limited in numbers and in influence on decision-making. There is even less inclusion of CSOs operating outside regulatory regimes, which will be impacted by generative AI and regulatory shifts.	– Policy-makers should ensure sectoral parity in discussions. – They should educate on the value of CSO community-driven insights. – They should strengthen outreach to vulnerable communities and relevant CSOs, including transnational CSOs, and engage international CSO forums, (e.g. C7, C20, African Union Civil Society Division).
	Delayed: CSOs engaged late in technical and governance processes.	– Policy-makers should ensure task forces, institutes etc. have CSO participation at formation.

2.2 Facilitate multistakeholder knowledge-sharing and interdisciplinary efforts

Governments should facilitate knowledge-sharing across stakeholder groups and with other governments to reduce duplicative efforts, offset expertise gaps and enable informed policies capable of addressing emerging, nuanced and wide-reaching generative AI challenges.

Ensure conditions for knowledge-sharing feedback loops

Knowledge-sharing requires nurturing of feedback loop conditions and proactive examination of challenges to those conditions that may prevent stakeholders from meaningfully participating, as described in Figure 2 and Table 6.

FIGURE 2 **Feedback loop conditions for effective multistakeholder participation**

TABLE 6 | **Challenges impacting feedback loop conditions (non-exhaustive)**

	Stakeholder challenges	Considerations for policy-makers
Trustworthy 	Industry may be wary of sharing models openly for fear of divulging trade secrets or exposure to legal liabilities.	– Policy-makers should provide safe harbour provisions and ensure discretion. – To ensure mutual benefit, all participants should be willing to share insights while preventing privileged access.
Communicative 	CSOs (that are more fluent in social impacts), industry (more fluent in technology) and government (more fluent in policy) may have difficulty understanding each other. Further complicating the issue, CSOs may often examine topics through the lens of human rights, whereas industry does so through risks.	– Policy-makers should use professional facilitators, invest in structured support for participation across sociotechnical conversations and increase incorporation of rights protections in frameworks (including in risk-based frameworks).
Representative 	Broad participation of actors is needed but can be difficult to coordinate, and its inputs can be hard to synthesize.	– Policy-makers could layer broad input models (e.g. written input) over narrow models (e.g. roundtable). – They could set ample time for input review and synthesis.
Independent 	The public may be concerned about regulatory capture or undue influence in boards or research partnerships.	– Policy-makers could set term limits for participation in boards. – They could make disclosure of extent of industry participation in research collaborations a requirement.
Consistent 	Sporadic touchpoints can leave non-industry participants playing catch-up on technological advances, and cause non-government participants to lag behind on policy changes.	– Policy-makers should align on frequency expectations and coordinate multiple feedback loops.
Transparent 	Participants and the public may be concerned that some stakeholders yield greater influence.	– Policy-makers could include equitable sectoral representation and provide transparency on feedback review processes with strengthened whistleblower protections.

Governments will need to coordinate multiple feedback models simultaneously to build holistic knowledge-sharing across issues and timelines (e.g. timing of AI model releases and legislative calendars), and to account for long-standing and emerging issues. Layering models is also necessary to address limited resources. For example, calls for inputs, which enable insights from numerous stakeholders, can require substantial resources to meaningfully review. Governments may consider combining routine calls for input with more narrow feedback mechanisms, such as advisory boards. The boards themselves may conduct interviews and roundtables to broaden representation of the insights they share with policy-makers.

In designing feedback loops, policy-makers should also consider how non-government stakeholders have limited resources. It is also crucial to explore how to simplify participation by, for instance, reducing unnecessary complexities in calls-for-input forms or merging similar calls for input from different agencies to reduce time requirements from participants.

Encourage interdisciplinary innovation

Generative AI innovation is built upon interdisciplinary research. For example, the development of ImageNet, a database that proved the importance of big data in training, emerged from the cross-pollination of ideas from linguistics, psychology, computer science and adjacent fields.[55] Despite the importance of interdisciplinary collaboration to generative AI innovation and addressing generative AI's sociotechnical challenges, industry and academia do not sufficiently cultivate environments that support this approach. Within private-sector tech companies, social scientists and humanities experts are often a fraction of the team. Despite maintained multi-disciplinary faculties within academic institutions, there are strong incentives for researchers to publish within discipline-specific journals, consequently encouraging isolated research. Policy-makers should consider levers to address these challenges, such as targeted academic research grants with interdisciplinary requirements or financial subsidies for interdisciplinary industry R&D.

Lead by example with responsible AI in public initiatives

Making use of AI, including generative AI, may improve governments' productivity, responsiveness and accountability.[56] However, its adoption requires responsible design, development, deployment and use, given its impact on individuals and society. Setting an example of responsible AI practices in government (including responsible procurement and acquisitions) could help to establish responsible AI norms[57] and secure the participation of industry, academia and civil society in creating a robust, responsible AI network. The City Algorithm Register, adopted across several cities in Europe, enables citizens to review algorithms employed by government agencies in public services, enhancing public oversight.[58] Jurisdictions such as Australia[59] and the US[60] have published internal policies for government AI practices aimed at advancing responsible innovation and managing risks.

Pillar 3

Plan future

Generative AI governance demands preparedness, agility and international cooperation to address evolving sociotechnical impacts and global challenges.

Generative AI's capabilities are rapidly evolving alongside other technologies and interacting with changing market forces, user behaviour and geopolitical dynamics. Bringing ongoing clarity to generative AI's changing short- and long-term uncertainties is critical for effective governance.

TABLE 7 **Government challenges and actions to keep pace with generative AI**

Compounding challenges	Strategic actions
Limited resources and expertise: Governments may struggle to prioritize investment in building state-of-the-art AI and generative AI expertise compared to other pressing needs.	**Targeted investments and upskilling**: Governments should be deliberate with limited resources in upskilling and hiring.
Rapid evolution: Governments may lack sufficient proximity to, and awareness of, generative AI evolution and adoption to effectively approximate sociotechnical impacts.	**Horizon scanning**: Governments should monitor emerging and converging generative AI capabilities and evolving interactions with society.
Uncertain futures: Technology, society and geopolitical uncertainties are outpacing traditional upskilling practices and policy development cycles.	**Strategic foresight**: Governments should ensure resilience though exercises that inform anticipatory policy.
Slow mechanisms: Government decision-making can be slow by design (e.g. due to separation of powers and oversight) or complicated by administrative procedures.	**Impact assessments and agile regulations**: Governments should prepare for the downstream effects of regulation and introduce agile dynamics into decision-making processes.
Global fragmentation: Limited resource-sharing and segregated jurisdictional governance activity can paralyse domestic investment and policy, and create non-interoperable international markets.	**International cooperation**: Governments should drive collective action to keep pace with generative AI innovation through harmonized standards and risk definitions, and sharing of knowledge and infrastructure.

3.1 Targeted investments and upskilling

- **Training on use**: Ensure officials who use generative AI technologies are trained in their varied capabilities and limitations.
- **Training on procurement**: Ensure officials who work with vendors are equipped to assess and test the AI capabilities of a product.
- **Adaptive upskilling**: Collaborate with industry and academia on adaptive upskilling of government in AI and foundational digital literacy.
- **Strategic hiring**: Recruit specialists for positions identified with amplified impact and, with limited resources, consider prioritizing sectors and use cases, for example, based on risk or domestic economic factors.
- **Hiring vs upskilling**: Consider how to appropriately balance hiring AI experts with AI upskilling of sector-specific experts (e.g. in agriculture and health).
- **AI body**: Carefully consider the need and scope of an AI-specific body or authority (see "Expansion of existing regulatory authority competencies" under section 1.4).
- **Guidance**: Examine where frameworks can be applied across sectors and where investment is needed for sector-specific guidance.

3.2 Horizon scanning

To anticipate and navigate novel risks and challenges posed by frontier generative AI, governance frameworks must continuously examine the horizon of generative AI innovation, including:

- **Emergence** of new generative AI capabilities
- **Convergence** of generative AI with other technologies
- **Interactions** with generative AI technologies

Documented, planned or forecasted emergence, convergence and interaction patterns can yield new waves of economic opportunities and novel approaches to addressing social and environmental challenges. Ongoing monitoring of opportunities and risks is critical to steering generative AI towards being a technology that benefits society. Multistakeholder knowledge-sharing (see Table 8) can enable informed horizon scanning.

Policy-makers should collaborate with industry to provide guidance on where disclosure of identified risks is needed and support oversight mechanisms to ensure compliance.

Emergence

As developers scale up generative AI models, the latter may exhibit qualitative changes in capabilities that do not present in smaller models. Such unexpected capabilities may include potentially risk-inducing abilities such as adaptive persuasion strategies, "power-seeking behaviours" to accrue resources and authority, and autonomous replication, adaptation and long-term planning capabilities. These emergent model properties must inspire appropriate governance benchmarks to effectively address unpredictable powers and potential pitfalls.

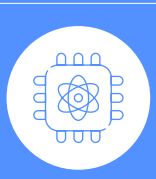
TABLE 8 **Generative AI emergent capabilities (non-exhaustive)**

Category	Example use	Example risks	Considerations for policy-makers
Multimodal generative AI Systems that synthesize and generate outputs across diverse data types and sensory inputs	Data analysed from radars, cameras, light detection and ranging (LiDAR), sensors and global positioning systems (GPS) in a safety-critical system (like a self-driving vehicle) to predict the behaviour of surrounding vehicles and pedestrians more accurately[61]	– Compounded data manipulation across input types – Amplification of potential flaws, biases and vulnerabilities – Novel systemic failures – Exacerbated societal disparities – Scaled and difficult-to-detect mis/disinformation – Novel persuasion techniques	– Focus on data integrity and secure-by-design frameworks, model architecture disclosures, responsible system design and impact assessment in public sectors – Examine readiness of existing policies and, if necessary, amend to address emerging privacy, security, safety, fairness, and IP rights and accountability
Multi-agent generative AI AI systems involving multiple agents that autonomously pursue complex goals with minimal supervision	Swarms of drones deployed for military and security purposes[62]	– Increased unpredictability and control complexity – Added accountability complexity – Challenges to traditional scenario planning and risk management – Potential for cascading failures – Novel adversarial attacks	– Develop guidelines for design and testing focused on robustness, security, safety, transparency, traceability and explainability – Establish accountability frameworks
Embodied generative AI AI systems embodied within physical entities such as robotics and devices capable of interacting with the real world	General-purpose humanoid robot with neural network-powered manual dexterity and ChatGPT 4's visual and language intelligence[63]	– Physical safety risks from control system failures – Security issues from malicious use of such systems – Novel physical manifestations of hallucinations	– Implement safety standards and security benchmarks – Encourage voluntary industry reviews and supplement with certification and audit practices, where appropriate

Convergence

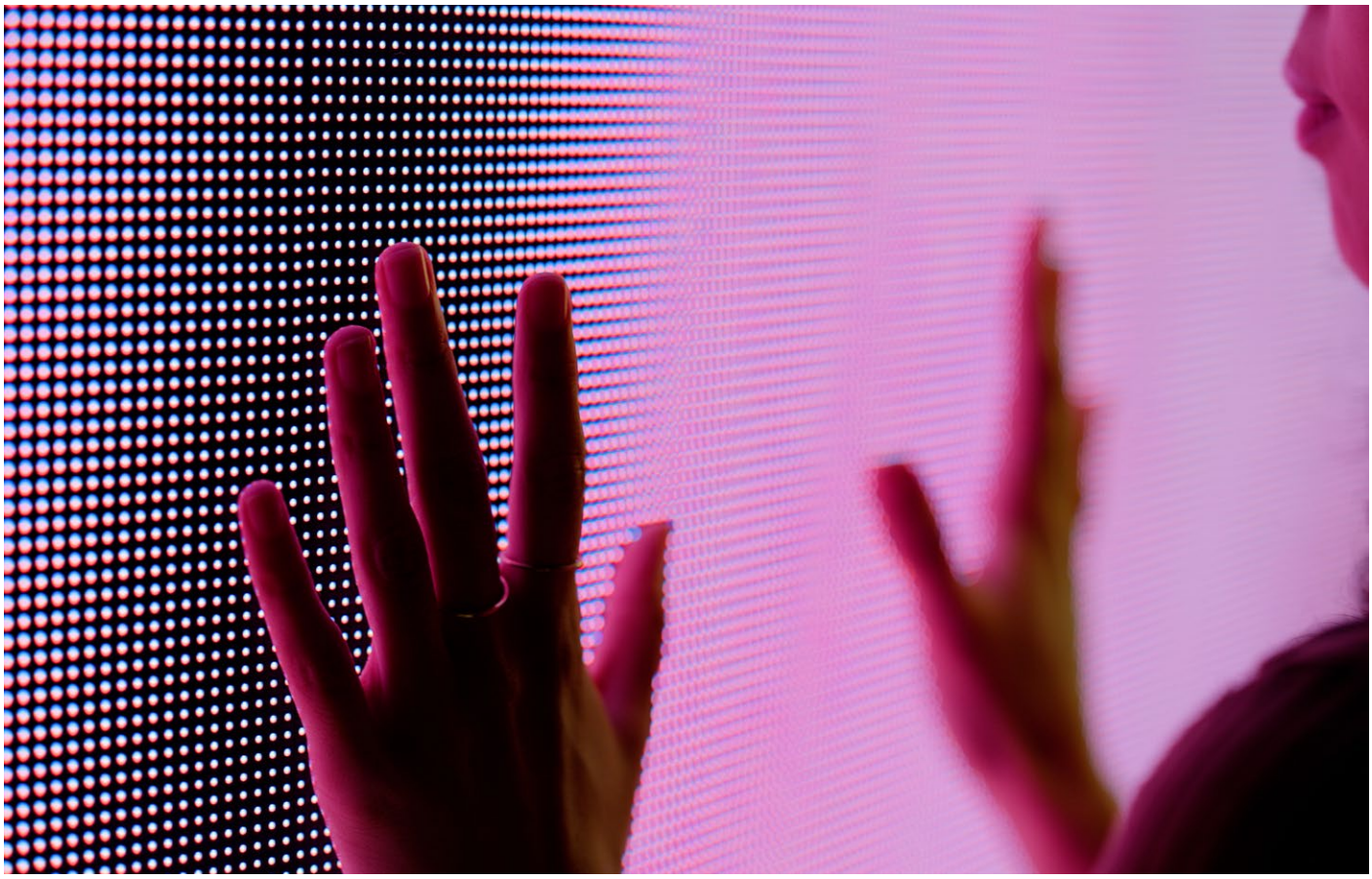
As a powerful general-purpose technology, generative AI can amplify other technologies, old and new, exposing complex governance challenges. For example, social media is under scrutiny due to its potential to distribute harmful AI-generated deepfakes,[64] such as non-consensual pornography[65] – including CSAM[66] – and election disinformation.[67] Looking ahead, the convergence of generative AI with advanced technologies can pose unprecedented opportunities and risks, as both the technologies and their governance frameworks are in the early stages.

TABLE 9 **Generative AI convergence with advanced technologies (non-exhaustive)**

Category	Example uses	Example risks	Considerations for policy-makers
 Synthetic biology	Generative AI is increasingly used in developing artificial analogues of natural processes, e.g. generation of genome sequences and cellular images, and simulations of genes and proteins. It is also used in building "virtual labs"[68] that can mitigate space and hazardous waste of real-world experimentations.	– Unintended ecological consequences – Gain-of-function research giving naturally occurring diseases new symptoms or capabilities like resiliency to medical treatments – Biosecurity risks and biological warfare – Novel ethical implications	– Robust bioethical frameworks – Tracking of the building and operation of various high-security disease labs globally – Restrictions on high-risk research – Strict containment protocols – International collaboration on safety standards – Refocusing of existing biological control laws
 Neurotechnology	Progress in generative AI, neuroscience and the development of brain-computer interfaces offers potential for increasing scientific discoveries, enabling paralysed individuals with communication, as well as addressing the burden of neurological disease and mental illnesses such as attention deficit hyperactivity disorder (ADHD), post-traumatic stress disorder (PTSD) and severe depression.	– Intentional abuse – Use in lethal autonomous weapon systems – Cognitive enhancement by brain-computer interfaces can amplify existing inequities – Behaviour modification and manipulation – Enfeeblement	– Review of privacy approaches that consider cognitive freedom, liberty and autonomy, and the establishment of new digital rights, if necessary – Establishment of assessment standards for model or neuroscientific accounts of disease on individuals, communities and society – Internationally harmonized ethical standards for biological material and data collection – Examination of moral significance of neural systems under development in neuroscience research laboratories – Context specification for neuroscientific technology use and deployment
 Quantum computing	Through optimizing code, generative AI may improve the design of hardware and quantum computing circuits, which are intended to solve problems too complex for classical computing. Quantum computing may accelerate generative AI training and inference and optimize parameter exploration.	– Advanced models beyond human comprehension – Impact on the environment due to increased energy and resource demands	– Review of legal provisions for controlled innovation that balance pace and safety without hindering progress – Incentivization of sustainable practices and energy-efficient technologies – Consideration of measures such as investing in research to strengthen the security and privacy of these systems

Interactions

Today, the integration of generative AI technologies into personal AI virtual assistants and companions raises new challenges that emerge from human interaction with and emotional reliance on these technologies. This issue highlights the need for responsible implementation, privacy, data protection and ethical human-AI interaction. For example, rapid advances and interactions with generative AI-enabled neurotechnology could become mainstream for many children, largely as consumer electronic devices that are not subject to rigorous oversight in clinical settings. The advancement and proliferation of voice chatbots, often with female-presenting voices, raise concerns about reinforced gender biases and stereotyping. Responsible and ethical development and regulation of these technologies, grounded in human rights, must therefore be an area of attention across stakeholder groups.



“ As generative AI applications become more complex and computationally powerful, the risk of emotional reliance between humans and generative AI applications tends to increase.

Emotional entanglement

Emotional AI aims to recognize, interpret and respond to human emotions, potentially improving human-computer interactions. As generative AI applications become more complex and computationally powerful, the risk of emotional reliance between humans and generative AI applications tends to increase.[69] Risks include dependency, privacy issues, coercion or manipulation leading to safety or psychological risks.[70] Such issues are exemplified by cases of users claiming that AI companies are interfering with their romantic relationships with chatbots.[71] The gravity of these phenomena is already evident in society, as seen in the case of a man who reportedly “ended his life following a six-week-long conversation about the climate crisis with an AI chatbot”.[72] Careful consideration of the ethical implications by policy-makers and legislators to ensure responsible AI use will be necessary.[73]

Synthetic data feedback loops

Human-created content scraped from the internet has been crucial in the training of large-scale machine learning, but this reliance is at risk due to the increasing prevalence of synthetic data generated by AI models.[74] Training models with synthetic data could lead to “model collapse”, where the quality of the generated content degrades over successive iterations, causing the performance of the models to deteriorate.[75] Policy-makers, in collaboration with industry, academia and CSOs, will need to consider how to stabilize these systems with human feedback, preserve human-created knowledge systems and incentivize the production and curation of high-quality data. Such considerations will need to be balanced against the requirements of substantial storage and processing resources, potentially impacting policy efforts related to sustainability.

3.3 Strategic foresight

Often, individuals and institutions rely on a default set of assumptions about the future. However, the future is inherently uncertain. For a technology as rapidly evolving (and with such complex geopolitics) as generative AI, unexamined assumptions can lead to miscalculations in governance.

Strategic foresight is a set of methodologies and tools that allow for an organized, scientific approach to thinking about, and preparing for, the future. Adoption of strategic foresight helps governments be agile – to move beyond assumptions of the future, systematically explore critical uncertainties, envisage potential solutions and risks, sandbox new ideas and articulate alternate visions of successful futures.

Strategic foresight has been adopted successfully by various governments. For example, in Finland, the *Government Report on the Future* sets parameters for long-term planning and decision-making.[76] In the United Arab Emirates, the Dubai Future Foundation (DFF) leads 13 councils,[77] each of which convenes government directors and experts to investigate the future of different sectors or issue areas (such as AI), and to identify the governance and capacity needed to drive positive change.

Although strategic foresight initiatives vary, best practices include:

- **Guided**: Use models or prompts to guide exercises, e.g. use scenario planning matrices to consider potential futures across axes of critical uncertainties.
- **Consistent**: Plan exercises on a recurring basis and identify organizational champions.
- **Multistakeholder**: Engage cross-functional internal and external stakeholders to mitigate biases and map multiple possible futures.
- **Transparent**: Track and measure adoption – for example, in Dubai, a numerical scale was developed to rank the effectiveness of each agency in integrating strategic foresight and rankings were then shared to increase healthy competition and incentivize adoption.

3.4 Impact assessments and agile regulations

Agile and flexible regulation is essential in AI to address evolving financial, economic and social impacts. Policy-makers must consider diverse stakeholder input to account for varied sectoral and community short- and long-term impacts. Governments should also study varied agile practices emerging globally and assess jurisdictional fit. For example, they should consider regulatory sandboxes for testing prior to broad deployment. Another approach is "complex adaptive regulations", which are designed to respond to the effects they create and require defined goals, success metrics and thresholds for how regulations will adapt to their own impacts.

Governmental structures can adopt the dynamics of tech companies to become more agile through: 1) a risk-based approach, 2) regular review of technology and marketplace challenges, 3) agile response to challenges,[78] and 4) review of response effects and adaptation.[79] Still, agile governance should not come at the expense of oversight or separation of powers, nor without regard to human rights and rights-based frameworks that ensure that generative AI development and deployment align with societal values and norms. Governments should avoid adopting a "move fast and break things" form of hyper-agility that has been criticized for prioritizing go-to-market testing over mitigation of harmful consequences.

3.5 International cooperation

The current international discussions on generative AI governance frequently lack meaningful participation from global majority countries. This can create significant knowledge gaps about the risks, opportunities and prospects of the generative AI supply chain in those underrepresented regions.[80] Principles and frameworks developed without their input may prove ineffective or even harmful. Unaddressed, these tensions could lead to a fragmentation of the global generative AI community into segregated, non-interoperable spheres.

Thus, international cooperation is essential in six areas (see Table 10) to harness the benefits of generative AI while managing its dangers equitably. This can be achieved through bilateral, regional and broader international mechanisms of cooperation, like those advanced by the World Economic Forum, the United Nations (UN), Group of 20 (G20), the Organisation for Economic Co-operation and Development (OECD) and the African Union High Level Panel on Emerging Technologies (APET).

TABLE 10 **Key areas requiring international cooperation between jurisdictions**

Standards	– Standards can help make abstract AI principles actionable, are more agile than regulations and can bolster global resilience while regulation processes are underway. – They are critical to regulatory interoperability.[81] – Quality assurance techniques and technical standards support cross-border trade. Provisions in free trade agreements (FTAs) are needed to address challenges facing AI innovators. Testing certifications should be interoperable where possible. – Anticipatory standards require increased inclusion of CSOs and academia, and coordination of standards bodies.[82]
Safety	– Strengthened R&D of safety techniques and evaluation tools is key to resilience. – It is crucial to coordinate AI safety institutes to maximize limited resources. An agreement signed by various jurisdictions at the AI Seoul Summit on a network of institutes is promising.[83] – It is additionally necessary to ensure that long-term risks are not prioritized at the expense of identified present AI harms.[84]
Risks	– Establishing mutual understanding of 1) taxonomy of risks, 2) definition and scope of mitigating risks, and 3) approaches is necessary to evaluate, quantify and determine if a model/application meets the risk mitigation threshold. – It is essential to embrace jurisdictional variability on risk tolerance and ethical principles,[85] while advancing risk management interoperability. This can be achieved by considering how standards may apply across high-risk cases while leaving the definition of "high-risk" to jurisdictions. – Collaboration across sectors is crucial for proactively identifying generative AI opportunities and risks (including critical-, systemic- and infrastructure-related). This could be achieved via a dedicated international observatory.
Prohibitions	– Lack of alignment on prohibitions increases the likelihood of generative AI misuse by state or non-state actors with severe global consequences. – Collaboration on treaties or other norm-building mechanisms is needed to establish clear prohibitions on specific forms of generative AI research, development, deployment and use.
Knowledge-sharing	– Participation in a platform, such as a global governance sandbox, enables the sharing of best practices, case studies (e.g. technical, ethical and legal) and tools that allow stakeholders to implement informed governance.
Infrastructure	– Many jurisdictions have limited access to compute and high-quality data for training and fine-tuning, leading to reliance on models prone to error in local languages or contexts. Even open models are not easily fine-tuned to a new language due to underlying tokenization. – Examination of opportunities for multilateral sharing, or shared ownership, of compute and data, alongside the mitigation of bad-actor access or certain other uses, e.g. military. – Developed countries should prioritize sharing resources, expertise and best practices to enable global majority countries to build their AI capabilities and participate effectively in international forums.

Conclusion

This paper is intended to provide policy-makers and regulators with a detailed, practical and implementable generative AI governance framework. Generative AI, like other technologies, is not neutral – it touches upon shared values and fundamental rights. Before introducing new AI regulations, it is crucial to evaluate the current regulatory landscape and enhance coordination among sectoral regulators to mitigate generative AI-induced tensions. Existing regulatory authorities should be assessed for their capability to respond to emerging generative AI challenges, and the trade-offs of a distributed governance approach versus a single dedicated agency should be considered. A comprehensive whole-of-society governance strategy should address industry, civil society and academic challenges, promoting cross-sector collaboration and interdisciplinary solutions. Looking ahead, future strategies need to account for resource limitations and global uncertainties, with adaptable foresight mechanisms and international cooperation through standardized practices and shared knowledge. By adopting a harmonized approach, generative AI challenges can be addressed more effectively at a global level.

Contributors

This paper is a combined effort based on numerous interviews, discussions, workshops and research. The opinions expressed herein do not necessarily reflect the views of the individuals or organizations involved in the project or listed below. Sincere thanks are extended to those who contributed their insights via interviews and workshops, as well as those not captured below.

Lead authors

Rafi Lazerson
Responsible AI Specialist, Accenture; Project Fellow, AI Governance Alliance

Manal Siddiqui
Responsible AI Manager, Accenture; Project Fellow, AI Governance Alliance

Karla Yee Amezaga
Lead, Data Policy and AI, World Economic Forum

World Economic Forum

Samira Gazzane
Policy Lead, Artificial Intelligence and Machine Learning

Accenture

Patrick Connolly
Responsible AI Research Manager

Kathryn White Krumpholz
Managing Director, Innovation Incubation; Executive Fellow, AI Governance Alliance

Andrew J.P. Levy
Chief Corporate and Government Affairs Officer

Valerie Morignat
Responsible AI Senior Manager, Accenture; Project Fellow, AI Governance Alliance

Charlie Moskowitz
Government Relations Senior Manager

Ali Shah
Managing Director, Responsible AI; Executive Fellow, AI Governance Alliance

Dikshita Venkatesh
Responsible AI Research Senior Analyst; Project Fellow, AI Governance Alliance

Acknowledgements

Sincere appreciation is extended to the following working group members, who spent numerous hours providing critical input and feedback on the drafts. Their diverse insights are fundamental to the success of this work.

Lovisa Afzelius
Chief Executive Officer, Apriori Bio

Hassan Al-Darbesti
Adviser to the Minister and Director, International Cooperation Department, Ministry of Information and Communication Technology (ICT) of Qatar

Uthman Ali
Global Responsible AI Officer, BP

Jason Anderson
General Counsel, Vice-President and Corporate Secretary, DataStax

Norberto Andrade
Professor and Academic Director, IE University

Jesse Barba
Head, Government Affairs and Policy, Chegg

Richard Benjamins
Co-Founder and Chief Executive Officer, OdiseIA

Saqr Binghalib
Executive Director, Artificial Intelligence, Digital Economy and Remote Work Applications Office of the United Arab Emirates

Anu Bradford
Professor, Law, Columbia Law School

Daniela Braga
Founder and Chief Executive Officer, Defined.ai

Michal Brand-Gold
Vice-President General Counsel, ActiveFence

Adrian Brown
Executive Director, Center for Public Impact

Melika Carroll
Head, Global Government Affairs and Public Policy, Cohere

Winter Casey
Senior Director, SAP

Daniel Castano Parra
Professor, Law, Universidad Externado de Colombia

Neha Chawla
Senior Corporate Counsel, Infosys

Simon Chesterman
Senior Director, AI Governance, AI Singapore, National University of Singapore

Quintin Chou-Lambert
Office of the UN Tech Envoy, United Nations

Melinda Claybaugh
Director of Privacy Policy, Meta Platforms

Francy Clement
Head, North America Region, Women in AI

Magda Cocco
Head, Practice Partner, Information, Communication and Technology, Vieira de Almeida and Associados

Amanda Craig
Senior Director, Responsible AI Public Policy, Microsoft

Renée Cummings
Data Science Professor and Data Activist in Residence, University of Virginia

Gerard de Graaf
Senior EU Envoy for Digital to the US, European Commission

Nicholas Dirks
President and Chief Executive Officer, The New York Academy of Sciences

Mark Esposito
Faculty Affiliate, Harvard Center for International Development, Harvard Kennedy School and Institute for Quantitative Social Sciences

Nita Farahany
Robinson O. Everett Professor of Law and Philosophy, Duke University; Director, Duke Science and Society

Max Fenkell
Vice-President, Government Relations, Scale AI

Kay Firth-Butterfield
Chief Executive Officer, Good Tech Advisory

Katharina Frey
Deputy Head, Digitalisation Division, Federal Department of Foreign Affairs (FDFA) of Switzerland

Alice Friend
Head, Artificial Intelligence and Emerging Tech Policy, Google

Tony Gaffney
President and Chief Executive Officer, Vector Institute

Eugenio Garcia
Director, Department of Science, Technology, Innovation and Intellectual Property (DCT), Brazilian Ministry of Foreign Affairs (Itamaraty)

Urs Gasser
Dean, TUM School of Social Sciences and Technology, Technical University of Munich

Justine Gauthier
Director, Corporate and Legal Affairs, MILA - Quebec Artificial Intelligence Institute

Debjani Ghosh
President, National Association of Software and Services Companies (NASSCOM)

Danielle Gilliam-Moore
Director, Global Public Policy, Salesforce

Anthony Giuliani
Global Head of Operations, Twelve Labs

Brian Patrick Green
Director, Technology Ethics, Markkula Center for Applied Ethics, Santa Clara University

Samuel Gregory
Executive Director, WITNESS

Koiti Hasida
Director, Artificial Intelligence in Society Research Group, RIKEN Center for Advanced Intelligence Project, RIKEN

Dan Hendrycks
Executive Director, Center for AI Safety

Benjamin Hughes
Senior Vice-President, Artificial Intelligence (AI) and Real World Data (RWD), IQVIA

Marek Jansen
Senior Director, Strategic Partnerships and Policy Management, Volkswagen

Jeff Jianfeng Cao
Senior Research Fellow, Tencent Research Institute

Sam Kaplan
Assistant General Counsel and Senior Director, Palo Alto Networks

Kathryn King
General Manager, Technology and Strategy, Office of the eSafety Commissioner Australia

Edward S. Knight
Executive Vice-Chairman, Nasdaq

James Laufman
Executive Vice-President, General Counsel and Chief Legal Officer, Automation Anywhere

Alexis Liu
Head, Legal, Weights and Biases

Caroline Louveaux
Chief Privacy and Data Responsibility Officer, Mastercard

Shawn Maher
Global Vice-Chair, Public Policy, EY

Gevorg Mantashyan
First Deputy Minister, High-Tech Industry, Ministry of High-Tech Industry of Armenia

Gary Marcus
Chief Executive Officer, Center for Advancement of Trustworthy AI

Gregg Melinson
Senior Vice-President, Corporate Affairs, Hewlett Packard Enterprise

Robert Middlehurst
Senior Vice-President, Regulatory Affairs, e& International

Satwik Mishra
Executive Director, Centre for Trustworthy Technology, Centre for the Fourth Industrial Revolution

Casey Mock
Chief Policy and Public Affairs Officer, Center for Humane Technology

Chandler Morse
Vice-President, Corporate Affairs, Workday

Henry Murry
Vice-President, Government Relations, C3 AI

Miho Naganuma
Senior Executive Professional, Digital Trust Business Strategy Department, NEC

Didier Navez
Senior Vice-President, Data Policy & Governance, Dawex

Dan Nechita
Former Head of Cabinet, MEP Dragoş Tudorache, European Parliament (2019-2024)

Jessica Newman
Director, AI Security Initiative, Centre for Long-Term Cybersecurity, UC Berkeley

Michael Nunes
Vice-President, Payments Policy, Visa

Bo Viktor Nylund
Director, UNICEF Innocenti Global Office of Research and Foresight, United Nations Children's Fund (UNICEF)

Madan Oberoi
Executive Director, Technology and Innovation, International Criminal Police Organization (INTERPOL)

Florian Ostmann
Head, AI Governance and Regulatory Innovation, The Alan Turing Institute

Marc-Etienne Ouimette
Lead, Global AI Policy, Amazon Web Services

Timothy Persons
Principal, Digital Assurance and Transparency of US Trust Solutions, PwC

Tiffany Pham
Founder and Chief Executive Officer, Mogul

Oreste Pollicino
Professor, Constitutional Law, Bocconi University

Catherine Quinlan
M&A Legal Integration Executive, IBM

Roxana Radu
Associate Professor of Digital Technologies and Public Policy, Blavatnik School of Government; Hugh Price Fellow, Jesus College University of Oxford

Martin Rauchbauer
Co-Director and Founder, Tech Diplomacy Network

Alexandra Reeve Givens
Chief Executive Officer, Center for Democracy and Technology

Philip Reiner
Chief Executive Officer, Institute for Security and Technology

Andrea Renda
Senior Research Fellow, Centre for European Policy Studies (CEPS)

Rowan Reynolds
General Counsel and Head of Policy, Writer

Sam Rizzo
Head, Global Policy Development, Zoom Video Communications

John Roese
Global Chief Technology Officer, Dell Technologies

Nilmini Rubin
Chief Policy Officer, Hedera Hashgraph

Arianna Rufini
ICT Adviser to the Minister, Ministry of Enterprises and Made in Italy

Crystal Rugege
Managing Director, Centre for the Fourth Industrial Revolution Rwanda

Joaquina Salado
Head, AI Ethics, Telefónica

Idoia Salazar
Professor, CEU San Pablo University

Nayat Sanchez-Pi
Chief Executive Officer, INRIA Chile

Mark Schaan
Deputy Secretary to the Cabinet (Artificial Intelligence), Privy Council Office, Canada

Thomas Schneider
Ambassador and Director of International Affairs, Swiss Federal Office of Communications, Federal Department of the Environment, Transport, Energy and Communications (DETEC)

Robyn Scott
Co-Founder and Chief Executive Officer, Apolitical

Var Shankar
Affiliate, Governance and Responsible AI Lab (GRAIL Lab), Purdue University

Navrina Singh
Founder and Chief Executive Officer, Credo AI

Scott Starbird
Chief Public Affairs Officer, Databricks

Uyi Stewart
Chief Data and Technology Officer, data.org

Charlotte Stix
Head, AI Governance, Apollo Research

Arun Sundararajan
Harold Price Professor, Entrepreneurship and Technology, Stern School of Business, New York University

Nabiha Syed
Executive Director, Mozilla Foundation

Patricia Thaine
Co-Founder and Chief Executive Officer, Private AI

V Valluvan Veloo
Director, Manufacturing Industry, Science and Technology Division, Ministry of Economy, Malaysia

Ott Velsberg
Government Chief Data Officer, Ministry of Economic Affairs and Information Technology of Estonia

Miriam Vogel
President and Chief Executive Officer, Equal AI

Takuya Watanabe
Director, Software and Information Service Industry Strategy Office, Ministry of Economy, Trade and Industry Japan

Andrew Wells
Chief Data and AI Officer, NTT DATA

Denise Wong
Assistant Chief Executive, Data Innovation and Protection Group, Infocomm Media Development Authority of Singapore

Kai Zenner
Head, Office and Digital Policy Adviser, MEP Axel Voss, European Parliament

Arif Zeynalov
Transformation Chief Information Officer, Ministry of Economy of the Republic of Azerbaijan

Sincere appreciation is also extended to the following individuals who contributed their insights for this report.

Basma AlBuhairan
Managing Director, Centre for the Fourth Industrial Revolution, Saudi Arabia

Abdulaziz AlJaziri
Deputy Chief Executive Officer and Chief Operations Officer, Dubai Future Foundation

Dena Almansoori
Group Chief AI and Data Officer, e&

Daniela Battisti
Senior Advisor and International Relations Expert, Department for Digital Transformation, Italian Presidency of the Council of Ministers

Daniel Child
Manager, Industry Affairs and Engagement, Office of the eSafety Commissioner Australia

Valeria Falce
Full Professor of Economic Law, Senior Advisor and Legal Expert, Department for Digital Transformation, Italian Presidency of the Council of Ministers

Lyn Jeffery
Distinguished Fellow and Director, Institute for the Future (IFTF)

Japan External Trade Organization

Genta Ando
Executive Director and Project Fellow, World Economic Forum

Hitachi America

Daisuke Fukui
Senior Researcher and Project Fellow, World Economic Forum

World Economic Forum

Minos Bantourakis
Head, Media, Entertainment and Sport Industry

Maria Basso
Portfolio Manager, Digital Technologies

Agustina Callegari
Lead, Global Coalition for Digital Safety

Daniel Dobrygowski
Head, Governance and Trust

Karyn Gorman
Communications Lead, Metaverse Initiative

Ginelle Greene-Dewasmes
Lead, AI and Energy

Bryonie Guthrie
Lead, Foresight and Organizational Transformation

Jill Hoang
Lead, AI and Digital Technologies

Devendra Jain
Lead, Artificial Intelligence, Quantum Technologies

Jenny Joung
Specialist, Artificial Intelligence and Machine Learning

Connie Kuang
Lead, Generative AI and Metaverse Value Creation

Benjamin Larsen
Lead, Artificial Intelligence and Machine Learning

Na Na
Lead, Advanced Manufacturing and Artificial Intelligence

Chiharu Nakayama
Lead, Data and Artificial Intelligence

Hannah Rosenfeld
Specialist, Artificial Intelligence and Machine Learning

Nivedita Sen
Initiatives Lead, Institutional Governance

Stephanie Smittkamp
Coordinator, AI and Data

Stephanie Teeuwen
Specialist, Data and AI

Kenneth White
Manager, Communities and Initiatives, Institutional Governance

Hesham Zafar
Lead, Business Engagement

Production

Louis Chaplin
Editor, Studio Miko

Laurence Denmark
Creative Director, Studio Miko

Cat Slaymaker
Designer, Studio Miko

Endnotes

1. Bielefeldt, H., & Weiner, M. (2023). *Declaration on the Rights of Persons Belonging to National or Ethnic, Religious and Linguistic Minorities*. United Nations. https://legal.un.org/avl/pdf/ha/ga_47-135/ga_47-135_e.pdf.
2. United Nations (UN). (1990). *The United Nations Convention on the Rights of the Child*. https://www.unicef.org.uk/wp-content/uploads/2010/05/UNCRC_PRESS200910web.pdf.
3. United Nations Office on Drugs and Crime. (n.d.). *Ad Hoc Committee to Elaborate a Comprehensive International Convention on Countering the Use of Information and Communications Technologies for Criminal Purposes*. https://www.unodc.org/unodc/en/cybercrime/ad_hoc_committee/home.
4. United Nations. (1992). *United Nations Framework Convention on Climate Change*. https://unfccc.int/files/essential_background/background_publications_htmlpdf/application/pdf/conveng.pdf; United Nations. (2015). *Paris Agreement*. https://unfccc.int/sites/default/files/english_paris_agreement.pdf.
5. Leslie, D., Burr, C., Aitken, M., Cowls, J., Katell, M., & Briggs, M. (2021). Artificial intelligence, human rights, democracy, and the rule of law: A primer. *SSRN.* https://doi.org/10.2139/ssrn.3817999.
6. World Economic Forum. (2023). *Data Equity: Foundational Concepts for Generative AI*. https://www3.weforum.org/docs/WEF_Data_Equity_Concepts_Generative_AI_2023.pdf.
7. World Economic Forum. (2020). *A New Paradigm for Business of Data*. https://www3.weforum.org/docs/WEF_New_Paradigm_for_Business_of_Data_Report_2020.pdf.
8. Van Bekkum, M., & Zuiderveen Borgesius, F. (2023). Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception? *Computer Law & Security Review*, vol. 48. https://doi.org/10.1016/j.clsr.2022.105770.
9. Reisner, A. (2023). Generative AI Might Finally Bend Copyright Past the Breaking Point. *The Atlantic.* https://www.theatlantic.com/technology/archive/2024/02/generative-ai-lawsuits-copyright-fair-use/677595/.
10. UK Government. (2023). *Pro-innovation Regulation of Technologies Review—Digital Technologies*. https://assets.publishing.service.gov.uk/media/64118f0f8fa8f555779ab001/Pro-innovation_Regulation_of_Technologies_Review_-_Digital_Technologies_report.pdf.
11. House of Lords Communications and Digital Committee. (2024). *Large language models and generative AI*. https://publications.parliament.uk/pa/ld5804/ldselect/ldcomm/54/54.pdf.
12. Shan, S., Ding, W., Passananti, J., Wu, S., Zheng, H., & Zhao, B. Y. (2024). Nightshade: Prompt-Specific Poisoning Attacks on Text-to-Image Generative Models. *Department of Computer Science, University of Chicago*. https://people.cs.uchicago.edu/~ravenben/publications/pdf/nightshade-oakland24.pdf.
13. Grynbaum, M. M., & Mac, R. (2023). The Times Sues OpenAI and Microsoft Over AI Use of Copyrighted Work. *The New York Times*. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.
14. Hays, K. (2023). Andreessen Horowitz would like everyone to stop talking about AI's copyright issues, please. *Business Insider.* https://www.businessinsider.com/marc-andreessen-horowitz-ai-copyright-2023-11.
15. Hoppner, T., & Ufues, S. (2024). On the Antitrust Implications of Embedding Generative AI in Core Platform Services. *CPI Antitrust Chronicles*, vol. 1, no. 12. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4904876.
16. European Commission. (2024, 23 July). *Joint Statement on Competition in Generative AI Foundation Models and AI Products* [Press release]. https://competition-policy.ec.europa.eu/about/news/joint-statement-competition-generative-ai-foundation-models-and-ai-products-2024-07-23_en.
17. Macko, M. S. (2024). *Applying Data Minimization to Consumer Requests*. California Privacy Protection Agency Enforcement Division. https://cppa.ca.gov/pdf/enfadvisory202401.pdf.
18. Office of the Privacy Commissioner of Canada. (2023). *Principles for responsible, trustworthy and privacy-protective generative AI technologies*. https://www.priv.gc.ca/en/privacy-topics/technology/artificial-intelligence/gd_principles_ai/.
19. Private AI. (n.d.). *Background on PII.* https://docs.private-ai.com/introduction/#background.
20. Ontario Securities Commission. (2024). *Data privacy and the Administrative Arrangement*. https://www.osc.ca/en/about-us/domestic-and-international-engagement/international-engagement/data-privacy-and-administrative-arrangement.
21. E-Safety Commissioner, Australian Government. (n.d.). *Tech Trends Position Statement – Generative AI*. Australian Government. https://www.esafety.gov.au/industry/tech-trends-and-challenges/generative-ai.
22. Government of Canada. (2023, 12 October). *Government of Canada launches consultation on the implications of generative artificial intelligence for copyright* [Press release]. https://www.canada.ca/en/innovation-science-economic-development/news/2023/10/government-of-canada-launches-consultation-on-the-implications-of-generative-artificial-intelligence-for-copyright.html.
23. US Copyright Office, Library of Congress. (2023). *Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence 37 CFR Part 202*. https://public-inspection.federalregister.gov/2023-05321.pdf.

24. Government of the United Kingdom. (2024). *CMA seeks views on Microsoft's partnership with OpenAI*. https://www.gov.uk/government/news/cma-seeks-views-on-microsofts-partnership-with-openai.
25. Atleson, M. (2023). Chatbots, deepfakes, and voice clones: AI deception for sale. *Federal Trade Commission.* https://www.ftc.gov/business-guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale.
26. Competition and Market Authority. (2023). *AI Foundation Models Review: Short Version*. https://assets.publishing.service.gov.uk/media/65045590dec5be000dc35f77/Short_Report_PDFA.pdf.
27. World Economic Forum. (n.d.). *Digital Trust Framework*. https://initiatives.weforum.org/digital-trust/framework.
28. Groves, L., Metcalf, J., Vecchione, B., & Strait, A. (2024). Auditing Work: Exploring the New York City algorithmic bias audit regime. *ACM Digital Library*. https://dl.acm.org/doi/10.1145/3630106.3658959.
29. Smith, B. (2023). How do we best govern AI? *Microsoft on the Issues*. https://blogs.microsoft.com/on-the-issues/2023/05/25/how-do-we-best-govern-ai/.
30. Schrepel, T., & Pentland, A. S. (2023). *Competition Between AI Foundation Models: Dynamics and Policy Recommendations*. Massachusetts Institute of Technology Connection Science. https://ide.mit.edu/wp-content/uploads/2024/01/SSRN-id4493900.pdf?x41178.
31. European Commission. (2024). *Commission Decision Establishing the European AI Office.* https://digital-strategy.ec.europa.eu/en/library/commission-decision-establishing-european-ai-office.
32. World Economic Forum. (2024). *AI Governance Alliance: Briefing Paper Series*. https://www.weforum.org/publications/ai-governance-alliance-briefing-paper-series/.
33. National Institute of Standards and Technology (NIST). (2024). *AI Risk Management Framework*. https://www.nist.gov/itl/ai-risk-management-framework.
34. Marcus, G. (n.d.). *AI Took My Career!* [Broadcast]. https://podcasts.apple.com/gb/podcast/ai-took-my-career/id1532110146?i=1000624493662.
35. World Economic Forum. (2024). *Responsible AI Playbook for Investors*. https://www3.weforum.org/docs/WEF_Responsible_AI_Playbook_for_Investors_2024.pdf.
36. The Forum's AI Governance Alliance is currently researching energy resources as part of the work of the AI Transformation of Industries pillar of work. Publications on this important topic will be released in coming months
37. Personal Data Protection Commission, Singapore. (n.d.). *Advisory Guidelines on use of Personal Data in AI Recommendation and Decision Systems*. https://www.pdpc.gov.sg/guidelines-and-consultation/2024/02/advisory-guidelines-on-use-of-personal-data-in-ai-recommendation-and-decision-systems.
38. Maslej, N. et al. (2024). *The AI Index 2024 Annual Report.* Institute for Human-Centered AI, Stanford University. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf.
39. National Institute of Standards and Technology (NIST). (n.d.). *Generative AI: Text-to-Text (T2T).* https://ai-challenges.nist.gov/t2t.
40. International Standards Organization. (2023). *ISO/IEC 42001:2023*. https://www.iso.org/standard/81230.html.
41. Li, F.-F. (2023). *Governing AI Through Acquisition and Procurement*. Stanford Institute for Human-Centered Artificial Intelligence (HAI), Stanford University. https://hai.stanford.edu/sites/default/files/2023-09/Fei-Fei-Li-Senate-Testimony.pdf.
42. European Commission. (2024). *Living guidelines on the responsible use of generative AI in research*. https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en.
43. World Economic Forum. (2024). *Shaping the Future of Learning: The Role of AI in Education 4.0* https://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Learning_2024.pdf.
44. Osloo, S., (2023, 22 August). *Why we must understand how generative AI will affect children*. World Economic Forum. https://www.weforum.org/agenda/2023/08/generative-ai-children-need-answers/.
45. Solyst, J., Yang, E., Xie, S., Hammer, J., Ogan, A., & Eslami, M. (2024). Children's Overtrust and Shifting Perspectives of Generative AI. *International Society of the Learning Sciences*. https://arxiv.org/pdf/2404.14511.
46. Theil, D. (2023). *Investigation finds AI image generation models trained on child abuse*. Stanford Cyber Policy Center, Stanford University. https://cyber.fsi.stanford.edu/news/investigation-finds-ai-image-generation-models-trained-child-abuse.
47. Thiel, D., Melissa, S., & Portnoff, R. (2023). *New report finds generative machine learning exacerbates online sexual exploitation*. Stanford Digital Repository, Stanford University. https://cyber.fsi.stanford.edu/io/news/ml-csam-report.
48. World Economic Forum. (2023). *Toolkit for Digital Safety Design Interventions and Innovations: Typology of Online Harms*. https://www3.weforum.org/docs/WEF_Typology_of_Online_Harms_2023.pdf.
49. Gruenhagen, J.H. et al. (2024). The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. *Computers and Education: Artificial Intelligence,* vol. 7. https://www.sciencedirect.com/science/article/pii/S2666920X24000766.
50. UNICEF. (n.d.). *Policy guidance on AI for children*. https://www.unicef.org/innocenti/reports/policy-guidance-ai-children.

51. Joint Research Centre, European Commission. (2022). Examining artificial intelligence technologies through the lens of children's rights. https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/examining-artificial-intelligence-technologies-through-lens-childrens-rights-2022-06-22_en.

52. World Economic Forum. (2022). *Artificial Intelligence for Children: Toolkit*. https://www3.weforum.org/docs/WEF_Artificial_Intelligence_for_Children_2022.pdf.

53. Shekhawat, G., & Livingstone, S. (2023). *AI and children's rights: A guide to the transnational guidance*. London School of Economics (LSE). https://blogs.lse.ac.uk/medialse/2023/11/01/ai-and-childrens-rights-a-guide-to-the-transnational-guidance/.

54. Dotan, R. et al. (n.d.). *Evaluating AI Governance: Insights from Public Disclosures*. TechBetter. https://www.techbetter.ai/_files/ugd/f83391_6aed42a5c87448b79821298183428a2e.pdf.

55. Li, F.-F. (2023). *The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI*. Flatiron Books.

56. Organisation for Economic Co-operation and Development. (2024). *Governing with Artificial Intelligence: Are governments ready?* https://doi.org/https://doi.org/10.1787/26324bc2-en.

57. Li, F.-F. (2023). *Governing AI Through Acquisition and Procurement*. Stanford Institute for Human-Centered Artificial Intelligence (HAI), Stanford University. https://hai.stanford.edu/sites/default/files/2023-09/Fei-Fei-Li-Senate-Testimony.pdf.

58. Cities for Digital Rights. (n.d.). *Nine European cities set a common data algorithm register standard to promote transparent AI*. https://citiesfordigitalrights.org/9-european-cities-set-common-data-algorithm-register-standard-promote-transparent-ai.

59. Australian Government Digital Transformation Agency. (2024). *Policy for the Responsible Use of AI in Government*. https://www.digital.gov.au/sites/default/files/documents/2024-08/Policy%20for%20the%20responsible%20use%20of%20AI%20in%20government%20v1.1.pdf.

60. The White House. (2024, 28 March). *Fact sheet: Vice President Harris Announces OMB Policy to Advance Governance, Innovation, and Risk Management in Federal Agencies' Use of Artificial Intelligence* [Press release]. https://www.whitehouse.gov/briefing-room/statements-releases/2024/03/28/fact-sheet-vice-president-harris-announces-omb-policy-to-advance-governance-innovation-and-risk-management-in-federal-agencies-use-of-artificial-intelligence/.

61. Reda, M., Onsy, A., Haikal, A., & Ghanbari, A. (2024). Path planning algorithms in the autonomous driving system: A comprehensive review. *Robotics and Autonomous Systems*, vol. 174. https://doi.org/10.1016/j.robot.2024.104630.

62. Hambling, D. (2024). Hives For U.S. Drone Swarms Ready to Deploy This Year. *Forbes.* https://www.forbes.com/sites/davidhambling/2024/05/16/hives-for-us-drone-swarms-ready-to-deploy-this-year/.

63. Heater, B. (2024). Figure's new humanoid robot leverages OpenAI for natural speech conversations. *TechCrunch.* https://techcrunch.com/2024/08/06/figures-new-humanoid-robot-leverages-openai-for-natural-speech-conversations/.

64. Interpol. (2024). *Beyond Illusions: Unmasking the threat of synthetic media for law enforcement*. https://www.interpol.int/content/download/21179/file/BEYOND%20ILLUSIONS_Report_2024.pdf.

65. Associated Press. (2024). X restores Taylor Swift searches after deepfake explicit images triggered temporary block. *AP News.* https://apnews.com/article/taylor-swift-x-searches-deepfake-images-adec3135afb1c6e5363c4e5dea1b7a72.

66. Thiel, D., Melissa, S., & Portnoff, R. (2023). *New report finds generative machine learning exacerbates online sexual exploitation*. Stanford Digital Repository, Stanford University. https://cyber.fsi.stanford.edu/io/news/ml-csam-report.

67. Shearer, J. (2024). Taylor Swift deepfakes on X falsely depict her supporting Trump. *NBC News. NBC Universal.* https://www.nbcnews.com/tech/internet/taylor-swift-deepfake-x-falsely-depict-supporting-trump-grammys-flag-rcna137620.

68. Martelloni, P.-H. (2021). *Modélisation et Simulation des systèmes complexes spatialisés. Utilisation de Systèmes Multi-Agents et Multi-composant pour la gestion des pêcheries.* Universite de Corse-Pascal Paoli. https://theses.hal.science/tel-03683015v1/document.

69. Samuel, S. (2024). People are falling in love with—And getting addicted to—AI voices. *Vox.* https://www.vox.com/future-perfect/367188/love-addicted-ai-voice-human-gpt4-emotion.

70. Skaug Saetra, H., & Mills, S. (2022). Psychological interference, liberty and technology. *Technology in Society*, *vol. 69*. https://doi.org/https://doi.org/10.1016/j.techsoc.2022.101973.

71. Tong, A. (2023). AI chatbot company Replika restores erotic roleplay for some users. *Reuters.* https://www.reuters.com/technology/ai-chatbot-company-replika-restores-erotic-roleplay-some-users-2023-03-25/.

72. Atillah, I. E. (2023). Man ends his life after an AI chatbot "encouraged" him to sacrifice himself to stop climate change. *Euro News.* https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-.

73. World Economic Forum. (2024). *Generative AI Governance: Shaping a Collective Global Future* https://www3.weforum.org/docs/WEF_Generative_AI_Governance_2024.pdf.

74. Taori, R., & Hashimoto, T. B. (2022). Data Feedback Loops: Model-driven Amplification of Dataset Biases. *Arxiv.* https://doi.org/10.48550/arXiv.2209.03942.

75. Shumailov, I. et al. (2024). The Curse of Recursion: Training on Generated Data Makes Models Forget. *Arxiv.* https://doi.org/10.48550/ARXIV.2305.17493.

76. Finnish Government. (2023). *Government Report on the Future*. https://valtioneuvosto.fi/en/foresight-activities-and-work-on-the-future/government-report-on-the-future.

77. Dubai Future Foundation. (n.d.). *Foreseeing Dubai's Future.* https://www.dubaifuture.ae/initiatives/future-foresight-and-imagination/dubai-future-councils/.

78. Lawfare, YouTube. (2024). *Lawfare Daily: Former FCC Chair Tom Wheeler on AI Regulation*. https://www.youtube.com/watch?v=Oodn1zEjLvI.

79. Organisation for Economic Co-operation and Development (OECD). (2024). *Regulatory Experimentation: Moving ahead on the Agile Regulatory Governance Agenda.* https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/regulatory-experimentation_fc84553c/f193910c-en.pdf.

80. Carnegie Endowment for International Peace. (2024). *AI Governance for the Global Majority: Understanding Opportunities and Challenges*. https://carnegieendowment.org/events/2024/05/ai-governance-for-the-global-majority-understanding-opportunities-and-challenges?lang=en.

81. World Economic Forum. (2024). *Generative AI Governance: Shaping a Collective Global Future.* https://www3.weforum.org/docs/WEF_Generative_AI_Governance_2024.pdf.

82. Alania, A. et al. (2022). *Looking Ahead: The Role of Standards in the Future of Artificial Intelligence (AI) Governance*. University College London. https://www.ucl.ac.uk/steapp/sites/steapp/files/looking_ahead_the_role_of_standards_in_the_future_of_ai_governance_v2.0.pdf.

83. Government of the United Kingdom. (2024, 21 May). *Global leaders agree to launch first international network of AI Safety Institutes to boost cooperation of AI* [Press release]. https://www.gov.uk/government/news/global-leaders-agree-to-launch-first-international-network-of-ai-safety-institutes-to-boost-understanding-of-ai.

84. World Economic Forum. (2024). *Generative AI Governance: Shaping a Collective Global Future*. https://www3.weforum.org/docs/WEF_Generative_AI_Governance_2024.pdf.

85. Carnegie Endowment for International Peace. (2024). *AI Governance for the Global Majority: Understanding Opportunities and Challenges*. https://carnegieendowment.org/events/2024/05/ai-governance-for-the-global-majority-understanding-opportunities-and-challenges?lang.

WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org