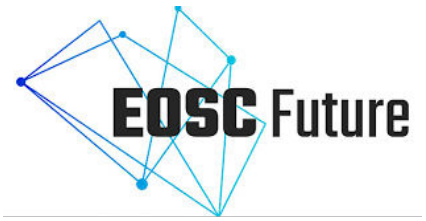


EOSC-Future & RDA Artificial Intelligence/Data Visitation Working Group

AI Bill of Rights Recommendation

Version 2.4 Nov 19, 2024

Prepared by AI Bill of Rights Team: Eyiuche Ezigbo(NG), Shiny Martis B(FR), Natalie Meyers(US), Ronit Purian-Lukatch(IL), Yeyang Su(CN)

Subject matter experts from across the globe developed these recommendations to promote fundamental human rights and advance trust in AI for Open Science for European and global platform communities from perspectives engaging ethics, law, medicine, and social considerations in Artificial Intelligence, Data Sharing, Data Visitation, and Open Science policy development. These recommendations are consensus-driven and actionable.

We hold as important tenets of our work that:

Adoption of AI Bills of Rights/policies within member organizations can create mutual trust and benefits for all.

And that

Attention to AI Governance can improve the longevity and relevance of research environments and communities like European Open Science Cloud(EOSC) and the Research Data Alliance(RDA).

Motivation: The AI Bill of Rights team goal is to present recommendations to the European Open Science Cloud(EOSC) Community and Research Data Alliance(RDA) on the needs for AI Governance/AI Bill of Rights in various jurisdictional, disciplinary and research scenarios taking into account the potential rights of data creators, model developers, model and data re-users, and citizens/communities/patients whose lives/privacy/wellbeing are impacted by AI and data sharing modalities.

Our team looks at how AI governance should shape what data is gathered and used in AI systems, what the rights of individual persons are inside this data, and how communities adopt and adapt to AI driven forms of decision making and creativity where ecological flourishing and human wellbeing are prioritized at the outset and throughout all design of any AI system as key metrics of societal and global prosperity and progress. It is our aim to inform an understanding of Inclusive AI Governance for EOSC constituencies and the RDA membership, as well as to encourage Civil Society Participation in Standards Development in AI to move beyond risk models in isolation to help shepherd AI systems to complement a flourishing future for all.

Table of Contents

Critical Tenets

We hold as critical tenets of our work that:

Adoption of AI Bills of Rights/ policies within member organizations will create mutual trust and benefits for all.

And that

Attention to AI Governance improves the longevity and relevance of research environments and communities like the European Open Science Cloud and the Research Data Alliance.

Introduction

Allowing and encouraging a culture of data access through data visitation[i] agreements for research, and to train models, can improve frequency of access to re-usable data sets and a better way of making machine actionable data use agreements. In the circumstance of creating or providing platforms for in silico trials and access to digital twins or "many worlds" for research purposes, certain rights can speed the progress of research or limit use of diagnostics, treatments/interventions, or forecasts created in such environments. Mutual data use agreements must therefore also take into consideration how other parties' governance of and use of AI may have used unconsented data to train models or have consumed and repurposed copyrighted data without permission inside generative AI models. Bias exists in models and training datasets. Decision making based on such system outputs can cause personal and community harms as a consequence. Therefore, it is of importance for communities and research partners interacting with models to communicate about and understand whether models are running under jurisdictional or other governance requiring AI governance, and/or transparency (or permitting opacity and to what extent) and whether they are compliant with BS ISO/IEC 42001:2023, ISO/IEC 23894:2023, or IEEE 7010-2020[1] or other technical standard(s). Information provided in such governance scenarios should include how current a model's training data is, the frequency with which models are retrained, what is the re-training periodicity (if-any), and which state, national or regional jurisdictional governance applies to the models and/or training data used in deployed AI systems.

Goal

The AI Bill of Rights team goal is to present recommendations to the EOSC Community and Research Data Alliance on the needs for AI Governance/AI Bill of Rights in various

[i] Data visitation means a process in which: i) data sets are subject to analysis within a host location; ii) without the data ever leaving the host location; iii) where the analytical framework according to which the data will be analysed can be submitted by a third party outside the host location; and iv) the results can be returned to that third party. See the federated learning example in Holloway et al. 2024 p 29 where a network of servers would create a secure space where federated learning can be carried out without the data leaving the premises or jurisdiction (DOI: 10.5281/zenodo.13831453).

jurisdictional, disciplinary and research scenarios taking into account the potential rights of data creators, model developers, model and data re-users, and citizens/communities/patients whose lives/privacy/wellbeing are impacted by AI and data sharing modalities. Our team looks at how AI governance can shape what data is gathered and used in AI systems, what the rights of individual persons are inside this data, and how communities adopt and adapt to AI driven forms of decision making and creativity. It is our aim to inform an understanding of Inclusive AI Governance for the EOSC community and the RDA membership, as well as to encourage Civil Society Participation in Standards Development in AI.

There are many efforts underway in various jurisdictional and disciplinary communities to govern AI and create policy that ensures a climate of progress for AI driven implementations in business, industry and research while at the same time balancing the needs of impacted people's rights to privacy that keeps lives harm from bias/prejudice. Concerns and tensions abound as use of AI begins to challenge previous notions of what it means to be a rights holder, or a model creator/deployer, and what it means to know and understand what kind of data is used to train models. In turn this leads to questions about what kind of AI should/could be used to create art, diagnose and treat patients, educate students, write papers and code, influence politics, surveil municipal safety or regional immigration, just to name a handful of use cases.

Relationship to other AIDV Working Group outputs/efforts

These recommendations proposed by the AIDV Working Group's Bill of Rights team are meant to be informed by and responsive to the concurrent outputs of AIDV WG team deliverables particularly the AIDV WG ***Guidance on Informed Consent***[2], and the AIDV WG ***Guidance for Ethics Committees Reviewing AI and Data Visitation***[3].

Recommendations

We recommend the below elements as essential to prioritize as parts of a Responsible AI[4] strategy for any group (organization, company, university, hospital, jurisdiction/community).

Essential Elements

1. AI & Acceptable Use Policies: Formalize the definition of AI in your community understanding that AI is a vague term and where necessary use more precise language (e.g., Large Language Model, Generative AI, Machine Learning, etc.). Document key roles and responsibilities related to AI, outline the acceptable and unacceptable uses of AI for your community.
2. AI Working Group (oversight committee): Form an oversight committee whose function it is to oversee AI systems, allocate resources, develop and maintain AI governance policies and procedures for your community.
3. AI Training and Education: Training and change management are integral to the successful operationalization of AI and AI policy. Offer learning opportunities on an Introduction to AI, on Ethical considerations for AI, as well as training on Development of AI strategy, Effective AI governance and Risk management, in addition to training on the necessary data infrastructure to enable effective AI work, including product/project management for how to run effective AI projects, and how to use technical tools in an AI landscape/cyberinfrastructure.
4. AI Impact Assessment
5. Ongoing Monitoring: Consider, will you have a Human in the loop, Human over the loop, Human out of the loop monitoring plan[5]?
6. Transparency, Notice & Disclosure

7. Vendor Due Diligence[ii]
8. Test AI systems for Bias: Identify different types of biases: (Systemic, Human, Computational) and recognize that mitigating bias requires action at each stage of the AI lifecycle

Impact Assessment

Because it is essential to govern systems that can potentially cause certain harms and it is essential to assess the Impacts of AI Systems, we recommend the below Impact assessment elements[6] as essential to prioritize as parts of a responsible AI strategy for any group (organization, company, university, hospital, jurisdiction/community):

- Determine what will be the AI system impact assessment process
- Documentation of AI system impact assessments
- Assessing AI System impact on individuals and groups of individuals
- Assessing societal impacts of AI systems
- AI system impact assessment process

AI Alignment

Any person or organization creating or adopting an AI policy needs to be able to discuss, understand and document what will be your strategy for AI Alignment. AI Alignment is an AI safety research/process that aims to ensure AI systems achieve desired outcomes. If you're creating or implementing an AI or AI policy you need to be able to define the system or policy's desired outcomes, and articulate how you'll document the processes for monitoring, measuring and logging where you achieve and fall short of achieving desired outcomes/benefits of AI or AI policy.

Any group creating or implementing an AI or AI policy should[7]:

- Consider auditable detection for bias. Bias detection exposes cases where decision logic is in violation of agreed upon ethics.
- Consider auditable detection for data changing by necessity over time and the AI model requiring re-validation
- Consider auditable detection for privacy violations. Privacy leak detection exposes cases where decision logic is in violation of agreed upon ethics ensuring privacy.
- Consider who is monitoring the model looking for drift after deployment - if drift is significant you have to retrain and revalidate/test.
- Have a control and escalation process to assign responsibility for investigation, a process for investigation, a process for notification and disclosure of violations, and a process for reconciliation/restitution to impacted individuals and groups to mitigate and correct consequences of non-aligned AI.

Recommendations for the EOSC Community and Research Data Alliance

To reiterate, we hold as a critical tenet of our work that **Adoption of AI Bills of Rights/ policies within member organizations will create mutual trust and benefits for all**

And that

[ii] i.e., Can your community easily switch vendors if necessary? Does your community have a voice in how the vendor operates? Is a vendor clear how community data will be treated?

Prioritization of AI Governance will improve the longevity and relevance of research environments and communities like European Open Science Cloud and the Research Data Alliance.

AI Governance can be approached from human rights, risk based, or safety-based motivations or a combination of all three. So far, most organizations and jurisdictions regulate artificial general intelligence based on the risks such models may give rise to during their development, deployment and dissemination. The 2023 State of AI Report mentions how "Amid the theoretical debate, labs are building in their own mitigations" and offers as examples how DeepMind's "proposed toolkit and associated workflow for extending standard model evaluations" assesses "for potentially dangerous capabilities (e.g. cyber-offense, self-proliferation) and propensity to cause harm." and how Anthropic's "Responsible Scaling Policy, with a risk-based list of safety commitments" (internal access controls, red-teaming, third-party evaluations, and tiered access for different AI Safety Levels) builds in "development breaks if safety measures fail to keep up with capabilities"[8].

Yet, as Coglianese argues, "risk-based regulation—like regulatory excellence more generally—is not a merely technical enterprise. It requires not only technical competence, but also principled decision-making, transparency, careful attention to empirical evidence and on-the-ground implementation"[9]. And, the application of AI raises fundamental questions regarding human rights. "At most, risk assessments inform regulators' decisions; they do not provide a full basis for them . . . Risk assessment provides scientific or empirical answers about probabilities, hazards, and their distribution; it does not supply the policy principle or normative reason needed to make regulatory or risk management decisions about these hazards"[10].

Groups like the RDA and the EOSC Community have a remit to go beyond this risk-based rhetoric that dominates corporate AI Governance strategies. Thus, in contrast to an exclusively risk-based governance, we hereby recommend adoption of a "human-rights first" based approach to AI governance. Organizations like RDA, due to their open international membership, have a need to prioritize and accommodate the perspectives and disparities of and between developing countries alongside the needs of companies and countries at the forefront of AI development and adoption. By taking a human rights first approach to AI governance, members of the RDA community will hold not only our own work accountable for current and future generations, but also provide actionable guidance to those guiding and regulating this fast-moving field towards the future.

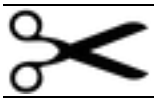
Considering the increasing call for global coordination on A(G)I governance, we call on the EOSC Community and RDA to take up leading roles in allying practitioners to center human-rights including the rights of model creators and deployers as core principles in AI and data governance, research, development, and deployment activities.

For the EOSC Community we recommend attention be paid to differentiating between persons as "model creators" separately from model deployers, or model users, and persons whose data is in training data sets, or persons who are impacted by 3rd party use of models for decision making. Bills of Rights and AI Governance recommendations tend to ignore, gloss over, or collapse treatment of the distinctive rights of model/AI creators and deployers. We call out the importance of acknowledging how persons can (and likely will be) in one or many AI impacted

classes over their lifetimes and careers. For example: as students, instructors, patients, customers, employees, employers, and citizens but also as model creator/deployers.

For the EOSC Community we recommend in-depth and documented consideration of the circumstances and rights of instructors teaching AI development and deployment, students studying AI, model creators, deployers and re-users, data collectors, providers, and sharers as individuals, research groups, organizations, and companies alongside current definitions of impacted communities and individuals in AI policy in ways that acknowledge the rights and well-being of those in the role of innovators/creators/implementers.

For Research Data Alliance: We acknowledge that RDA is well positioned to champion the primacy of protection of human rights in globally impactful AI Governance mechanisms. We recommend the membership and organization consider hosting and offering AI Governance learning opportunities and feedback sessions prior to introduction of any membership-wide attempt to develop an RDA specific organizational AI Bill of Rights. In an organization like RDA with such an international membership it will be of paramount importance that members acknowledge the complexities of jurisdictional pluralism and opportunities of cross-national policy layers when it comes to AI policy making. RDA has a duty to honor and not to ignore differences and protections already granted to its individual members as citizens, residents, researchers, and students under the regional, national, state and municipal laws that may already govern its members where they live, work, and study.



We propose for RDA Output Endorsement the aforementioned recommendations on lines 1-182. The next sections of this document are informational in nature. If the aforementioned recommendations are endorsed, the document can be truncated here and the following sections can be maintained as a supporting output.

Keeping up with the evolving AI landscape

The AIDV-WG's research on emerging Jurisdictional and Special Interest AI Rights and Protections encompasses resources that intersect human rights, education, ethics, health/medicine, regulation/governance and risk. The AIDV WG's open, shared citation library,[11] gathers emerging and codified rights for AI/LLM/Model creators, deployers, and harmed/ impacted parties at the International, Regional, National, and Municipal levels and disciplinary scenarios.

Figure 1 Shared Citation Library (credit AIDV-WG)

We have categorized these resources by whether they relate to the categories of: “Disciplinary or Special Interest Rights and Protections”, Education, Ethics, “Healthcare and Medicine”, “Generative AI Publishing/Authorship and Copyright”, “Informed Consent”, “Jurisdictional Rights Declarations and Acts”, Legal, “Privacy and Surveillance”, Risks, and “Related News Stories”. The assets in our shared library are tagged by country and jurisdiction where relevant.

AI Regulation accelerates and AI Risk dominates the governance conversation

Why did we tag by “risk”? If an AI risk mitigation regulation or policy is adopted it becomes necessary to accompany it with a methodology for identifying, assessing, and mitigating these risks.

Therefore, the risks posed by Artificial Intelligence (AI) can easily dominate policy priorities, AI conversations, press coverage, and AI governance spending. For example, the MIT Future Tech’s AI Risk Repository, at the time of this report, is a live, open database of ~777 risks extracted from 43 taxonomies with an accompanying taxonomy and preprint.[12] A flurry of model management tools has arisen to organize AI governance and risk management tasks alongside a burgeoning ecosystem of responsible AI startups.[13,14]

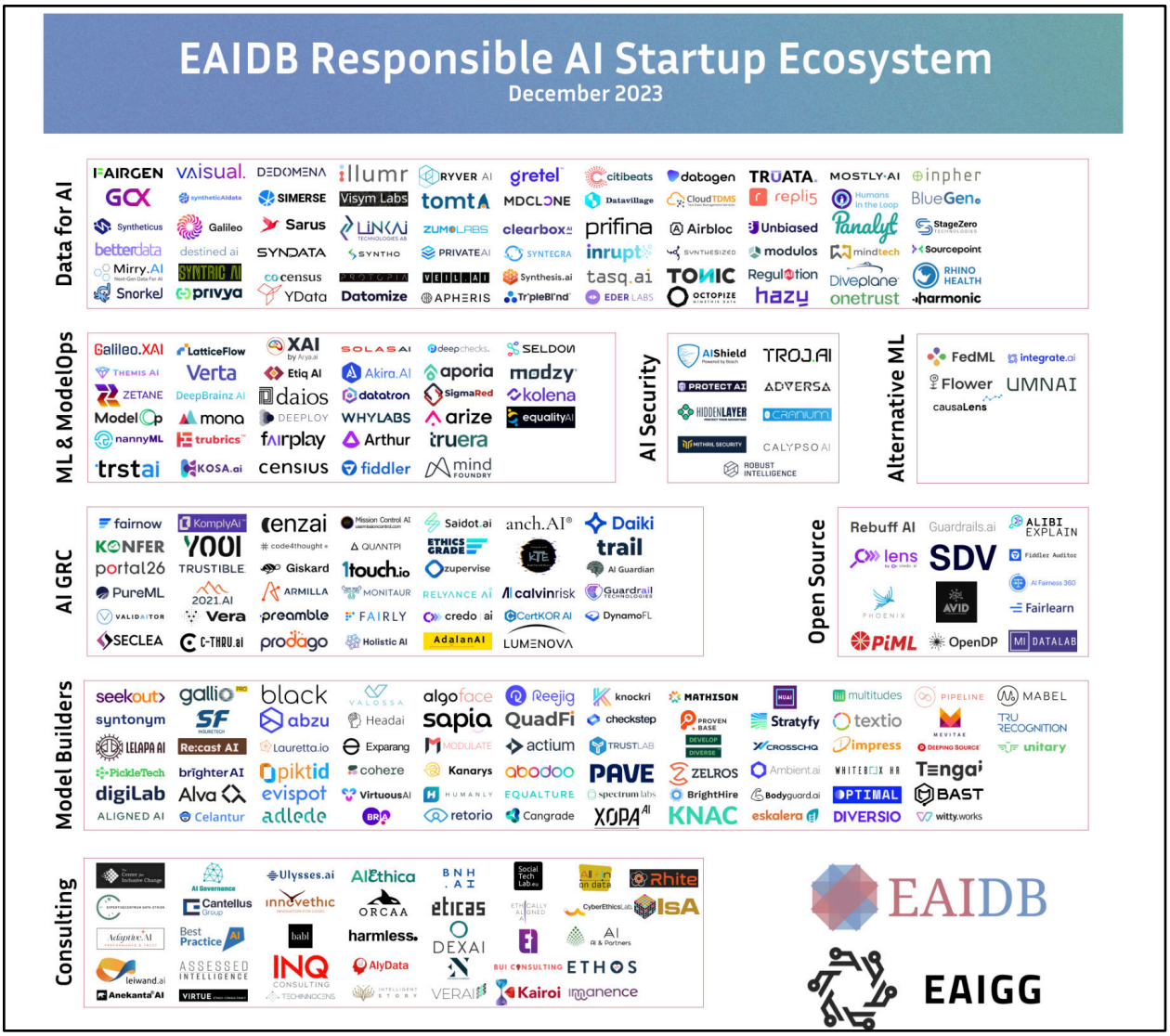


Figure 2. EAIDB Responsible AI Startup Ecosystem credit https://www.eaidb.org/map

The landscape of Standards and Jurisdictional AI Policy

Below, arising from our literature search and team research effort, we summarize global, international, regional, country, municipal level, disciplinary and special interest policies and foci and give attention to relevant technical standards (e.g. ISO, NIST, IEEE) for the benefit of the EOSC Community and the Research Data Alliance.

UNESCO: Recommendations on AI Ethics

On 23 November, 2021, after three years of preparations, UNESCO adopted the Recommendation on the Ethics of Artificial Intelligence, first-ever comprehensive UN guidelines for AI[15], which were adopted by 193 countries[16]. All Member states of the UNESCO

adopted the recommendations that define common values and principles needed to ensure the healthy development of AI. Signatories agree that Member States are to:

> "apply on a voluntary basis the provisions of this Recommendation by taking appropriate steps, including whatever legislative or other measures may be required, in conformity with the constitutional practice and governing structures of each State, to give effect within their jurisdictions to the principles and norms of the Recommendation in conformity with international law, including international human rights law" and that "Member States engage all stakeholders, including business enterprises, to ensure that they play their respective roles in the implementation of this Recommendation; and bring the Recommendation to the attention of the authorities, bodies, research and academic organizations, institutions and organizations in public, private and civil society sectors involved in AI technologies, so that the development and use of AI technologies are guided by both sound scientific research as well as ethical analysis and evaluation."

adopted a landmark resolution on the promotion of "safe, secure and trustworthy" AI systems that will also benefit sustainable development for all. The resolution calls on the Member States to refrain from using AI systems that cannot comply with international human rights law, and ensure that people's rights online are protected the same way as they are offline. The resolution emphasized protection of personal data, monitoring of AI risks, and the need to close the digital divide.

The Group of Friends of the Implementation of the Recommendation on the Ethics of Artificial Intelligence was launched on 3 February 2022 as an informal and open-ended network of Member States. The Group of Friends meets regularly to support UNESCO's work on implementation and to create a space to share experiences and good practices for the full implementation of the Recommendation. At their first meeting they also stressed "the importance of including in this Group of Friends external stakeholders including the private sector, academics and others, on a needs basis".

Regional, country and municipal level policies and foci

Several regional and country level AI governance commonalities and differences can be seen in examples from the OECD, China, Africa, Nigeria, the Middle East and North Africa region, Israel, The European Union, the Council on Europe, Italy, the United Kingdom, the United States, California, New York City, Canada, Latin America, Australia, and Taiwan. Below we offer brief overviews of some noteworthy components of these.

OECD: "Ethics Guidelines for Trustworthy AI"

The Organisation for Economic Co-operation and Development (OECD) AI Principles are among the first such principles signed up to by governments. They promote AI that is innovative and trustworthy and that respects human rights and democratic values. They were adopted in May 2019 by OECD member countries when they approved the OECD Council Recommendation on Artificial Intelligence. They include concrete recommendations for public policy and strategy, and their general scope ensures they can be applied to AI developments

around the world. The principles interpret fairness through the lenses of equal access, inclusive design processes, and equal treatment: diversity, nondiscrimination, fairness, inclusion – throughout the entire AI system's life cycle.

The OECD guidelines for ethical AI – promote seven governance principles: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, nondiscrimination and fairness, (6) environmental and societal well-being, and (7) accountability.

Governments that have Committed to the AI Principles

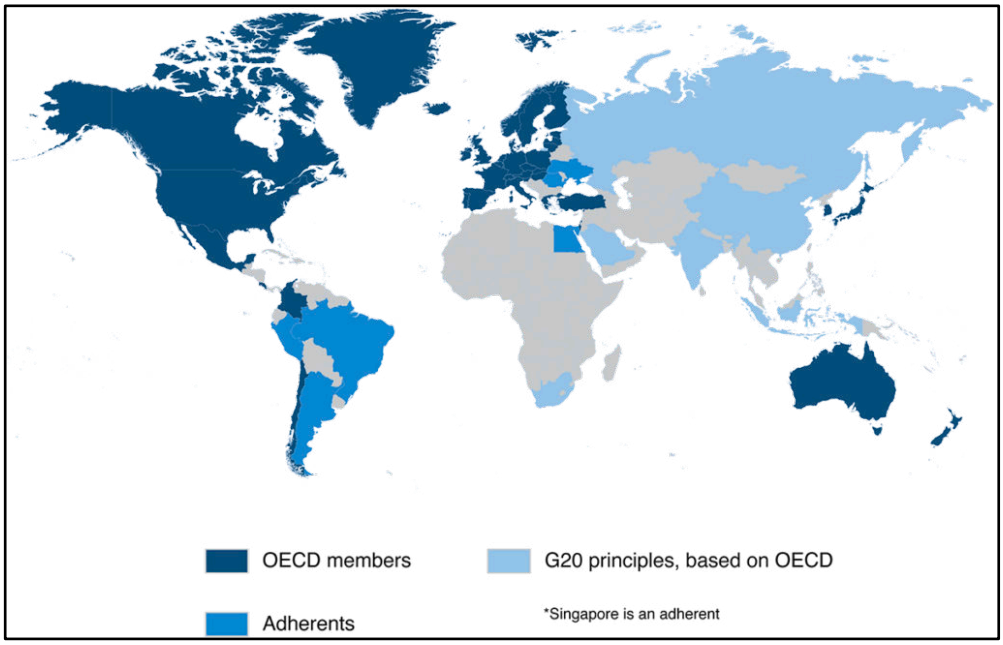


Figure 3. "Governments that have Committed to the AI Principles" © 2023 OECD. All rights reserved

Global Partnership on Artificial Intelligence (GPAI)

GPAI[17], organized in 2020 around a shared commitment to the above OECD *Recommendation on Artificial Intelligence*, brings together experts from science, industry, civil society, governments, international organisations and academia to foster international cooperation related to AI and convenes working groups on both Data Governance and Responsible AI.

GPAI's **Working Group on Data Governance**[18] collates evidence, shapes research, undertakes applied AI projects and provides expertise on data governance, to promote data for AI being collected, used, shared, archived and deleted in ways that are consistent with human rights, inclusion, diversity, innovation, economic growth, and societal benefit, while seeking to address the UN Sustainable Development Goals. This group has released numerous outputs toward these ends, including "***Data Justice in Practice: A Guide for Policymakers***[19] and a companion ***Data Justice in Practice: A Guide for Impacted Communities***[20].

The GPAI **Responsible AI Working Group**[21] is grounded in a vision of AI that is human-centred, fair, equitable, inclusive and respectful of human rights and democracy, and that aims at

contributing positively to the public good. The GPAI countries have put climate action and biodiversity preservation at the top of their agenda, positing that as a general-purpose technology, AI can be harnessed responsibly to accelerate positive environmental action. In 2021 the WG released an action-oriented set of recommendations entitled ***CLIMATE CHANGE AND AI: Recommendations for Government Action*** to guide policy makers developing AI related climate action strategies[22].

In 2022, the GPAI Responsible AI Working Group and Project RAISE (Responsible AI Strategy for the Environment)[23] began to implement some of the recommendations of the WG's AI & climate roadmap by assessing the environmental impacts of AI compute and preparing AI readiness booklets for key industries to achieve net-zero. The scope was later expanded to start building a roadmap on AI and biodiversity preservation. Three WG deliverables produced within this mandate include:

1. ***The AI Footprint: Measuring the environmental impacts of AI Compute and Applications***[24], a guide drafted in consultation with RAI experts and the OECD.AI Expert Group on AI Compute and Climate. This report provides actionable information for policymakers who wish to implement the principles and priorities of data justice in their policymaking activities.
2. AI for Net Zero electricity[25]
3. Biodiversity and AI: Opportunities and Recommendations for Action[26]

Regulation on AI in China: A brief overview

Ethical governance of artificial intelligence (AI) has been integrated in China's 'new generation artificial intelligence development plan' since 2017[27]. The effort to develop guiding ethical principles and standards has also been integrated into China's recent and wider efforts to develop ethical governance of science and technology since late 2018. Since AI is an interdisciplinary and intersectoral field that involves the scientific and industry communities and concerns the public at large, multiple agencies are involved in overseeing the development of standards for ethical governance of the research, development and deployment of AI in China. Leading agencies include and not limited to the Ministry of Industry and Information Technology (MIIT), the Ministry of Science and Technology (MOST), the Cyberspace Administration of China (CAC) and the State Administration for Market Regulation.

Even though at the national level, no law has yet been made specifically on AI, the policy makers have paid great attention on regulating data security, protecting personal information, and regulating the cyberspace. For instance, the Data Security Law of People's Republic of China[28], and the Personal Information Protection Law of the People's Republic of China[29] were released and put into effect in the year of 2021. A White Paper that summarizes the Chinese approach in regulating the cyberspace was recently published by the State Council Information Office[30].

A series of efforts have been made to steer ethical governance of AI development at the national level in China by entities such as the Expert Committee for Ethical Governance of the Next Generation AI and National Informational Security Standardisation Technical Committee. These efforts have led to the publication of *the Governance Principles for* a *New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence* in 2021[31]. Eight guiding principles are 1) harmony and friendliness, 2) fairness and justice, 3) inclusiveness and sharing, 4) respect for privacy,5) security and controllability, 6) shared responsibility, 7) open cooperation and 8) agile governance. Standards have also been developed in both key technical

areas and in ethical terms, such as guidelines for prevention of ethical & security risks of artificial intelligence in 2021[32]and guidelines for the standardisation of artificial intelligence ethical governance in 2023[33].

Notably, China has taken a proactive approach in regulating generative AI. In November 2022, the CAC, the MITT and the Ministry of Public Security jointly issued the *Administrative Provisions on Deep Synthesis in Internet-Based Information Services* that were put into effect on 1 January 2023[34]. In response to the rising popularity and elicited concern on generative AI, on 11 April 2023, the CAC released a draft regulation on the use of generative AI for public consultation[35] and seven governmental agencies have subsequently jointly issued the regulation in July.

Lastly, at the municipal level, Shenzhen and Shanghai have developed policies to steer the development of AI industries in responsible manners within their jurisdiction[36]. Given the widespread interest in developing AI industries around China, it is likely other provinces and/or cities will follow suit. And at the international level, it is worth noticing that, not only China has endorsed WHO's *Ethics & Governance of Artificial Intelligence for Health* (2021)[37] and UNESCO's *Recommendation on the Ethics of Artificial Intelligence* (2022)[38], but it has also begun to actively promote dialogue and cooperation on the governance of AI. In late 2022, the Ministry of Foreign Affairs issued the *Position Paper of the People's Republic of China on Strengthening Ethical Governance of Artificial Intelligence*, stating that 'China is committed to building a community with a shared future for mankind in the domain of AI, advocating a people-centered approach and the principle of *AI for good'* [39]. The ethos is echoed in the aforementioned white paper that was published by the State Council Information Office this year.

In recent years, China has made considerable efforts to steer ethical development of AI research, innovation and related industries. Even though there is yet no overarching law on AI, with China's proactive approach in regulating AI, the ethical and regulatory framework it has established is to large extent able to cover the evolving fields of AI research, development and deployment and to agilely respond to emerging challenges. More importantly, the ethical principles China have adopted are not that distinctive from other countries and regions, and the Chinese government has shown great interest in promoting collaboration in developing international guidance and standards on ethical AI in line with its commitment to building a community with a shared future for mankind.

Taiwan

Tawan initiated the "AI Action Plan" in 2018 with the aim to prosper AI development and to build solid digital infrastructure. The government is drafting an act to govern artificial intelligence (AI) and expects to have it ready for the legislature in September 2023. The National Science and Technology Council (NSTC) announced the draft act will cover the legal definition of AI, privacy protections, data governance, risk controls and ethical principles related to AI[40].

Africa

Africa has many organizations dedicated to offering AI services, training, research. Early among them are: The Machine Intelligence Institute of Africa (MIIA)[41], an African non-profit organization founded in 2016. AI Kenya formed soon after in 2017, as a community aimed at bringing together like-minded people to discuss, build and grow the Artificial Intelligence and Data Science ecosystem in Kenya and East Africa.

The Forum on Artificial Intelligence in Africa gathered in Benguérir (Kingdom of Morocco) on 12 and 13 December 2018, to reflect and debate on the different dimensions of AI within the African context[42].

Later, in 2019 the In Union (AU) established a working group on Artificial Intelligence with a mandate to study the creation of a common African stance on AI, the development of an Africa-wide capacity-building framework, and to establish an AI think tank to assess and recommend projects to collaborate on in line with the AU's Agenda 2063 and the United Nations ("UN") Sustainable Development Goals (SDGs).

Africa's AI Centre of Excellence Africa (AICE) was founded in Kenya in 2020 to create value and sustainable impact within the African Intelligence and Machine Learning space.

Responsible AI in Africa: Challenges and Opportunities (2023) is an open access collection edited by Damian Okaibedi Eke, Kutoma Wakunuma, and Simisola Akintoye[43] that offers a good overview of regional progress and challenges including pre COVID-19 years.

In the forward, Mark Coeckelbergh calls the reader's attention to the responsibility to study when and where "tech innovation and use of digital technologies actually works in African contexts, including cases in which that might already be responsible." He extolls how:

> "There are examples of good strategies, inside and outside Africa. Ethics should not only be about what could go wrong, but also about what already goes right and why, and about how we can shape a good common future. These comments also raise the question: who should develop responsible AI in Africa? The answer must be: in the first place Africans themselves. Neocolonialism is game over, and AI ethics should not be the place to perpetuate or renew it. Unfortunately, as Stahl et al. remind us in this volume, AI strategies are often dominated by powerful global actors. Such structural problems need to be addressed[44]. Yet it should not be an excuse to refrain from developing one's own policy ecosystem with regard to AI and similar technologies. This is about creating opportunities for responsible innovation. This is about harnessing the benefits of AI for citizens. This is about taking responsibility for one's own technological future."

AI strategies have been released by the North African countries of Algeria, Egypt and Tunisia.

Table 1 Overview of existing North-African AI strategies

Country	*Title of AI strategy document*	*Status*	*Policy areas/priority segments or sectors*	*Regulatory/ethical considerations*
Algeria	The national strategy of research and innovation on Artificial Intelligence (2020–2030)	Completed	– Higher education – Health – Energy – Technologies	
Egypt	National AI Strategy	Completed	– Agriculture/environment and water management – Healthcare – Natural Language Processing – Economic planning – Manufacturing and infrastructure management	Track and monitor implementation of strategy, laws and regulations, ethical principles and guidelines
Tunisia	National Artificial Intelligence Strategy	In progress		Sustainable, equitable development, and ethical challenges

Tunisia

Work started early in Tunisia which created a Task Force in 2018 to "devise a methodology and an action plan to produce the country's National AI Strategy (Ministère de l'Enseignement Supérieur et de la Recherche Scientfique 2018).

Egypt

In crafting its national strategy, the Egyptian government formed the National Council for AI in 2020 "as a partnership between the governmental institutions, prominent academics and practitioners from leading businesses in the field of AI" with a responsibility amongst others, to "identify AI applications that provide smart, safe and sustainable solutions and services". The country expects to "track and monitor the implementation of the strategy, laws and regulations, ethical principles and guidelines" (MCIT 2020).

Algeria

For example, Algeria presented its National Artificial Intelligence Strategy 2020–2030 on the 18th of January, 2021, to "improve Algerian skills in the field of AI through education, training

and research, on the one hand, and strengthen these capacities as a development tool allowing socio-economic sectors to iron out the obstacles hindering the digital transition underway, on the other hand" ("Strategy for research in artificial intelligence launched" 2021)

In 2021 an AI Blueprint[45] project was led by the Republic of South Africa, and the Smart Africa Secretariat (SAS) with contributions from the AU, The Republic of Rwanda, Uganda, and many others to help guide African member-states towards developing policies, strategies and plans best ensuring growth and prosperity. There is a section within the *AI Blueprint* which guides the member-states on what they could or should implement around the matters requiring AI intervention as a technology of choice.

The recent 2024 Lawyers Hub report, *5 Years of AI Regulation in Africa,* puts much of the above progress in continental and regional context providing comparative details for different stages of AI implementation, policy making, and regulation.[46]

Figure 4. Five years of AI Regulation in Africa © 2024 Bonyo et. al

Nigeria

Nigeria published their draft National Artificial Intelligence Strategy (NAIS) in August 2024 which emphasizes the potential of AI to drive economic growth, enhance national security, and address social challenges. “One of the key objectives of the NAIS is to create a regulatory environment that supports innovation while ensuring ethical standards are maintained. This involves developing guidelines for AI research and development, data privacy, and security, as well as establishing mechanisms for monitoring and evaluating the impact of AI initiatives” (Bonyo et. al 2024). The below timeline illustrates milestones along the path to Nigeria’s policy.

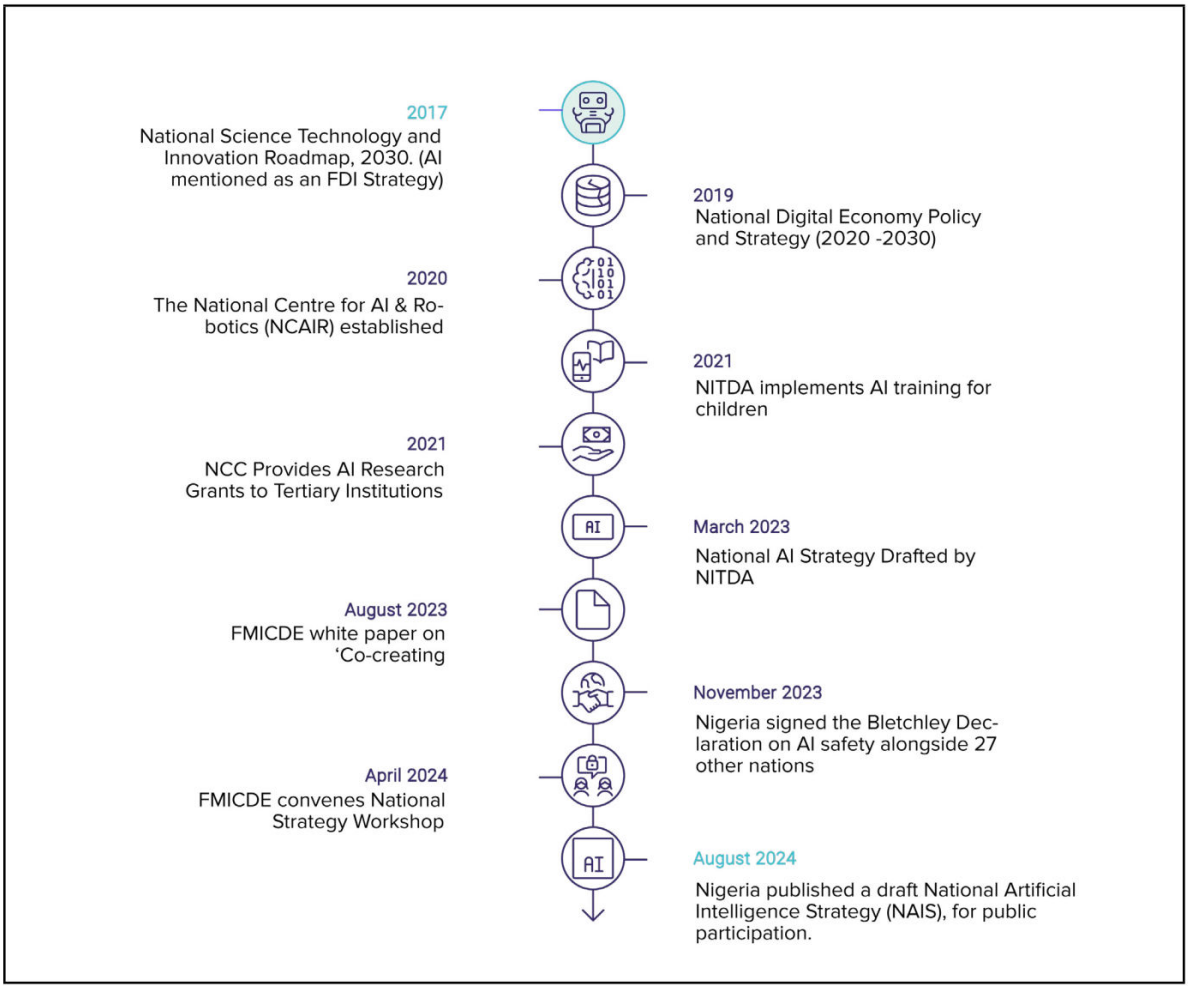


Figure 5. AI Policy Timelines for Nigeria 2017-2018 © 2024 Bonyo et. al

The evolution of AI policy in Nigeria builds on the groundwork laid when the National Information Technology Development Agency (NITDA)[46] was established in April 2001 following the Federal Executive Council's (FEC) approval of the Nigerian National IT Policy. A subsequent responsibility was to implement the Nigerian Information Technology Policy and coordinate general IT development in Nigeria. The NITDA Act (2007)[47] mandated creation of a framework for the planning, research, development, standardization, application, coordination,

monitoring, evaluation and regulation of Information Technology practices, activities and systems in Nigeria.

The role

- Is to develop, regulate and advise on Information technology in the country through regulatory standards, guidelines and policies.
- NITDA is the clearing house for all IT projects and infrastructural development in the country.
- It is the prime Agency for e-government implementation, Internet governance and general IT development in Nigeria.

With the advent of the National Digital Economy Policy and Strategy (NDEPS), which was formally unveiled by the immediate past government on 28th November 2019, it became necessary for the NITDA to chart a new vision for itself. The implementation of the NDEPS has also required that the Agency create a new department designated as the Department of digital economy, to be dedicated fully to the implementation of the policy.

The various arms of NITDA related to AI now engage with:

- The National Centre for Artificial Intelligence and Robotics (NCAIR)[48] which was launched in November 2019 to spur development of an AI ecosystem in the Nigeria;
- The National Digital Economy Policy and Strategy (NDEPS);[49]
- Artificial Intelligence Centre of Excellence Africa (AICE)[50] was founded in 2020 in partnership with the Massachusetts Institute of Technology (MIT) to support the development of AI skills and talent in the country;
- The NCAIR is a collaborative effort between the government, academia, and industry. It **aims** to promote the development and deployment of AI and robotics technologies in Nigeria. Its **objectives** include fostering research and development in AI and robotics, developing AI and robotics talent, and promoting the use of AI and robotics to solve national challenges.

NITDA's track record prior to release of its AI Policy includes publication of technical frameworks and policy guidelines and a remit to monitor their compliance alongside the continuing development of information technology in Nigeria. The instruments serve as a minimum benchmark in the development and implementation of information technology in Nigeria, complement the new AI policy and are enforceable by law.

- NDPR Implementation Framework[51]
- Guidelines for The Management of Personal Data By Public Institutions In Nigeria, 2020[52]
- Guidelines for Nigerian Content Development in Information and Communication Technology (ICT) [53]
- Framework and Guidelines for Public Internet Access[54]
- Guidelines for Clearance of Information Technology (IT) Project by Public Institutions[55]
- Guidelines for Registration of ICT Service Providers/Contractors for Delivery of It Services to MDAs[56]
- Nigeria e-Government Interoperability Framework (Ne-GIF)[57]

- Framework and Guidelines for Information and Communication Technology (ICT) Adoption in Tertiary Institutions[58]
- Framework and Guidelines for the Use of Social Media Platforms in Public Institutions[59]
- Nigerian Government Enterprise Architecture

Nigeria's policy for ethical AI deployment[60] and NITDA's above guidelines for data protection provide the bedrock on which the nation's businesses and organizations can implement responsible AI. The early guidelines and recently issued policy outline principles such as fairness, transparency, and accountability. The aim is to ensure that AI deployment is done in a manner that is consistent with human rights and ethical principles.

Identified Challenges to be addressed are the continuing need for infrastructure investment, and skills development to support AI research and development.[61]

Middle East and South Asia

Riyadh AI Call for Action Declaration (RAICA) Digital Cooperation Organization (DCO)

The DCO brings together the Ministries of Communications and IT of sixteen nations – Bahrain, Bangladesh, The Republic of Cyprus, Djibouti, The Gambia, Ghana, Greece, Jordan, Kuwait, Morocco, Nigeria, Oman, Pakistan, Qatar, Rwanda, and Saudi Arabia. Following a Ministerial roundtable at the Global AI Summit 2022, member states of the Digital Cooperation Organization (DCO) agreed to adopt the Riyadh AI Call for Action Declaration (RAICA)[62], which seeks to use AI technology to benefit people, communities, nations, and the world as a whole. The declaration was signed by all members of DCO.

> As AI technology continues to advance and become embedded in every aspect of daily life, it becomes increasingly necessary for leaders to ensure that it is utilized solely to benefit people, communities, nations, and the planet.
>
> The Riyadh AI Call for Action Declaration advances this commitment to identify and address present, emerging, and future humanitarian issues in the field of AI and highlight ways to use AI as a tool to benefit lives around the world, improve the quality of work, better design public policies, and bring efficiency into the ecosystem.
>
> This is to be achieved through seven key pillars, each comprising principles that address methods to ensure that all enjoy the benefits of AI and the harm is felt by none. To advance this goal, each principle is built upon one overarching objective, which serves as the foundation of this Declaration: AI FOR THE GOOD OF HUMANITY.

India

The Indian government plans to release a draft regulatory framework for Artificial Intelligence (AI) by mid-year, aimed at promoting economic growth while addressing risks. This initiative, announced at the Nasscom Leadership Summit, also includes efforts to develop AI skills among individuals and to establish global governance for AI safety and trust.

Israel

In Israel, despite the publication of several plans related to AI over the last decade, developed by multi-participant expert committees, their implementation has mostly been delayed due to political instability[63].

A few noteworthy exceptions include: C4IR NETWORK, The Israeli hub of the WEF C4IR network, advancing smart regulation for emerging technologies such as AI and blockchain, through collaboration with government regulators and the tech industry in Israel. The AI Strategy Governmental Team: This cross-governmental team was set up in 2020 to devise recommendations for a policy plan to promote AI R&I activities in Israel.

The most recent draft[64] was open for public consultation, by the previous government. However, the current Prime Minister has recently recommenced on preparing a new national policy for generative AI. The prime minister made comments about AI regulation in a recent interview where he made clear his lack of trust in success for global governance of AI in part because of the need for every country to competitively enter AI full force.[65]

As of today, Israel's regulation of research data in academia follows the European regulatory framework, mainly the General Data Protection Regulation (GDPR). Related documents: Tel Aviv University complies with the GDPR (document in Hebrew)[66].

European Union (EU)

The EU's General Data Protection Regulation (GDPR)

The GDPR (2016/679) Regulation in EU law on data protection and privacy[67] is an important component of EU privacy law and of human rights law, building on Articles 7 and 8 of the EU Charter of Fundamental Rights, recognising respect for private life and protection of personal data as closely related but separate fundamental rights[68]. It also addresses the transfer of personal data outside the EU and EEA areas. GDPR does not limit research objectives, as long as the objectives are defined – and accordingly, define fair and transparent data collection and processing.

The GDPR notably sets two guiding principles:

- "Privacy by Default" principle prioritizes the alternative that better protects privacy;
- "Privacy by Design" principle precedes the implementation of privacy considerations to the initial stage of system design and research plan.

These two principles are summarized into the **data minimization** guideline, "collect the minimum amount of personal data you need". Both principles rely on the implementation of a risk analysis from the data controller, on the setting up of technical and organisational measures adapted to the field, to the processing characteristics, purposes and technological environment in which it is taking place. The objectives being to ensure data subject's rights and personal data protection throughout the processing in compliance with GDPR and to conceive privacy-preserving technologies as enablers. For an further analysis of the changes brought by this

regulation in the EU, we refer readers to: The impact of the EU general data protection regulation on scientific research (Chassang, 2017)[69] and to the *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*[70].

GDPR Data sharing across countries

The GDPR restricts data transfer from EU countries; and prohibits its transfer to a country that does not maintain an adequate level of personal information protection. Data sharing among researchers and research centers established in different EU member States of the EU countries is allowed, but data transfer outside the borders of those countries is subject to additional conditions for ensuring compliance with the EU regulations (see Chapter V GDPR). One of the mechanisms planned for allowing such data transfers consists in the recognition of the third country as ensuring an adequate level of personal data protection (via a European Commission's adequacy decision) as complying with the European regulations.

As of autumn 2023, the European Commission has so far recognised Andorra, Argentina, Canada (commercial organisations), Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Republic of Korea, Switzerland, the United Kingdom under the GDPR and the LED, the United States (commercial organisations participating in the EU-US Data Privacy Framework) and Uruguay as providing adequate protection[71].

With the exception of the United Kingdom, these adequacy decisions do not cover data exchanges in the law enforcement sector which are governed by the Law Enforcement Directive (Article 36 of Directive (EU) 2016/680)[72].

The European Data Governance Act (DGA) and Data Act

The EU DGA (November 2020) is a cross-sectoral instrument aiming to make more data available by regulating the re-use of publicly/held, protected data, by boosting data sharing through the regulation of novel data intermediaries and by encouraging the sharing of data for altruistic purposes. Both personal and non-personal data are in scope of the DGA, and wherever personal data is concerned, the General Data Protection Regulation (GDPR) applies.

The Data Act on harmonised rules on fair access to and use of data (adopted as Regulation (EU) 2023/2854 Dec 13, 2023) complements the Data Governance Act of November 2020 by clarifying who can create value from data under which conditions. It includes:

- Measures that enable users of connected devices to access the data generated by these devices and related services
- Measures to provide protection from unfair contractual terms that are unilaterally imposed
- Mechanisms for public sector bodies to access and use data held by the private sector
- New rules that grant customers the freedom to switch between various cloud data-processing service providers
- Measures to promote the development of interoperability standards

The Data Act's aim is to give both individuals and businesses more control over their data through a reinforced portability right, enabling copying or transferring data easily from across

different services, where the data are generated through smart objects, machines, and devices. Uses by connected products using AI are ostensibly a case covered by the Data Act.

Italy

In February 2023 the Italian Data Protection Authority (Garante per la protezione dei dati personali) banned AI Replika on the basis that it presented a risk to minors and to those considered emotionally fragile.[73] At the end of March 2023 Garante then imposed an immediate temporary limitation on the processing of Italian users' data by OpenAI/ChatGPT.[74,75,76] In the case of OpenAI, the action was prompted by a data breach related to ChatGPT conversations and payments and several concerns arose, including that: 1)no information was being made available to users and data subjects whose data are being collected by Open AI; 2)lack of legal basis underpinning the massive collection and processing of personal data in order to 'train' the algorithms on which the ChatGPT platform relies. 3)Concern over how information made available by ChatGPT does not always match factual circumstances, so that inaccurate personal data can be and are processed. 4)Lack of age verification mechanism allows use by minors and may expose children to receiving responses that are absolutely inappropriate to their age and awareness. By April 2023 Italy had lifted its ChatGPT ban, and the service was once again online in Italy[77]. The ChatGPT privacy policy was now accessible to people before they registered and there was a new tool to verify the age of users and OpenAI had explained that it would also provide a new form for European Union users to exercise their right to object to its use of personal data to train its models. An Italian Garante Spokesperson said they would be planning and conducting an information campaign to inform Italians of what happened as well as of their right to opt-out from the processing of their personal data for training algorithms."

AI Act (aka EU AIA)[78]

The EU acknowledges that "the use of AI, with its specific characteristics (e.g. opacity, complexity, dependency on data, autonomous behaviour), can adversely affect a number of fundamental rights and users' safety". The European Union's AI Act lays the foundations for the regulation of AI in the EU. The EU AI Act is now finalized and published on 13 July 2024: Regulation - EU - 2024/1689 - EN - EUR-Lex (europa.eu) It comes into force on 1 August 2024. The AI Act Explorer also contains the Artificial Intelligence Act, Official Journal version of 13 June 2024 and enables users to explore the contents of the Act or search for parts that are most relevant to them. Priorities for the act are to "make sure that AI systems used in the EU are safe, transparent, traceable, non-discriminatory and environmentally friendly" and that "AI systems should be overseen by people, rather than by automation, to prevent harmful outcomes"[79].

In general, the rules establish obligations for providers and users depending on the level of risk from artificial intelligence. AI applications would be regulated only as strictly necessary to address specific levels of risk whereby legal intervention is tailored to concrete level of risk. The draft AI act distinguishes between AI systems posing (i) unacceptable risk, (ii) high risk, (iii) limited risk, and (iv) low or minimal risk.

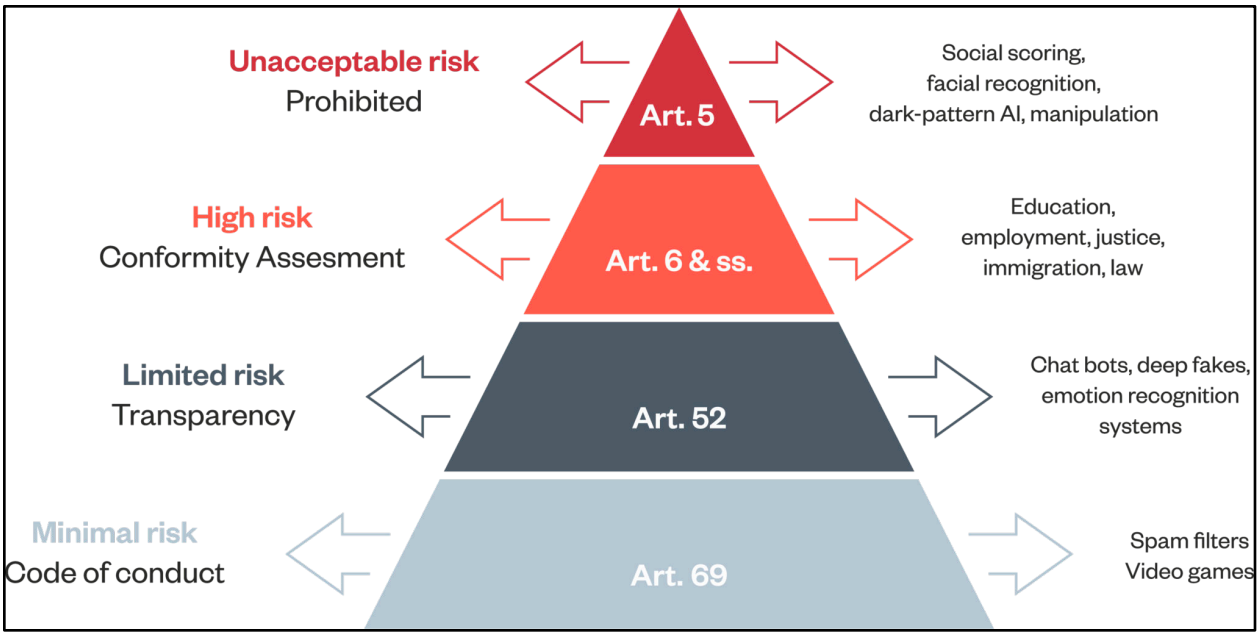


Figure 6. *Pyramid of Risks.* Edwards. (2022). Ada Lovelace Institute. CC-BY-4.0

AI Systems considered a threat to people will be banned and characterized as "Unacceptable Risk" AI systems. They include:

- Cognitive behavioural manipulation of people or specific vulnerable groups: for example, voice-activated toys that encourage dangerous behaviour in children
- Social scoring: classifying people based on behaviour, socio-economic status or personal characteristics
- Real-time and remote biometric identification systems, such as facial recognition

AI systems that negatively affect safety or fundamental rights will be characterized as "High Risk". High-risk AI systems will be assessed before being put on the market and also throughout their lifecycle and will be divided into two categories:

1. AI systems that are used in products falling under the EU's product safety legislation[80]. This includes toys, aviation, cars, medical devices and lifts.
2. AI systems falling into eight specific areas that will have to be registered in an EU database:
    1. Biometric identification and categorisation of natural persons
    2. Management and operation of critical infrastructure
    3. Education and vocational training
    4. Employment, worker management and access to self-employment
    5. Access to and enjoyment of essential private services and public services and benefits
    6. Law enforcement
    7. Migration, asylum and border control management
    8. Assistance in legal interpretation and application of the law.

Under the EU AI Act, Generative AI, like ChatGPT, must comply with transparency requirements:

- Disclosing that the content was generated by AI
- Designing the model to prevent it from generating illegal content
- Publishing summaries of copyrighted data used for training

Limited risk AI systems will be required to comply with minimal transparency requirements that would allow users to make informed decisions. After interacting with the applications, the user can then decide whether they want to continue using it. Users should be made aware when they are interacting with AI. This includes AI systems that generate or manipulate image, audio or video content, for example deepfakes.

The AIA allows the free use of minimal-risk AI. This includes applications such as AI-enabled video games or spam filters. The vast majority of AI systems currently used in the EU fall into this category.

Responses to the AIA before its adoption were varied and numerous. Some common concerns were voiced in The Civil Society Statement for Fundamental rights in EU AI Act and issued in Nov 2021 by European Digital Rights (EDRi) and 119 civil society organisations[81] who bring to attention that:

> "In Europe, we have already witnessed the negative impact of AI when governed incorrectly. For example, discriminative AI uses at the border have facilitated the deportation of people on the move and denied them access to vital services such as health-care and social security. We have also seen how the use of predictive policing systems has led to increased over-policing of racialised communities, and how poor, working-class and migrant areas are being wrongfully targeted by fraud detection systems. The use of facial recognition and similar systems have been used across Europe in ways that lead to biometric mass surveillance.
>
> By fostering mass surveillance and amplifying some of the deepest societal inequalities and power imbalances, AI systems are putting our fundamental rights and democratic processes and values at great risk. That is why the European Union (EU) institutions' proposal for an AIA is a globally significant step. But the AIA must address the structural, societal, political and economic impacts of the use of AI. This will ensure that the law is future-proof, and prioritises the protection of fundamental rights."

The Civil Society Statement for Fundamental rights in EU AI Act also emphasizes importance to govern the development and use of AI systems so they are deployed in a "sustainable, resource-friendly way which respects our planetary boundaries".

> As a first step towards addressing environmental dimensions of sustainability, we need transparency about the level of resources needed to develop and operate AI systems. To address this, we recommend: The introduction of horizontal, public-facing transparency requirements on the resource consumption and greenhouse gas emission impacts of AI systems – irrespective of risk level – in relation to design, data management and training, application, and underlying infrastructures (hardware, data centres, etc.).

Council of Europe

The Council of Europe established in 1949 is composed of 46 member states and six observer states (Canada, Holy See, Israel, Japan, Mexico, United States of America). The Council and the European Union (established 1957 & now comprised of 27 states) share the same fundamental values – human rights, democracy and the rule of law – but are separate entities which perform different, yet complementary, roles[82]. The Council of Europe brings together governments from across Europe – and beyond – to agree to form minimum legal standards in a wide range of areas. It then monitors how well countries apply the standards that they have chosen to sign up to.

Council of Europe's Framework Convention and the Primacy of Human Rights

The *Framework Convention on Artificial Intelligence, Human Rights, Democracy and The Rule Of Law* (CETS No. 225) was signed Sept 2024[83] during a conference of Council of Europe Ministers of Justice in Vilnius by Andorra, Georgia, Iceland, Norway, the Republic of Moldova, San Marino, the United Kingdom, Israel, the United States of America and the European Union(on behalf of its twenty seven member states) making it the first-ever international legally binding treaty aimed at ensuring that the use of AI systems is fully consistent with human rights, democracy and the rule of law. The *Framework* applies to the development, design and application of AI systems with a focus on ensuring the continued seamless application of human rights especially in contexts where AI systems assist or replace human decision-making or perform other tasks relevant in such contexts and works to assure that AI systems shall only be used in such a way that they do not, directly or indirectly, endanger or undermine democratic processes.

Background: In 2019 the Council of Europe's ad hoc Committee on Artificial Intelligence (CAHAI) was tasked with examining the feasibility of negotiating a legally binding treaty to ensure that activities within the lifecycle of artificial intelligence systems are "fully consistent with human rights, democracy and the rule of law, while being conducive to technological progress and innovation". In December 2021, the outcome, CAHAI's paper on the "Possible elements of a legal framework on artificial intelligence, based on Council of Europe's standards on human rights, democracy and the rule of law[84]" was adopted [and] the Committee of Ministers of the Council of Europe instructed a succeeding Committee on Artificial Intelligence (CAI) to move forward with drafting the treaty. The *Framework* itself was then drafted by the 46 member states of the Council of Europe, with the participation of all observer states: Canada, Japan, Mexico, the Holy See and the United States of America, as well as the European Union, and a significant number of non-member states: Australia, Argentina, Costa Rica, Israel, Peru and Uruguay.

The Framework acknowledges policy pluralism and Chapter II Article 4 39 is a list of the main global and regional international human rights instruments and treaties to which various states that negotiated the Framework Convention may be parties.

Who is covered by the framework?

- The Framework Convention covers the use of AI systems by public authorities – including private actors acting on their behalf – and private actors.
- The Convention offers Parties two modalities to comply with its principles and obligations when regulating the private sector: Parties may opt to be directly obliged by the relevant Convention provisions or, as an alternative, take other measures to comply with the treaty's provisions while fully respecting their international obligations regarding human rights, democracy and the rule of law.
- Parties to the Framework Convention are not required to apply the provisions of the treaty to activities related to the protection of their national security interests but must ensure that such activities respect international law and democratic institutions and processes. The Framework Convention does not apply to national defense matters nor to research and development activities, except when the testing of AI systems may have the potential to interfere with human rights, democracy, or the rule of law.

European Ethical Charter on the use of Artificial Intelligence in Judicial systems

In December 2018, The European Commission for the Efficiency of Justice (CEPEJ) adopted the Working Group on quality of justice's (GT-QUAL)[85] "European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment" as one of the first European texts setting out ethical principles relating to the use of artificial intelligence (AI) in judicial systems and the main principles to be observed when developing AI applications in alignment with Human Rights.

The CEPJ Charter takes the strong position that "it is essential to ensure that AI remains a tool in the service of the general interest and that its use respects individual rights." The Charter identifies five core principles to be respected in the field of AI and justice and asserts that "compliance with these principles must be ensured in the processing of judicial decisions and data by algorithms and in the use made of them." The Charter principles are:

1. Principle of respect of fundamental rights: ensuring that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights;
2. Principle of non-discrimination: specifically preventing the development or intensification of any discrimination between individuals or groups of individuals;
3. Principle of quality and security: with regard to the processing of judicial decisions and data, using certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment;
4. Principle of transparency, impartiality and fairness: making data processing methods accessible and understandable, authorising external audits;
5. Principle "under user control": precluding a prescriptive approach and ensuring that users are informed actors and in control of their choices.

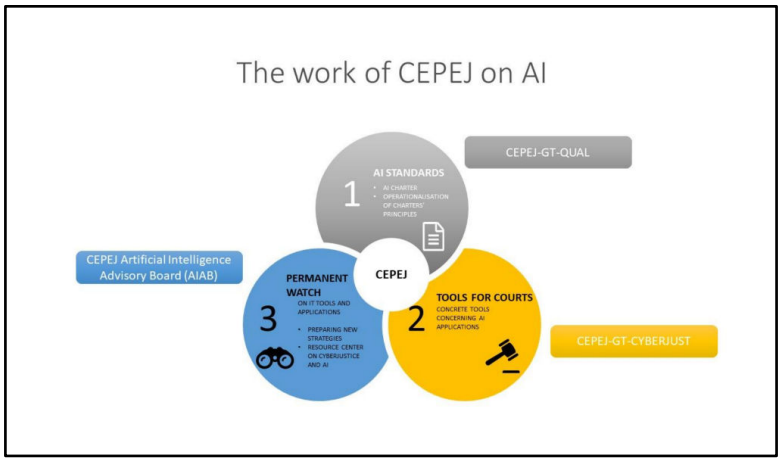


Figure 7. "The work of CEPEJ on AI." CEPEJ (2021)

The CEPEJ Artificial Intelligence Advisory Body (AIAB)[86] (see above) monitors the actual emergence of artificial intelligence applications in the justice sector and provides expert guidance on the operationalisation of the principles of the Charter. The AIAB is in the process of developing an AI compliance "Assessment Tool" checklist and guidelines which will give authorities in charge of implementing judicial AI systems more practical guidance on how to apply the CEPEJ Charter's five principles.

The CEPEJ has also established a Resource Centre on Cyberjustice and Artificial Intelligence (ReC)[87]. The ReC maintains a publicly accessible registry of AI systems and other key cyberjustice tools used in judiciary cyberinfrastructure. "The Centre serves as a starting point for detailed examinations of specific applications with a view to their compliance with European standards of democracy, human rights and rule of law. Relevant authorities, researchers and the interested public are invited to use the displayed information for that purpose"[88].

UK Algorithmic Transparency Standard

The Cabinet Office's Central Digital and Data Office has developed an algorithmic transparency standard for government departments and public sector bodies with the Centre for Data Ethics and Innovation making the UK one of the first countries in the world to develop a national algorithmic transparency standard[89]. The standard has been piloted by some public sector organisations and will be further developed based on feedback.

The UK is also adopting a pro-innovation approach to AI regulation through a framework that aims to build public confidence and to provide "a clear and unified approach to regulation" without introduction of new legislation[90]:

> Initially, we do not intend to introduce new legislation. By rushing to legislate too early, we would risk placing undue burdens on businesses. But alongside empowering regulators to take a lead, we are also setting expectations. Our new monitoring functions will provide a real time assessment of how the regulatory framework is performing so that we can be confident that it is proportionate.

The UK framework is underpinned by 5 principles to guide and inform the responsible development and use of AI in all sectors of the economy. These five principles will be issued on a non-statutory basis and implemented by existing regulators.

- Safety, security and robustness
- Appropriate transparency and explainability
- Fairness
- Accountability and governance
- Contestability and redress

The proposed framework is aligned with, and supplemented by, a variety of tools for trustworthy AI, such as assurance techniques, voluntary guidance and technical standards. The UK Government will promote the use of such tools and is collaborating with partners like the UK AI Standards Hub[91] to encourage responsible AI innovation (see part 4 of the Framework for details). The Framework also acknowledges risk and promises regulatory intervention to ensure that AI does not cause harm at a societal level, threatening democracy[92] or UK values:

> We have made an initial assessment of AI-specific risks and their potential to cause harm, with reference in our analysis to the values that they threaten if left unaddressed. These values include safety, security, fairness, privacy and agency, human rights, societal well-being and prosperity.

> Our assessment of cross-cutting AI risk identified a range of high-level risks that our framework will seek to prioritise and mitigate with proportionate interventions. For example, safety risks include physical damage to humans and property, as well as damage to mental health. AI creates a range of new security risks to individuals, organisations, and critical infrastructure. Without government action, AI could cause and amplify discrimination that results in, for example, unfairness in the justice system. Similarly, without regulatory oversight, AI technologies could pose risks to our privacy and human dignity, potentially harming our fundamental liberties.

United States

Post GDPR Data sharing history between EU the United States

Despite the profound difference between Europe and the United States, data can be transferred between these two giant economies of the Western world.

The United States has a special status. The Federal Trade Commission in the United States and the European Commission formulated the Privacy Shield – an arrangement that allows researchers to voluntarily comply with the European regulation.

By virtue of this arrangement many companies, among them Microsoft, Google, Amazon and Facebook, have joined Privacy Shield to permit data transfer to them. Research data can be stored in Privacy Shield cloud services. Otherwise, researchers should refrain from transferring GDPR data to a server in the United States.

US/EU TTC Joint Roadmap for trustworthy AI and risk management

The U.S.-EU Joint Statement of the Trade and Technology Council released a “Joint Roadmap” on evaluation and measurement tools for trustworthy AI and risk management December 1,

2022[93]. This Joint Roadmap aims to guide the development of tools, methodologies, and approaches to AI risk management and trustworthy AI by the EU and the United States and to advance our shared interest in supporting international standardization efforts and promoting trustworthy AI on the basis of a shared dedication to democratic values and human rights. The roadmap takes practical steps to advance trustworthy AI and uphold our shared commitment to the Organisation for Economic Co-operation and Development (OECD) *Recommendation on AI* (see OECD section above).

In *The EU and U.S. diverge on AI regulation: A transatlantic comparison and steps to alignment*[94] Brookings considers the broad approaches of the U.S. and the EU to AI risk management, compares policy developments across eight key subfields, and discusses collaborative steps taken so far, especially through the above-mentioned EU-U.S. Trade and Technology Council. Of particular interest to the bill of rights team was the below application-based comparison table (Engler, 2023):

Application	**Examples**	**EU policy developments**	**U.S. policy developments**
AI for human processes/socioeconomic decisions	AI in hiring, educational access, and financial services approval	GDPR requires human in the loop for significant decisions. High-risk AI applications in Annex III of EU AI Act would need to meet quality standards, implement risk management system, and perform conformity assessment	AI Bill of Rights and associated Federal Agency Actions have created patchwork oversight for some of these applications.

Application	Examples	EU policy developments	U.S. policy developments
AI in consumer products	AI in medical devices, partially autonomous vehicles, and planes	EU AI Act considers AI implemented within products that are already regulated under EU law to be high risk and further would have new AI standards incorporated into current regulatory process.	Individual federal agency adaptations, such as by FDA for medical devices; DOT for automated vehicles; CPSC for consumer products
Chatbots	Sales or customer service chatbots on commercial websites	EU AI Act would require disclosure that a chatbot is an AI (i.e., not a human).	NA
Social media recommender & moderation systems	Newsfeeds and group recommendations on TikTok, Twitter, Facebook, or Instagram	EU Digital Services Act creates transparency requirement for these AI systems; also enables independent research and analysis	NA

Application	Examples	EU policy developments	U.S. policy developments
Algorithms on e-commerce platforms	Algorithms for search or recommendation of products and vendors on Amazon or Shopify	EU Digital Markets Act will restrict self-preferencing algorithms in digital markets. Individual anti-trust actions (e.g., against Amazon, and Google Shopping) to reduce self-preferencing in E-commerce algorithms and platform design.	NA
Foundations models/generative AI	Stability AI's Stable Diffusion and OpenAI's GPT-3	Draft proposals of the EU AI Act consider quality and risk management requirements.	NA
Facial recognition	Clearview AI, PimEyes, Amazon Rekognition	EU AI Act will include restrictions on remote facial recognition and biometric identification. EU Data Protection Authorities have fined facial recognition companies under GDPR.	NIST's AI Face Recognition Vendor Test program contributes efficacy and fairness information to the market for facial recognition software.

Application	Examples	EU policy developments	U.S. policy developments
Targeted advertising	Algorithmically targeted advertising on websites and phone applications	GDPR has fined Meta for using personal user data for behavioral ads. The Digital Services Act bans targeted advertising to children and certain types of profiling (e.g., by sexual orientation). It requires targeted ads have explanations and users have control over what ads they see.	Individual federal agency lawsuits have slightly curtailed some targeted advertising. This includes the DOJ and HUD, who successfully sued Meta for discriminatory housing ads and an FTC penalty against Twitter for using security data for targeted ads.

Table 1. Comparison of EU and U.S. AI risk management by application type. Engler, 2020

Engler observes that:[95]"

> "The EU and U.S. strategies share a conceptual alignment on a risk-based approach, agree on key principles of trustworthy AI, and endorse an important role for international standards. However, the specifics of these AI risk management regimes have more differences than similarities. Regarding many specific AI applications, especially those related to socioeconomic processes and online platforms, the EU and U.S. are on a path to significant misalignment.

NIST AI Risk Management Framework

On March 30, 2023 The National Institute of Standards and Technology(NIST), U.S. Department of Commerce, launched the Trustworthy and Responsible AI Resource Center(AIRC),[96] which will facilitate implementation of, and international alignment with, the NIST Artificial Intelligence Risk Management Framework(AI RMF, 2023).

As directed by the National Artificial Intelligence Initiative Act of 2020 (P.L. 116-283), NIST developed the *AI RMF* "to offer a voluntary resource to the organizations designing, developing, deploying, or using AI systems to help manage the many risks [to individuals, organizations, and society] of AI and promote trustworthy and responsible development and use of AI systems"[97,98]

> The Framework is intended to be voluntary, rights-preserving, non-sector specific, and use-case agnostic, providing flexibility to organizations of all sizes and in all sectors and throughout society to implement the approaches in the Framework. The AI RMF is

designed to equip organizations and individuals – AI actors – with approaches that increase the trustworthiness of AI systems.

In the NIST AI RMF[99], risk refers to:

> The composite measure of an event's probability of occurring and the magnitude or degree of the consequences of the corresponding event. The impacts, or consequences, of AI systems can be positive, negative, or both and can result in opportunities or threats (Adapted from: ISO 31000:2018). When considering the negative impact of a potential event, risk is a function of 1) the negative impact, or magnitude of harm, that would arise if the circumstance or event occurs and 2) the likelihood of occurrence (Adapted from: OMB Circular A-130:2016).
>
> Negative impact or harm can be experienced by individuals, groups, communities, organizations, society, the environment, and the planet.

The AI RMF[100] provides examples of potential harms (and harmed parties) that can be related to AI systems.

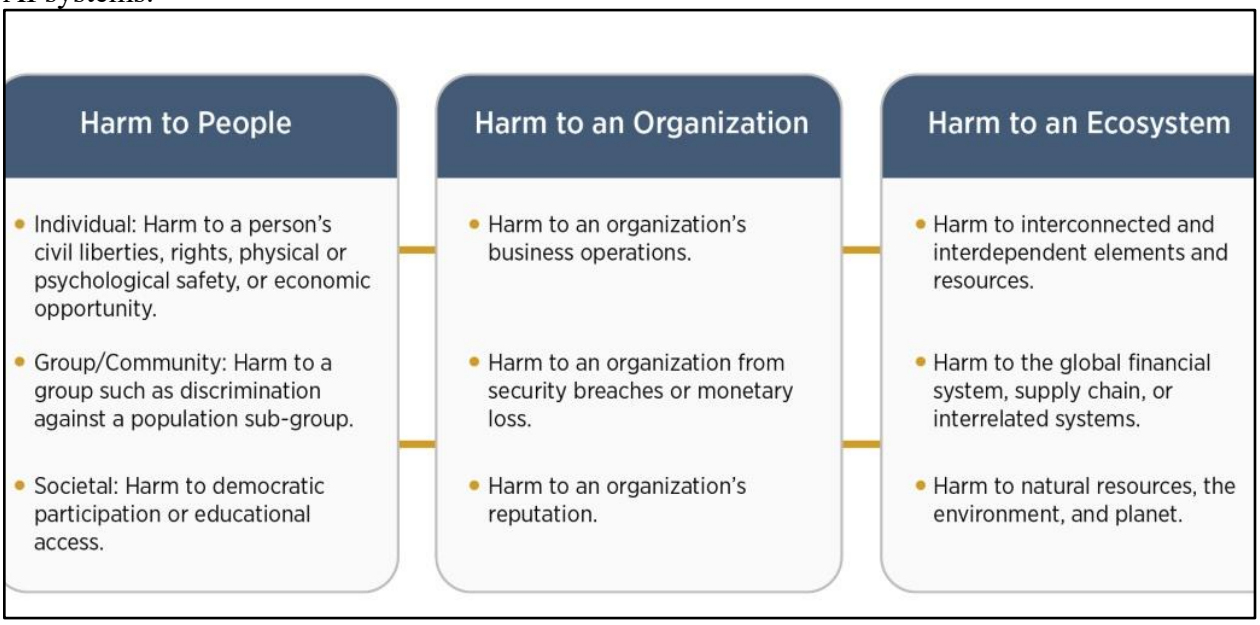


Figure 8. AI RMF NIST

Appendix A in the AI RMF[101] offers detailed descriptions of AI Actor tasks including those associated with AI Design, Development, and Deployment separate from "Operation and Monitoring", and "Test, Evaluation, Verification and Validation" (TEVV) , as well as end users, affected individuals/communities, third party entities, and the general public.

Ten priorities to advance the AI RMF and its use are detailed in the below *Roadmap for the NIST AI RMF*[102] starting with standards alignment (e.g., ISO/IEC 5338, ISO/IEC 38507, ISO/IEC 22989, ISO/IEC 24028, ISO/IEC DIS 42001, and ISO/IEC NP 42005) and illustrating how the Trustworthy and Responsible AI Resource Center(AIRC) is positioned to offer support:

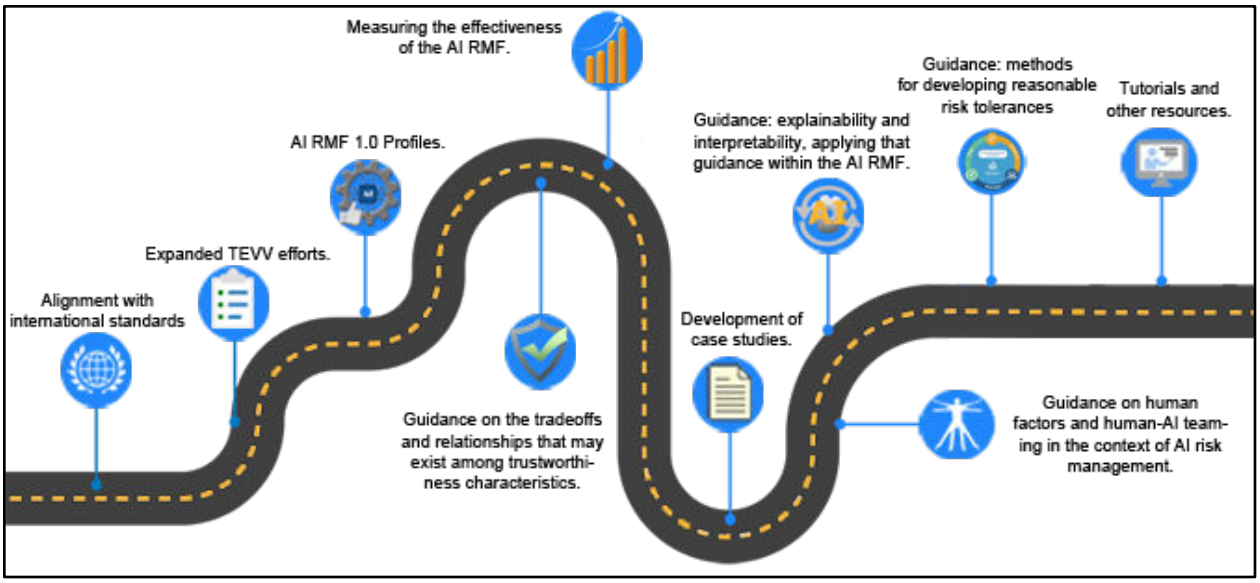


Figure 9. *Roadmap for the NIST AI RMF*

Blueprint for an AI Bill of Rights[103]

In October 2022 the *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People* was published by the White House Office of Science and Technology Policy. This framework was released one year after OSTP announced the launch of a process to develop "a bill of rights for an AI-powered world" and much public engagement. The Blueprint is a set of far-reaching goals aimed at averting harms caused by the rise of artificial intelligence systems, including guidelines for how to protect people's personal data and limit surveillance[104].

The AI Bill of Rights declares five principles that should guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence. This offers a promising vision of community-oriented equity unique amongst peer countries[105].

> "This is the Biden-Harris administration really saying that we need to work together, not only just across government, but across all sectors, to really put equity at the center and civil rights at the center of the ways that we make and use and govern technologies," said Alondra Nelson, deputy director for science and society at the White House Office of Science and Technology Policy[106].

The framework is accompanied by a technical companion—a handbook for those seeking to incorporate these protections into policy and practice, including detailed steps toward actualizing these principles in the technological design process.

The Blueprint on its own is non-binding, and without concrete legislation, some anticipate the private sector may ignore it[107] while others "voice concerns that the non-binding guidelines could lead to stifling regulation concerning artificial intelligence"[108]. Nandita Sampath, a policy analyst focused on algorithmic accountability and bias at Consumer Reports commended the report as a step in the right direction saying[109]:

> American consumers need protections from these complex systems that are used to evaluate them without their knowledge and without a meaningful explanation as to why they arrived at certain decisions. Consumers deserve to know why an automated decision system denies them an opportunity, particularly in

> sensitive cases. We hope to see more federal and state agencies implementing these recommendations and also ask Congress to codify these recommendations into law.

What happens next? As the US codifies policy and law from the blueprint, with the recent passage of the EU AI Act (EU AIA), United States companies will be subject to the EU AIA when they export to the EU. There are likely to be a multiplicity of cases similar to the OpenAI and ChatGPT situation in Italy. Lilian Edwards writes that "An even more profound issue is that the EU standards may by reciprocity or osmosis become global standards (e,g, via ISO). So just as the [EU AI] Act may become a global regulatory model, these standards may also become global rules in effect[110]." In the meantime, there are several more U.S. Acts and Policies, some binding and codifying, others proposed, and still others unbinding yet influencing policy and practice of AI in the US. The AI Index Report 2024[111] summarizes that:

> The number of AI related regulations in the U.S. has risen significantly in the past year and over the last five years. In 2023, there were 25 AI-related regulations, up from just one in 2016. Last year alone, the total number of AI-related regulations grew by 56.3%.

Some highlights are featured in the list below.

Other AI related US Acts and Policies

- AI in Government Act 2020 (DIVISION U, TITLE I) [112]
  This codifies into law the GSA Artificial Intelligence (AI) Center of Excellence (CoE)[113], which was launched in 2019. It further calls on the Office of Management and Budget to provide guidance for agency use of AI, and for the Office of Personnel Management to update the occupational series for AI for Federal employees.
- National AI Initiative Act (NAII 2021) [114]
  The mission of the National AI Initiative is to ensure continued U.S. leadership in AI research and development, lead the world in the development and use of trustworthy AI in the public and private sectors, and prepare the present and future U.S. workforce for the integration of AI systems across all sectors of the economy and society and became law on January 1, 2021 as part of the National Defense Authorization Act for FY 2021, HR 6395, Division E.
- Four Principles of Explainable AI (NIST 2021) [115]
  In this paper NIST introduces four principles that comprise fundamental properties for explainable AI systems. NIST recognizes that not all AI systems may require explanations. However, for those AI systems that are intended or required to be explainable, they propose that those systems adhere to the following four principles:
  - Explanation: A system delivers or contains accompanying evidence or reason(s) for outputs and/or processes.
  - Meaningful: A system provides explanations that are understandable to the intended consumer(s).
  - Explanation Accuracy: An explanation correctly reflects the reason for generating the output and/or accurately reflects the system's process.
  - Knowledge Limits: A system only operates under conditions for which it was designed and when it reaches sufficient confidence in its output.
- National AI Research Resource (NAIRR) (2023) An Implementation Plan for a National Artificial Intelligence Research Resource Jan 24, 2023
- Government-wide policy Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence[116] (OMB March 28, 2024)

- Food and Drug Administration (FDA): Clinical Decisions Support Software (2022)
- H.R.5628 - Algorithmic Accountability Act / S.2892 (2023-2024))- Proposed
- H.R. 8818 American Privacy Rights Act (2023-2024)- Proposed
- Assessing Adverse Impact in Software, Algorithms, and Artificial Intelligence Used in Employment Selection Procedures Under Title VII of the Civil Rights Act of 1964[117] (05-18-2023)

  A Technical assistance document, not new policy, which applies principles already established in the Title VII statutory provisions. The contents of this publication do not have the force and effect of law and are meant only to provide clarity to the public regarding existing requirements under the law. Here the EEOC provides guidance on the use of AI in employment selection procedures emphasizing that:

  To ensure fair and unbiased AI usage in the workplace, employers should consider the following recommendations:

  1. Conduct a job analysis by identifying the essential functions, duties, skills, qualifications, and competencies required for each position. Ensure that the AI system aligns with the identified criteria to avoid any implicit bias that already may be present in an industry.
  2. Validate the AI system by testing and evaluating the AI system to ensure its accuracy, validity, and freedom from bias for the intended purpose and target population. Conduct disparate impact analyses to help identify any marginalized groups and update any AI systems regularly to address any issues.
  3. Involve human decision-makers who can review and override the AI system's results if necessary. Provide training and guidance to these decision-makers on how to use the system properly and ethically.

State level legislation on AI in the U.S.

In recent years, Alabama, Colorado, Illinois, and Vermont have passed bills creating a commission, task force, or oversight position to evaluate the use of AI in their states and make recommendations regarding its use. In the 2023 legislative session, at least 25 states, Puerto Rico and the District of Columbia introduced artificial intelligence bills, and 14 states and Puerto Rico adopted resolutions or enacted legislation[118]. 2024 is shaping up to have been a busy year as well, at least 45 states, Puerto Rico, the Virgin Islands and Washington, D.C., introduced AI bills, and 31 states, Puerto Rico and the Virgin Islands adopted resolutions or enacted legislation.[119]

California

ACR-215 23 Asilomar AI Principles[120]

Adopted in California (2018), the Asilomar AI Principles offer guiding values for the development of artificial intelligence and of related public policy.

In 2023 California's Governor Newsom signed an executive order to prepare California for the progress of artificial intelligence with a focus on shaping the future of ethical, transparent, and trustworthy GenAI.[121] In Fall 2024 the Governor has signed over a dozen concerning GenAI technology[122] including:

AB 1008 by Assemblymember Rebecca Bauer-Kahan (D-Orinda) – Clarifies that personal information under the California Consumer Privacy Act (CCPA) can exist in various formats, including information stored by AI systems.

AB 1831 by Assemblymember Marc Berman (D-Menlo Park) – Expands the scope of existing child pornography statutes to include matter that is digitally altered or generated by the use of AI.

AB 1836 by Assemblymember Rebecca Bauer-Kahan (D-Orinda) – Prohibits a person from producing, distributing, or making available the digital replica of a deceased personality's voice or likeness in an expressive audiovisual work or sound recording without prior consent, except as provided.

AB 2013 by Assemblymember Jacqui Irwin (D-Thousand Oaks) – Requires AI developers to post information on the data used to train the AI system or service on their websites.

AB 2355 by Assemblymember Wendy Carrillo (D-Los Angeles) – Requires committees that create, publish, or distribute a political advertisement that contains any image, audio, or video that is generated or substantially altered using AI to include a disclosure in the advertisement disclosing that the content has been so altered.

AB 2602 by Assemblymember Ash Kalra (D-San Jose) – Provides that an agreement for the performance of personal or professional services which contains a provision allowing for the use of a digital replica of an individual's voice or likeness is unenforceable if it does not include a reasonably specific description of the intended uses of the replica and the individual is not represented by legal counsel or by a labor union, as specified.

AB 2655 by Assemblymember Marc Berman (D-Menlo Park) – Requires large online platforms with at least one million California users to remove materially deceptive and digitally modified or created content related to elections, or to label that content, during specified periods before and after an election, if the content is reported to the platform. Provides for injunctive relief.

AB 2839 by Assemblymember Gail Pellerin (D-Santa Cruz) – Expands the timeframe in which a committee or other entity is prohibited from knowingly distributing an advertisement or other election material containing deceptive AI-generated or manipulated content from 60 days to 120 days, amongst other things.

AB 2876 by Assemblymember Marc Berman (D-Menlo Park) – Require the Instructional Quality Commission (IQC) to consider AI literacy to be included in the mathematics, science, and history-social science curriculum frameworks and instructional materials.

AB 2885 by Assemblymember Rebecca Bauer-Kahan (D-Orinda) – Establishes a uniform definition for AI, or artificial intelligence, in California law.

AB 3030 by Assemblymember Lisa Calderon (D-Whittier) – Requires specified health care providers to disclose the use of GenAI when it is used to generate communications to a patient pertaining to patient clinical information.

SB 896 by Senator Bill Dodd (D-Napa) – Requires CDT to update report for the Governor as called for in Executive Order N-12-23, related to the procurement and use of GenAI by the state; requires OES to perform a risk analysis of potential threats posed by the use of GenAI to California's critical infrastructure (w/high-level summary to Legislature); and requires that the use of GenAI for state communications be disclosed.

SB 926 by Senator Aisha Wahab (D-Silicon Valley) – Creates a new crime for a person to intentionally create and distribute any sexually explicit image of another identifiable person that was created in a manner that would cause a reasonable person to believe the image is an authentic image of the person depicted, under circumstances in which the person distributing the image knows or should know that distribution of the image will cause serious emotional distress, and the person depicted suffers that distress.

SB 942 by Senator Josh Becker (D-Menlo Park) – Requires the developers of covered GenAI systems to both include provenance disclosures in the original content their systems produce and make tools available to identify GenAI content produced by their systems.

SB 981 by Senator Aisha Wahab (D-Silicon Valley) – Requires social media platforms to establish a mechanism for reporting and removing "sexually explicit digital identity theft."

SB 1120 by Senator Josh Becker (D-Menlo Park) – Establishes requirements on health plans and insurers applicable to their use AI for utilization review and utilization management decisions, including that the use of AI, algorithm, or other software must be based upon a patient's medical or other clinical history and individual clinical circumstances as presented by the requesting provider and not supplant health care provider decision making.

SB 1288 by Senator Josh Becker (D-Menlo Park) – Requires the Superintendent of Public Instruction (SPI) to convene a working group for the purpose of exploring how artificial intelligence (AI) and other forms of similarly advanced technology are currently being used in education.

SB 138

1 by Senator Aisha Wahab (D-Silicon Valley) – Expands the scope of existing child pornography statutes to include matter that is digitally altered or generated by the use of AI.

Colorado

Colorado's new *Consumer Protections for Artificial Intelligence* law (SB24-205)[123] enacted in May 2024 effective on and after Feb 1, 2026 requires a person doing business in this state, including a deployer or other developer, that deploys or makes available an artificial intelligence system that is intended to interact with consumers to ensure disclosure to each consumer who interacts with the artificial intelligence system that the consumer is interacting with an artificial intelligence system. It requires companies using AI to help make consequential decisions for Americans to annually assess their AI for potential bias; implement an oversight program within the company; tell the state attorney general if discrimination was found; and inform to customers when an AI was used to help make a decision for them, including an option to appeal.

Indiana

Indiana's governor Holcomb signed two AI related bills in 2024: (House Bill 1133) that requires disclaimers on any political materials that use AI generated content and (SB 150) which establishes a Task Force for studying the landscape of AI use and mandating the development of cybersecurity policies for public entities.

U.S. Municipalities govern AI in hiring and its use by Law Enforcement Agencies

New York City (NYC)

Effective July 5, 2023 the New York City Local Law 144[124] **regulates AI and automated decision-making tools for employee hiring and promotions**. The bill was signed into law on

December 11, 2021 and prohibits employers and employment agencies from using an automated employment decision tool in NYC unless they ensure a bias audit was done annually and provide required notices. Job candidates who are New York City residents must receive notice that the employer or employment agency uses an AEDT. Employers and employment agencies subject to the law must publish:

- The distribution date of the AEDT.
- A summary of the results of the most recent bias audit. AND
- The date of the most recent bias audit of the AEDT. AND
- The source and explanation of the data used to conduct the bias audit. AND
- The number of individuals the AEDT assessed that fall within an unknown category. AND
- The number of applicants or candidates, the selection or scoring rates, as applicable, and the impact ratios for all categories.

Baltimore, Maryland

Act 21-038 – Surveillance Technology in Baltimore

This Act introduced 1/11/21, approved: 6/14/21, signed: 8/16/21 and codified in the city's police ordinances as Article 19 § 18.2 prohibits use of face surveillance technology and prevents the Baltimore city government from obtaining or contracting with another entity that provides certain face surveillance technology, prohibits any person in Baltimore City from obtaining or using face surveillance technology, and requires the Director of Baltimore City Information and Technology to submit an annual report to the Mayor and City Council regarding the use of surveillance by the Mayor and City Council.

Bellingham, Washington

Initiative Bill 2021-19 - use of facial recognition technology and predictive policing technology

This bill adopted 8/2/2021, passed 11/10/21 and codified 4/5/2022 in the Bellingham Municipal code as Chapter 2.24 article II prohibits government use of facial recognition and predictive policing technologies. Bellingham residents voted to prohibit the city from acquiring or using facial recognition technology or contracting with a third party to use facial recognition technology on the city's behalf. The measure also restricts use of illegally obtained data in policing or trials.

Canada

- Canada's Directive on Automated Decision-Making - In Effect
- Canada's Artificial Intelligence and Data Act (AIDA) – at consideration

AIDA is an example of an Internationally aligned AI Governance Strategy. Its key definitions and concepts were designed to coincide with evolving international norms– including the previously covered EU AI Act, OECD AI Principles and US NIST Risk Management Framework (RMF) – while integrating with existing Canadian legal frameworks.

Latin America and the Caribbean

The CAF-Development Bank of Latin America and UNESCO signed a letter of intent in 2022 to work together on the implementation of the Recommendation on the Ethics of Artificial Intelligence (AI) in Latin America and the Caribbean[125].

On March 10, 2023 the Montevideo Declaration on Artificial Intelligence[126] was presented at the Latin American Meeting on Artificial Intelligence KHIPU 2023, to urge governments and companies to make AI developments put at the service of the people, reflecting the particularities and problems of Latin America. In the declaration, more than 100 specialists call for artificial intelligence technologies to serve the common good, improving people's quality of life and aligning with democratic principles and human rights, concluding with the observation that:“There is no social value in technologies that simplify tasks for a few people, generating a hi gh risk to the dignity of many others, limiting their development opportunities, their access to res ources, and their rights. “

On 9 August 2024, the “Cartagena de Indias Declaration for Governance, the construction of AI ecosystems and the promotion of AI education in an Ethical and Responsible manner in Latin America and the Caribbean”[127] was signed by the governments of 16 countries of the region (Argentina, Brazil, Chile, Colombia, Costa Rica, Curacao, Ecuador, Guatemala, Guyana, Honduras, Panama, Paraguay, Peru, Dominican Republic, Surinam, and Uruguay). The Declaration states the commitment to promote AI governance frameworks and ecosystems for a safe, inclusive, ethical, and responsible development of AI.

At the country level, Argentina, Brazil, Chile, Colombia, Mexico, Peru, Dominican Republic, and Uruguay have released national AI strategies and/or have bills under review. The 2024 Latin American Artificial Intelligence Index[128] evaluates 19 of the region’s countries, measuring each one’s level of preparation in relation to artificial intelligence (AI). On a maximum of 100 points, Chile obtained the top spot in the report’s country ranking, with a score of 73.07, followed by Brazil (69.30) and Uruguay (64.98)[129]. Other countries in the region follow in the “adopters” category. These include Argentina (55.77), Colombia (52.64) and Mexico (51.40). The report identifies the top three as pioneers in AI noting that they have:

> *not only made progress on implementing AI-based technologies, but are also orienting their national strategies towards the consolidation and expansion of these technologies in all the sectors of their economy and society. Furthermore, they have a conducive environment that fosters the research, development and adoption of technologies, promoting innovation and the application of AI.*

A translated table of Strategy Indicators from the report is shared below:

On a more somber note, according to the study, even though the region has doubled the percentage of AI talent concentration in the workforce on average over the last eight years, no country has reached the levels that countries from the Global North had at the start of the same period.

Brazil

Draft law No. 2,338 of 2023123 is the result of the work done by the Commission of Jurists on AI and created by the Brazilian Senate, is the most comprehensive proposal. It establishes general rules for the development, deployment, and responsible use of AI systems.

Chile

Chile is the first country in the world to apply and complete UNESCO's Readiness Assessment methodology[130]. This milestone has provided valuable insights that will serve as a foundation not only for Chile but also for other nations pursuing similar paths. On May 2, 2024 Chile launched an updated National AI Policy and action plan, and on May 7, 2024 the president sent a bill regarding regulation to the country's Chamber of Deputies. The AI Regulation Bill is aligned with UNESCO's recommendations on ethics of AI covered earlier in this document is aimed at regulating and promoting the ethical and responsible development of this technology.

Colombia

Colombia was first in Latin America to publish a standalone AI ethics framework.[131] This document "seeks to provide a soft law guide of recommendations and suggestions to Public Entities, to address the formulation and management of projects that include the use of Artificial Intelligence (AI)" and " brings together international practices that are not mandatory for any of the entities but that are presented as recommendations". Colombia has also been investing in new AI policy initiatives, including an AI strategy, and has formed an international AI council.

Argentina

In Argentina, changes in political leadership have resulted in the new administration putting the previously released national AI strategy on hold but it went ahead with Resolution 90/2021, which mentions the need to address AI adoption in the public sector and skills development. Several draft legislative initiatives were presented in 2023[132] including Draft law 2505-D-2023: "Legal framework for the regulation of the development and use of Artificial Intelligence" which seeks to regulate "an ethical development and use of AI that will benefit society as a whole." Draft law No and 3161-D-2023: to create the "Federal Council of Artificial Intelligence", it proposes "a hierarchical and democratic structure" to guarantee, promote and strengthen Argentina's potential in the field of AI".

Buenos Aires City Plan

In Buenos Aires, Argentina there is now a city plan[133] for AI that "seeks to generate a positive impact in all areas of citizens' lives through the development and use of artificial intelligence."

Mexico

Mexico began work on an AI strategy in 2017[134] and its strategy was among the first to be presented in the region when the UK Embassy in Mexico, with support from the Office of the Mexican President, commissioned Oxford Insights to draft a national AI plan which was published as a white paper in June 2018 entitled: "Towards An Ai Strategy In Mexico: Harnessing the AI Revolution."[135] Then, an IA2030Mx citizen coalition was created, founded by nine institutions from all sectors. The IA2030Mx effort addressed the need to have a national, multidisciplinary, multisectoral and collaborative exercise to develop an AI action plan which

could serve government, academia, civil society and industry. This output was presented as the Mexican National Artificial Intelligence Agenda[136] in 2020.

Australasia

Australia

In Australia there are several efforts of relevance on AI governance and cross-domain information sharing. Here we describe a few:

The Australian computer society (ACS) convenes working groups on data sharing and AI ethics. A data taskforce was created to address the overarching challenge of developing ethical and privacy-preserving frameworks that support automated data sharing to facilitate the creation and deployment of smart services. This framework will seek to address technical, regulatory, and authorising frameworks. The intention is to identify, adopt, adapt, or develop frameworks for data governance, privacy protection, and practical data sharing that facilitates smart service creation and cross-jurisdictional data sharing between governments. Three publications arising from this work will be of interest to this audience: 1) ***Data Sharing Frameworks*** *(ACS, 2017)*[137] 2) ***Privacy in Data Sharing: A Guide for Government and Business*** *(ACS, 2018)*[138]. This paper describes a framework for privacy-preserving data sharing which builds on the 2017 paper, expanding the concept of a Personal Information Factor (PIF) and introducing a Data Safety Factor with recommendations for threshold settings. This paper further develops the concept of a quantified 'Five Safes' data analytics framework and briefly examines the implications of such frameworks when artificially intelligent algorithms are used to analyse data. And 3) ***Privacy-preserving data sharing frameworks: People Projects, Data and Output***[139]. This paper tackles the challenge of balancing data sharing benefits with citizens' privacy rights.

Another ACS collaboration takes as its charge "***Making the Invisible Visible"***, and this group is discussing ways of training data stewards to 'make the invisible visible' by remaining alert to who (and what) is missing, under-represented or mis-represented in data in any given context. They aspire to enrich the application of FAIR principles in the "making" of data.

Particular to AI, In 2019 Australia adopted an Artificial Intelligence Ethics Framework[140], a set of voluntary ethics principles to ensure AI applications are safe, secure and reliable. The Australian AI Ethics principles are:

- *Human, societal and environmental wellbeing:* AI systems should benefit individuals, society and the environment.
- *Human-centred values:* AI systems should respect human rights, diversity, and the autonomy of individuals.
- *Fairness:* AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.
- *Privacy protection and security:* AI systems should respect and uphold privacy rights and data protection, and ensure the security of data.
- *Reliability and safety:* AI systems should reliably operate in accordance with their intended purpose.
- *Transparency and explainability:* There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI, and can find out when an AI system is engaging with them.

- *Contestability:* When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.
- *Accountability:* People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

In 2023, just four years later, the Australian Government opened a period of open feedback on responsible AI, informing its public that as "Many other countries are already looking at new governance arrangements to ensure AI is used responsibly" they are seeking input on "what else the Australian Government can do to support the safe and responsible use of AI." Which could be through either or both: "voluntary approaches, like tools, frameworks and principles" or "enforceable regulatory approaches, like laws and mandatory standards". The period of input closed at the end of July 2023.

In June 2024, the Australian government published "National framework for the assurance of artificial intelligence in government"[141] developed jointly by Australian federal, state and territory governments. It aims to establish a consistent approach for the safe and responsible use of AI in government, based on Australia's AI Ethics Principles. The framework is organized around eight core ethics principles, including human, societal, and environmental wellbeing, human-centered values, fairness, privacy protection, reliability, safety, transparency, contestability, and accountability. It emphasizes clear governance, data quality, transparency, human oversight, and continuous improvement.

Australian Government's policy for the responsible use of AI in government[142] became effective from 1 Sep 2024. The policy aims to promote AI's use responsibly, ensuring its safe, ethical, and ethical use. The policy is mandatory for Non-corporate Commonwealth entities and takes effect on 1 September 2024. It aims to address public concerns about AI use, such as transparency, accountability, and data protection. The policy is designed to be flexible and evolve with AI advancements and regulatory changes, ensuring its relevance in a rapidly changing technological landscape. It applies to all Non-Corporate Commonwealth Entities under the Public Governance, Performance and Accountability Act 2013. The policy follows the "Enable, Engage, Evolve" framework, requiring agencies to designate accountable officials for AI implementation, provide AI training to staff, engage responsibly, disclose AI adoption strategies, ensure ethical use, and regularly review their AI policies to ensure compliance with evolving governance and risk management frameworks.

The Australian government has proposed ten mandatory guardrails for artificial intelligence in high-risk settings as it moves to curb dangers posed by the nascent technology. These are articulated in the *Proposals paper for introducing mandatory guardrails for AI in high-risk settings*[143] published September, 2024 and aim to support safe and responsible AI in Australia by addressing the risks associated with AI, especially in high-risk contexts. The paper proposes three regulatory approaches for regulating AI: adapting existing frameworks, introducing framework legislation, or creating a new AI-specific act. It also emphasizes interoperability and global standards, aligning with international standards like the European Union's AI Act and Canada's AI and Data Act. The Australian Government seeks feedback from various stakeholders to refine the definitions of high-risk AI and the proposed mandatory guardrails. The goal is to create a regulatory framework that balances innovation with safety, ensuring AI is used responsibly in high-risk settings while providing clarity and certainty for businesses operating in Australia.

A recent nationally representative survey[144] has revealed Australians are deeply concerned about the risks posed by artificial intelligence with 80% of Australians identifying AI risk is a global priority.

A standalone Artificial Intelligence Act similar to EU AI Act is being considered after more than 12 months of consultation. The Australian government plans to adopt a risk-based approach to AI regulation, focusing on high-risk contexts and allowing lower-risk forms to flourish. The government plans to develop mandatory requirements through public and industry consultation and an advisory body, focusing on testing and audit, transparency, and accountability. Interim measures to ensure AI safety include developing an AI safety standard and options for watermarking AI-generated content. The government's response to the Paper is unclear in areas such as AI-related policy, legislative, or budgetary matters. The introduction of AI-specific legislation was a topic of intense debate in submissions to the Paper. The government's response also raises questions about how to future-proof AI regulation, as comprehensive regulation is difficult due to the constantly evolving nature of the technology. There is no mention of a dedicated regulator, raising concerns about whose jurisdiction AI will fall under, and the potential inconsistency between regulatory approaches across sectors. The government's response to the Paper raises questions about the best way to design a regulatory regime that provides certainty and consistency while being flexible and adaptable to changes in technology.

New Zealand

There are no AI specific laws in New Zealand (NZ), but there is a rich ecosystem of AI policy research and collaboration. The AI Forum founded in 2017 as a non-governmental organization brings together New Zealand's AI community, including innovators, end users, investors, regulators, researchers, educators, entrepreneurs and interested public. They make a "curated collection of AI governance resources available to equip members with knowledge and the guidance to address the most critical issues in AI systems, with the ultimate goal of maximising value and fostering trust and confidence in AI technologies while minimising risks and negative societal impacts"[145].

Meanwhile, most government agencies have signed up to NZ's 2020 Algorithm Charter [146]. Signatories to the Algorithm Charter have agreed to apply certain principles in how they use algorithms, especially in designing access to public services but it doesn't address newer technologies such as the LLMs.

The Algorithm Charter is part of a wider ecosystem works together with existing tools, networks and research, including[147]:

- Principles for the Safe and Effective Use of Data and Analytics (Privacy Commissioner and Government Chief Data Steward, 2018)
- Government Use of Artificial Intelligence in New Zealand (New Zealand Law Foundation and Otago University, 2019)
- Trustworthy AI in Aotearoa – AI Principles (AI Forum New Zealand, 2020)
- Open Government Partnership, an international agreement to increase transparency.
- Data Protection and Use Policy (Social Wellbeing Agency, 2020)
- Privacy, Human Rights and Ethics Framework (Ministry of Social Development).

In July 2023, New Zealand's Government Chief Digital Officer (GCDO) released early, interim, very practical Initial Guidance on Generative AI in the public service.[148] This advice from joint system leads and its attached A3 are "intended to support agencies to make more informed decisions about using GenAI, balancing benefits and risks[149]". The guidance emphasizes the benefits of GenAI, such as improved efficiency, service delivery, cyber defense, innovation, and policy development. Key recommendations include not using GenAI for sensitive data, avoiding business-critical information, and exercising caution with both free and paid GenAI tools. The guidance also encourages agencies to develop their own GenAI policies and standards, emphasize privacy and security, and encourage continuous learning.

Page two of the handout (aka A3) offers a concise crosswalk of NZ policy concerns compared to US, EU and other nations[150].

Comparison of global approaches to AI risk management

	EU	US	Other	NZ
Dedicated AI legislation?	AI Act in in the works and expected to become law later in the year; legislation will require creation of standards Existing GDPR covers algorithmic decision making and targeted ads, and Digital Services and Digital Market Acts target transparency and fair market competition	White House has produced AI Bill of Rights and other guidelines; these have no mechanism to compel compliance Algorithm Accountability Act before both chambers of Congress; not clear whether has political legs to progress Some state legislatures have passed algorithm accountability legislation; NYC has imposed requirements on the use of algorithms in hiring/promotion Some existing legislation has implications for AI e.g. around fair trading practices, anti-discrimination	China: Has taken a "vertical" approach; individual pieces of legislation on algorithmic recommendations, deep synthesis, and generative AI. Legislation around generative AI and deepfakes has created a compulsory registry; it is expected this will be a part of future legislation for different AI Canada: Digital Charter Implementation Act (Bill C-27) passed second reading in lower house; long process to go before adopted Australia: Consultation process for creating legislation launched June 2023, including rapid evidence review and a report to inform public submissions UK: no dedicated legislation. AI white paper describes possible future regulation in context of wider strategic approach to AI.	No dedicated legislation. Some existing legislation has implications for AI e.g. Privacy Act, Human Rights Act, as well as Te Tiriti o Waitangi. Algorithm Charter has been adopted by most government agencies.
AI for human processes/socioeconomic decisions *AI in hiring, educational access, and financial services approval*	GDPR requires human in the loop for significant decisions. "High-risk" AI applications AI Act would need to meet quality standards, implement risk management system, and perform conformity assessment	AI Bill of Rights and associated Federal Agency Actions have created patchwork oversight for some of these applications. Notable gap even in Algorithm Accountability Act is that some sectors are out of scope, including public services NYC requires impact assessment of hiring and promotion decisions that involve algorithms; in practice, these requirements are poorly defined, and deadlines have been repeatedly pushed back as a result	Canada: Directive on Automated Decision Making applies to government services and imposes requirements around transparency when there are AI components in decision making UK: Piloting Algorithmic Transparency Reporting Standard in some parts of government	Algorithm Charter has been adopted by most government agencies
AI in consumer products *AI in medical devices, partially autonomous vehicles, and planes*	AI Act considers AI implemented within products that are already regulated under EU law to be high risk; new AI standards to be incorporated into current regulatory process.	Individual federal agency adaptations, such as by FDA for medical devices; DOT for automated vehicles; CPSC for consumer products		Existing laws such as Consumer Guarantees Act, Human Rights Act, Privacy Act apply as relevant
Chatbots	AI Act would require disclosure that a chatbot is an AI (i.e., not a human).	California BOT Act makes it an offense to pretend to be a person to sell products of influence elections.		Existing laws such as Fair Trading Act 1986, Human Rights Act, Privacy Act 2020apply as relevant, as does sector-specific regulation (e.g. in the financial services sector)
Social media recommender algorithms *Newsfeeds and group recommendations on social media*	Digital Services Act creates transparency requirement; also enables independent research and analysis		China: Consumer must be informed that an algorithm has been used	NZ Code of Practice for Online Safety and Harms applies as relevant if a company has adopted it
Algorithms on e-commerce platforms *Algorithms for search or recommendation of products and vendors*	Digital Markets Act restricts self-preferencing algorithms in digital markets		China: Prohibits use of personal info in price setting	
Foundation models/generative AI *DALL-E; ChatGPT*	Draft proposals of the EU AI Act consider quality and risk management requirements.		China: Output must be true, unbiased, and conform with state ideology; developers responsible for all content produced (even by a different end user). Developers responsible for ensuring training data are unbiased, objective and accurate	Privacy Act applies; Office of the Privacy Commissioner released guidance on this in May 2023.
Facial recognition	AI Act will include restrictions on remote facial recognition and biometric identification. Data Protection Authorities have fined facial recognition companies under GDPR.	NIST's AI Face Recognition Vendor Test program contributes efficacy and fairness information to the market for facial recognition software.		Office of the Privacy Commissioner is exploring a code of practice on biometrics.
Targeted advertising *Algorithmically targeted advertising on websites and phone applications*	Meta has been fined under GDPR for using personal user data for behavioural ads. The Digital Services Act bans targeted advertising to children and certain types of profiling (e.g., by sexual orientation). It requires targeted ads have explanations and users have control over what ads they see.	Individual federal agency lawsuits have slightly curtailed some targeted advertising. This includes the DOJ and HUD, who successfully sued Meta for discriminatory housing ads and an FTC penalty against Twitter for using security data for targeted ads.		Existing law such as Unsolicited Electronic Messages Act 2007 and Privacy Act 2020 apply as relevant.

Adapted from Engler 2023 The EU and U.S. diverge on AI regulation: A transatlantic comparison and steps to alignment. Brookings Institution. Available at https://www.brookings.edu/research/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/

Table 2 - GCDO, NZ

In October 2023, NZ government released guidelines on AI in healthcare for Aotearoa New Zealand[151] highlighting the potential benefits of AI integration, focusing on improving healthcare delivery, addressing health inequities, and fostering innovation. Key points include ethical principles, safe, effective, and responsible AI, and the need for robust governance, data security, and a solid legal framework. AI can be used in various domains, including back-office operations, image analysis for diagnostics, medical research, and predictive and generative

systems. The 2023 Therapeutic Products Act will play a key role in regulating AI used in healthcare, with strict criteria for safety and performance. The guidelines provide 22 recommendations, categorized into eight themes: mapping the healthcare and AI landscape in New Zealand, maintaining human interaction in healthcare, ensuring public trust in AI technologies, and addressing data governance, privacy, and Māori data sovereignty.

Discussion

Jurisdictional Pluralism and Cross-National Layering

As efforts like the European Open Science Cloud, and membership organizations like the Research Data Alliance navigate and claim rights, there is a need to have informed conversations about balancing and navigating jurisdictional pluralism and the opportunities of influencing cross-layered rights. Over time, cross-national layers have the potential to transform domestic institutions and in turn global rules[152]. Literature tagging and matrix approaches are aids for observing jurisdictionally layered use cases and scenarios to better understand how overlapping rights, misalignments and gaps protect or leave vulnerable the interests of individuals and communities in an increasingly AI influenced world. Indeed, the NZ example above adapted from Engler on EU/US policy inspires making.

Cross-Policy Layering and Regulation of AI in Academia - Two Examples

AI governance decisions in academia take into account the policies of researchers' affiliated institution(s), and funder policies, in addition to any jurisdictionally cross layered regulations. The University of Notre Dame and the University of California in the United States offer illustrative examples of how model centric/artificial Intelligence research work can face extra but disparate compliance burdens through signatory commitments as well as through national, state, or local law.

For example, researchers, staff and students at University of Notre Dame near South Bend, Indiana in the United States abide by the relevant US jurisdictional policies and international collaborators' policies which all overlay one another, but in addition the University is also a signatory of the *Rome Call for AI Ethics* (2022)[153,154,155] and has issued its own *Generative AI policy for students*[156] and *AI recommendations for Instructors*.[157] The University of Notre Dame also has a public *Statement on Generative AI and Academic Integrity*[158] and an *AI Policies and Guidelines resource*[159].

Researchers' work at University of California is informed by the Asilomar AI Principles[160,161] (because the State of California is a signatory) and by other state laws related to AI as well. The University of California has made public the *UC Responsible AI Principles*[162] and offers a resource page on *Applicable law and UC Policy*[163] listing 33 relevant policies and laws at state, national, and international levels. This landscape is about to get more complicated. As described earlier in this document, California's governor has signed 17 bills in Fall 2024 alone covering deployment and regulation of AI technology[164]. Comparatively, in University of Notre Dame's home state of Indiana only two were signed over the same year. The complicated pluralistic policy and regulatory landscape faced by research collaborators even within the same country as

we observe in the above examples, coupled with cost of training and running artificial intelligence models partially explains how industry outcompetes academia in AI research.

In The Artificial Intelligence Index Report 2024 it is reported that:[165]

> Until 2014, academia led in the release of machine learning models. Since then, industry has taken the lead. In 2023, there were 51 notable machine learning models produced by industry compared to just 15 from academia (The AI Index 2024 Annual Report, Figure 1.3.1). Significantly, 21 notable models resulted from industry/academic collaborations in 2023, a new high. Creating cutting-edge AI models now demands a substantial amount of data, computing power, and financial resources that are not available in academia. This shift toward increased industrial dominance in leading AI models was first highlighted in last year's AI Index report. Although this year the gap has slightly narrowed, the trend largely persists.

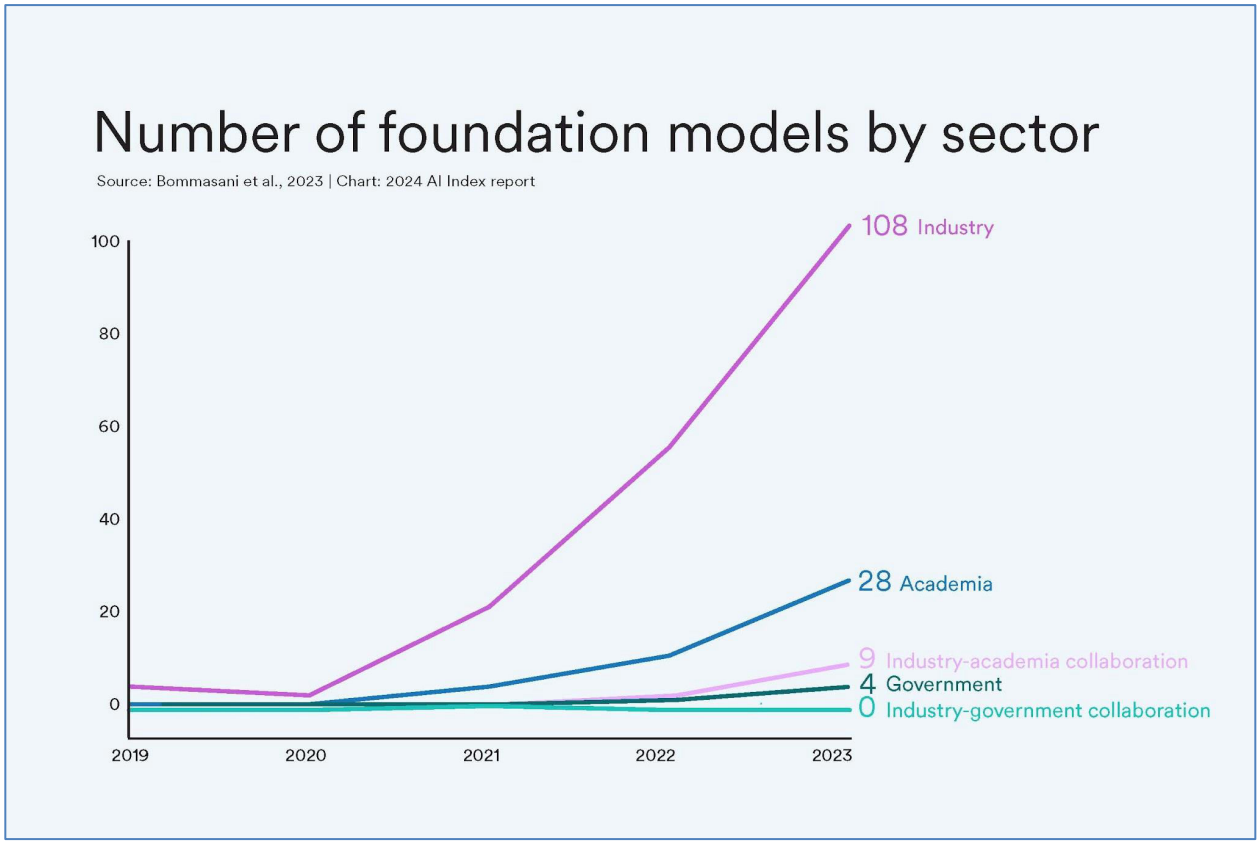


Figure 10. Number of foundation models by sector

Creating policy pathways and access to resources to pursue research in AI and Computational modeling generally is going to be essential to assuring a scientifically sound, rigorous basis for delivering AI research and instruction for professors and students alike.

AI in Migration

The use of AI for Migration surveillance demands attention to cross national policy layering and global tolerance for policy pluralism. Artificial intelligence can be used to enhance surveillance and forecasts of people on the move in origin and transit countries. When AI driven surveillance in Middle East and North African countries is traced back to EU funding through instruments like the EU Emergency Trust Fund for Africa and/or Neighbourhood, Development and

International Cooperation Instrument – Global Europe (NDICI – GE) whose human rights and AI bill of rights are in play as people cross and attempt to cross borders?[166] What about at the US Mexico border? Where is it okay to run such models and use them to aid migration decision-making or border policy? Do state, municipal, country or regional AI policies take precedence? The policies for use of AI in law enforcement or intelligence agencies or both?

The below table of use cases presented in the 2020 report on ***Opportunities and Challenges for the Use of Artificial Intelligence in Border Control*** EU shows a range of migration & security related uses of AI[167]:

Table 19: Scaling complexity categorisation based upon feasibility

Use case ID	Short name	Feasibility score	Scaling complexity
VISA-1	Application chatbot	3,9	LOW
VISA-3	Application triaging	3,1	HIGH
VISA-8	Identification of irregular travelling patterns	3,0	MEDIUM
VISA-9	Tailored application form	2,8	HIGH
ETIAS-1	Risk assessment	2,5	HIGH
ETIAS-4	Application chatbot	3,6	MEDIUM
LTSTAY-1	Application chatbot	3,7	MEDIUM
LTSTAY-3	Application triaging	3,5	HIGH
LTSTAY-9	Moving within the Schengen area	3,4	HIGH
ASYLUM-3	Vulnerability assessment	2,9	HIGH
ASYLUM-5	Registration chatbot	3,9	MEDIUM
ASYLUM-7	Abscondment risk assessment	3,0	HIGH
ASYLUM-11	Refugee allocation	3,2	HIGH
ASYLUM-14	Intelligent search engine	2,8	MEDIUM
SISSIRENE-1	Alert detection	3,5	HIGH
SISSIRENE-4	Knowledge search/management tools	3,8	MEDIUM
SISSIRENE-6	Automatic form completion	3,9	MEDIUM
SCHENGEN-3	Triaging border crossings	3,3	HIGH
SCHENGEN-4	Border flow analytics	2,9	HIGH
OPS-3	Incident prediction	3,6	MEDIUM
OPS-5	Triaging chatbot for L1/L2	4,0	LOW
OPS-7	Learning chatbot	3,9	LOW
POLICY-2	Linking regulations	4,7	LOW
POLICY-5	Clustering of regulations	4,4	LOW
POLICY-8	Automated newsgathering	3,0	HIGH
POLICY-9	Effective stakeholder communication	4,3	LOW
POLICY-12	Predicting policy acceptance	3,2	MEDIUM
CROSS-1	Multi-lingual translation	4,0	LOW
CROSS-6	Forged supporting document detection	2,7	HIGH
CROSS-7	Historical case reasoning	2,6	HIGH
CROSS-8	Ethical monitoring	3,6	HIGH*
CROSS-12	Forged travel document detection	3,6	MEDIUM
CROSS-20	Post application monitoring	3,9	MEDIUM
CROSS-23	General EU chatbot	4,6	LOW
CROSS-25	Biometric matching	3,2	HIGH

*classified as High complexity due to expected technical aspects and related challenges

Table 3. European Commission, Directorate-General for Migration and Home Affairs (2020)

To have positively impactful humane discussions about migration in the AI age we must be able to consider AI from the lens of jurisdictional layering. The alternative is to wend our way through loopholes between regions and countries and their border policies which allow companies and decision makers to collect migration related data, analyze it, using AI for decision making in the gaps between governed countries and governed human rights. Bircan & Korkmaz posit that: “A central issue concerns who is absent from the decision-making table—namely,

scholars and data scientists but also civil society and migrants themselves"[168]. They recommend that "To be able to overcome the existing challenges and be prepared for future complications in global migration governance, the discussions should rely on 'human rights' and there is a need for clear identification for the stakeholders who will closely be involved with decisions and act as a controlling body".

Just as in the AIDV WG's *Guidance for Ethics Committees Reviewing AI and DV* we argue here in the *AI Bill of Rights Recommendations* that stakeholders governing migration and their review boards will have to now include those making AI policy, those governing AI rights, and those who are concerned with AI policy compliance. Jurisdictional layering affords us an opportunity to consider immigration and AI in such a way that we take account the provenance of where models influencing immigration are run, whose data they were trained on, how that data was collected and who will be impacted by decision making based on those models.

Health/Medicine: in silico clinical trials & Digital Twins

In recent years, *in silico* approaches especially digital twins have demonstrated their ability to accurately predict changes in individuals' health using clinical endpoints. This has led to the emergence of computer modeling and simulation technologies in the field of *in silico* medicine. These models and digital twins of systems can be constructed using different approaches, depending on available technologies, data quality, and knowledge. Knowledge-driven models rely on scientific understanding of the human body's biophysics, biochemistry, and physiology, while data-driven models are developed directly from data without making causal assumptions, often utilizing artificial intelligence (AI) methods.

While AI methods offer significant power and continue to advance, it is important to exercise caution and avoid exaggerated promises that could harm this emerging field. Both knowledge-driven and data-driven models have their strengths and limitations, and the choice of approach depends on the specific question, context, and available data and knowledge. Combining different *in silico* technologies is often necessary to address complex healthcare scenarios.

Several factors should be considered when selecting the most suitable modeling technology for a given health scenario. The availability of reliable mechanistic knowledge is crucial, and in its absence, data-driven models are preferred. Computational cost varies depending on the application, with real-time simulations requiring different resources compared to regulatory approval processes or medical intervention planning. Each modeling strategy has specific data requirements for building, running, and validating predictive models. Obtaining large, high-quality datasets remains a challenge in healthcare, with data acquisition, preparation, and management costs becoming significant, especially for data-driven models. Additionally, establishing the credibility of predictions differs for knowledge-driven and data-driven models, as does the risk associated with using them beyond their validation domain or in cases of concept drift.

In silico or computer modeling and simulation (CM&S) enables the simulation of intricate biological systems and the prediction of critical parameters such as drug efficacy and toxicity by developing digital twins. By doing so, it reduces the reliance on animal testing and expensive

clinical trials. This technology has the capacity to streamline the regulatory approval process, minimize the environmental footprint associated with drug development, and ultimately enhance patient outcomes.

However, the application of *in silico* methods in regulatory processes, post-market surveillance, and investigation of adverse events is still limited. The lack of a well-defined process for the widespread adoption of *in silico* results in regulatory contexts remains a significant barrier. It is crucial to clearly establish the specific context in which *in silico* methods are utilized and demonstrate their validity for addressing specific questions of interest.

When we consider the data flow into *in silico* approaches such as integrating data-driven elements into knowledge-driven models to complement, assess, or expedite the modeling process, particularly when mechanistic knowledge is limited. Personalizing knowledge-driven models with patient-specific data into a representative digital twin often requires advanced data-driven methods when simple fitting techniques are inadequate. Knowledge-driven models can be partially or fully replaced by data-driven surrogate models to meet simulation accuracy and speed requirements. Conversely, mechanistic elements can significantly enhance data-driven models. Building powerful data-driven models typically requires large amounts of data, but in medical applications, the available clinical data may be insufficient. Knowledge-driven models can serve as data sources to improve the predictive power of data-driven models, especially for rare events. Additionally, knowledge-driven models can be used as benchmarks for testing the accuracy of developed algorithms.

The challenges of using silico models in healthcare can be categorized into several dimensions. Firstly, the computational cost, which refers to the effort required to obtain predictions within the specified time limit. Real-time models used in emergency rooms need to provide quick answers, while models for planning surgeries can take longer. Data-driven models have high computational costs during construction but low costs during execution, while knowledge-driven models have moderate construction costs but potentially higher execution costs. Secondly, data requirements play a crucial role. Predictive models need observational data for development and additional data for validation and testing. Patient-specific data is also necessary for personalized predictions and development of digital twins. Different modeling technologies have varying data requirements, and the availability and access to reliable data pose significant challenges in healthcare. The cost associated with data acquisition and the impact of privacy restrictions, such as GDPR, add to the complexity.

Credibility assessment is another dimension to consider. Guidance documents and standards provide procedures for assessing the credibility of knowledge-driven models. The identification of the question of interest, defining the context of use (CoU), and assessing the model's risk are essential steps. The credibility goals are achieved through verification, validation, and applicability analysis. The focus is on a "fit-for-purpose" approach, acknowledging that models may not be perfect but can still serve a specific CoU given certain quality criteria. For closed data-driven models, credibility assessment can be similar to that of knowledge-driven models. Some AI models have obtained FDA permission for clinical use. However, self-learning models pose challenges as new data can change the entire model, requiring re-validation. Regulatory discussions are ongoing to address this challenge, with the FDA considering a lifecycle-based regulatory framework for adaptive technologies. Explainability of AI is also an important aspect,

as requested by the EU-GDPR. Modern AI algorithms, like deep learning, may lack inherent explainability, but efforts are being made to develop additional methods to address this issue. Knowledge-driven models, on the other hand, are inherently explainable.

Overall, the challenges in using *in silico* models in healthcare include computational costs, data requirements, credibility assessment, and explainability, with ongoing discussions and developments to address these challenges and ensure the safe and effective use of such models in medical decision-making. Leveraging *in silico* and digital twin development approaches prior to animal testing or clinical trials to refine results, identify relevant animal models and populations, reduce in vivo testing, optimize drug dosing, and provide better insights for testing on real patients. *In silico* medicine holds promise in overcoming socio-economic and technological barriers to medical innovation and improving patient safety.

Changing Employment and Education Landscapes

The proliferation of AI Policy and Law at Regional, National, State, and Municipal levels will impact education, hiring and careers. There are laws emerging to regulate AI and mitigate bias in hiring (see NYC). Both education admissions and broader employment landscapes demand Responsible use of AI Systems in candidate selection and bias mitigation. Use of AI in Education itself requires thoughtful governance.

In the United States for example, it's been identified that policies are urgently needed to implement the following[169]:

1. leverage automation to advance learning outcomes while protecting human decision making and judgment;
2. interrogate the underlying data quality in AI models to ensure fair and unbiased pattern recognition and decision making in educational applications, based on accurate information appropriate to the pedagogical situation;
3. enable examination of how particular AI technologies, as part of larger edtech or educational systems, may increase or undermine equity for students; and
4. take steps to safeguard and advance equity, including providing for human checks and balances and limiting any AI systems and tools that undermine equity.

Current ethics expectations for students, faculty, and employees will need to be expanded and rethought to encompass the breadth of challenges presented by AI. There are new career paths emerging in AI operations and governance. These will demand skills in ethics, policy, model creation, deployment, management of data for AI systems, and data/model governance.

Acknowledgments

We'd like to thank Gnana Bharathy, Gauthier Chassang, John C. Havens, Mary O'Brien Uhlmansiek and Ryan O'Connor for their contributions to this content.

This work was developed by the EOSC-Future & RDA Artificial Intelligence and Data Visitation Working Group (AIDV-WG) with funding received from the European Commission

under the Horizon2020 Programme with Grant Agreement Number 101017536 through the European Open Science Cloud – Future Project (EOSC-Future). The project was designed and coordinated by Francis P. Crawley as Principal Investigator with the Good Clinical Practice Alliance - Europe (GCPA). The AIDV-WG enjoys further support from the RDA TIGER programme funded under the Horizon Europe Grant agreement ID: 101094406.

[document end]

Endnotes

1. 'IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being'. May 01, 2020. doi: 10.1109/IEEESTD.2020.9084219. Available: https://ieeexplore.ieee.org/document/9084219
2. RDA AIDV WG, Cambon-Thomsen A, Chassang G, Hackett K, Retanan LJ, Dubruel N (May 14 2024) Guidance on Informed Consent in Artificial Intelligence and Data Visitation. Research Data Alliance. doi: 10.15497/RDA00121 https://www.rd-alliance.org/group_output/guidance-for-informed-consent-in-the-context-of-artificial-intelligence-and-data-visitation/
3. RDA AIDV WG, (May 15 2024). Guidance for Ethics Committees Reviewing AI and Data Visitation. Research Data Alliance. doi: 10.15497/RDA00121 https://www.rd-alliance.org/group_output/guidance-for-informed-consent-in-the-context-of-artificial-intelligence-and-data-visitation/
4. Piovesan, C.; Avon, R. "The Workshop on Assessing and Operationalizing Responsible AI" DataConnect Conference, 2023.
5. Docherty, B. *Losing Humanity: The Case against Killer Robots*; Human Rights Watch, 2012.
6. Piovesan & Avon 2023
7. Ibid.
8. Benaich, N.; Chalmers, A.; Sebbouh, O.; Gurau, C. State of AI Report, 2023. https://www.stateof.ai/2023-report-launch & https://docs.google.com/presentation/d/156WpBF_rGvf4Ecg19oM1fyR51g4FAmHV3Zs0WLukrLQ/edit#slide=id.g24daeb7f4f0_0_3373 (accessed 2023-10-19).
9. Coglianese, C. *What Does Risk-Based Regulation Mean?*. The Regulatory Review. https://www.theregreview.org/2019/07/08/coglianese-what-does-risk-based-regulation-mean/ (accessed 2023-08-17).
10. Ibid
11. Artificial Intelligence and Data Visitation (AIDV) Working Group, 'AIDV-WG Library', *Zotero | Groups > AIDV-WG*. Research Data Alliance. Available: https://www.zotero.org/groups/4922635/aidv-wg. [Accessed: Aug. 16, 2024]
12. P. Slattery *et al.*, 'The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks From Artificial Intelligence'. Aug. 14, 2024. doi: 10.13140/RG.2.2.28850.00968
13. https://www.eaidb.org/map
14. Kherroubi Garcia, I. 'Automating Safety, Industrialising Ethics; A critical lens on AI ethics solutions for research integrity', Jul. 23, 2024. doi: 10.5281/zenodo.12788336. p 29 Available: https://zenodo.org/records/12788336. [Accessed: Jul. 23, 2024]
15. UNESCO (2021, November). UNESCO Recommendations on AI Ethics. UN Educational, Scientific and Cultural Organization (UNESCO). https://unesdoc.unesco.org/ark:/48223/pf0000377897
16. The US and Israel were not members of UNESCO in Spring 2023, or the time of the adoption of the UNESCO AI Recommendations in 2021. The United States initially joined UNESCO at its founding in 1945. but withdrew for the first time in 1984 in protest against alleged financial mismanagement and perceived anti-U.S. bias, returning almost 20 years later in 2003 under President George W. Bush, who then said the agency had undertaken needed reforms, then stopped paying in 2011 when Palestine became a full member because such funding is barred by U.S. law. Israel and The United States formally withdrew from the U.N. cultural agency again in December 2018 over accusations of anti-Israel bias and mismanagement. Washington owed $542 million when it quit. An agreement reached at the U.S. Congress in December 2022 makes it possible for Washington to re-start financial contributions to UNESCO and re-admission of the United States into UNESCO is planned for July 2023 pending approval of the financial plan.
17. GPAI. Global Partnership on Artificial Intelligence - GPAI. https://gpai.ai/
18. GPAI. Data Governance Working Group. https://gpai.ai/projects/data-governance/
19. GPAI. (2022) Data Justice in Practice: A Guide for Policymakers. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4080050
20. GPAI. (2022) Data Justice in Practice: A Guide for Impacted Communities. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4080046
21. GPAI. Responsible AI Working Group. https://gpai.ai/projects/responsible-ai/
22. GPAI. (2021) Climate Change And Ai: Recommendations for Government Action. https://gpai.ai/projects/responsible-ai/environment/climate-change-and-ai.pdf

23. GPAI Responsible AI Working Group and Project RAISE. A responsible AI strategy for the environment. *GPAI*, https://gpai.ai/projects/responsible-ai/environment/
24. OECD (2022) Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint. https://doi.org/10.1787/7babf571-en
25. GPAI 2022 (2022) AI for Net Zero Electricity: Assessing the Electricity Sector's Readiness for AI. https://gpai.ai/projects/responsible-ai/environment/AINetZeroElectricity_2022.pdf
26. GPAI 2022 (2022) Biodiversity & Artificial Intelligence, Opportunities and Recommendations. https://gpai.ai/projects/responsible-ai/environment/biodiversity-and-AI-opportunities-recommendations-for-action.pdf
27. F. Wu *et al.*, 'Towards a new generation of artificial intelligence in China', *Nature Machine Intelligence*, vol. 2, no. 6, Art. no. 6, Jun. 2020, doi: 10.1038/s42256-020-0183-4.
28. The National People's Congress of the People's Republic of China, 'Data Security Law of the People's Republic of China'. The National People's Congress of the People's Republic of China, 2021. Accessed: May 19, 2023. [Online]. Available: http://www.npc.gov.cn/englishnpc/c23934/202112/1abd8829788946ecab270e469b13c39c.shtml
29. The National People's Congress of the People's Republic of China, 'Personal Information Protection Law of the People's Republic of China'. The National People's Congress of the People's Republic of China, 2021. Accessed: May 19, 2023. [Online]. Available: http://en.npc.gov.cn.cdurl.cn/2021-12/29/c_694559.htm
30. State Council Information Office, 'China's Law-Based Cyberspace Governance in the New Era', State Council Information Office, Beijing, Mar. 2023. Accessed: May 19, 2023. [Online]. Available: http://english.scio.gov.cn/node_9001076.html
31. L. Zhang, 'China: AI governance principles released', *Library of Congress*, 2019. Accessed: May 19, 2023. [Online]. Available: https://www.loc.gov/item/global-legal-monitor/2019-09-09/china-ai-governance-principles-released/
32. 全国信息安全标准化技术委员会秘书处, '关于发布《网络安全标准实践指南—人工智能伦理安全风险防范指引》的通知', 全国信息安全标准化技术委员会, Jan. 05, 2021.
33. 国家人工智能标准化总体组 and 全国信标委人工智能分委会, '人工智能伦理治理标准化指南(2023版)'. 中国电子技术标准化研究院, Mar. 2023. [Online]. Available: https://www.aipubservice.com/airesource/fs/%E4%BA%BA%E5%B7%A5%E6%99%BA%E8%83%BD%E4%BC%A6%E7%90%86%E6%B2%BB%E7%90%86%E6%A0%87%E5%87%86%E5%8C%96%E6%8C%87%E5%8D%97.pdf
34. L. Zhang, 'China: Provisions on deep synthesis technology enter into effect', *Library of Congress*, Apr. 26, 2023. Accessed: May 19, 2023. [Online]. Available: https://www.loc.gov/item/global-legal-monitor/2023-04-25/china-provisions-on-deep-synthesis-technology-enter-into-effect/
35. Y. Cao, 'Cyber watchdog opens generative AI draft regulation to public opinions', *China Daily*, Apr. 11, 2023. Accessed: Apr. 12, 2023. [Online]. Available: https://www.chinadaily.com.cn/a/202304/11/WS64350bdea31057c47ebb981c.html
36. P. Schildkraut and H. Zhang, 'What to know about China's new AI regulations', Arnold&Porter, Apr. 2023.
37. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021. License: CC BY-NC-SA 3.0 IGO.
38. UNESCO World Committee on the Ethics of Scientific Knowledge and Technology. *Recommendation on the Ethics of Artificial Intelligence*; 0000380455; UNESCO, 2021; p 21. https://unesdoc.unesco.org/ark:/48223/pf0000380455.locale=en (accessed 2023-01-23).
39. Ministry of Foreign Affairs of the People's Republic of China, 'Position Paper of the People's Republic of China on Strengthening Ethical Governance of Artificial Intelligence (AI)', *Ministry of Foreign Affairs of the People's Republic of China*, Nov. 17, 2022. Accessed: May 19, 2023. [Online]. Available: https://www.fmprc.gov.cn/eng/wjdt_665385/wjzcs/202211/t20221117_10976730.html
40. Staff writer, with CNA (2023) Government drafting basic law to regulate AI. *Taipei Times*, https://www.taipeitimes.com/News/taiwan/archives/2023/07/06/2003802773
41. MIIA The Machine Intelligence Institute of Africa (MIIA). *MIIA*, Available at: https://miiafrica.org/
42. UNESCO (2018) Forum on artificial intelligence in Africa. *Forum on artificial intelligence in Africa*, https://en.unesco.org/artificial-intelligence/africa-forum & Participants of the Forum on Artificial Intelligence in Africa (2018) Outcome Statement of the Forum on Artificial Intelligence in Africa. https://en.unesco.org/sites/default/files/ai_outcome-statement_africa-forum_en.pdf

43. Bernd Carsten Stahl, Tonii Leach, Oluyinka Oyeniji, George Ogoh (2023) AI Policy as a Response to AI Ethics? Addressing Ethical Issues in the Development of AI Policies in North Africa. Responsible AI in Africa: Challenges and Opportunities, DOI: https://doi.org/10.1007/978-3-031-08215-3
44. Bernd Carsten Stahl, Tonii Leach, Oluyinka Oyeniji, George Ogoh (2023) AI Policy as a Response to AI Ethics? Addressing Ethical Issues in the Development of AI Policies in North Africa. *Responsible AI in Africa: Challenges and Opportunities*, https://doi.org/10.1007/978-3-031-08215-3
45. (2021) "'AI For Africa"' Blueprint. https://smart.africa/board/login/uploads/70029-eng_ai-for-africa-blueprint.pdf
46. National Information Technology Development Agency (NITDA), Nigeria https://nitda.gov.ng/
47. NITDA Act (2007) https://nitda.gov.ng/wp-content/uploads/2020/11/NITDA-ACT-2007-2019-Edition1.pdf
48. National Center for Artificial Intelligence and Robotics – NITDA. https://nitda.gov.ng/ncair/
49. Nigerian Communications Commission (2020) National Digital Economy Policy and Strategy (NDEPS)(2020-2030). https://ncc.gov.ng/media-centre/public-notices/822-national-digital-economy-policy-strategy & https://ndpb.gov.ng/Files/Policy-National_Digital_Economy_Policy_and_Strategy.pdf
50. Artificial Intelligence Center of Excellence Africa (AICE). https://www.aiceafrica.com/
51. NITDA (2020) Nigeria Data Protection Regulation 2019: Implementation Framework. https://nitda.gov.ng/wp-content/uploads/2021/01/NDPR-Implementation-Framework.pdf
52. NITDA (2020) Guidelines For The Management Of Personal Data By Public Institutions In Nigeria, 2020. https://nitda.gov.ng/wp-content/uploads/2020/11/GuidelinesForImplementationOfNDPRInPublicInstitutionsFinal11.pdf
53. NITDA (2019) Guidelines for Nigerian Content Development in Information and Communication Technology (ICT). https://nitda.gov.ng/wp-content/uploads/2020/11/GNCFinale2211.pdf
54. NITDA (2019) Framework And Guidelines For Public Internet Access (PIA). https://nitda.gov.ng/wp-content/uploads/2020/11/FrameworkAndGuidelinesForPublicInternetAccessPIA1.pdf
55. NITDA (2018) Guidelines For Clearance Of Information Technology (It) Projects By Public Institutions. https://nitda.gov.ng/wp-content/uploads/2020/11/Guidelines-For-Clearance-Of-Information-Technology-IT-Projects-By-Public-Institutions21.pdf
56. NITDA (2018) GUIDELINES FOR REGISTRATION OF ICT SERVICE PROVIDERS/CONTRACTORS FOR DELIVERY OF IT SERVICES TO MDAS. https://nitda.gov.ng/wp-content/uploads/2020/11/GuidelinesForRegistrationServiceProviders-311.pdf
57. NITDA (2019) Nigeria e-Government Interoperability Framework (Ne-GIF). https://nitda.gov.ng/wp-content/uploads/2020/11/Ne-GIFFinal1.pdf
58. NITDA (2019) Framework and Guidelines for Information and Communication Technology (ICT) Adoption in Tertiary Institutions. https://nitda.gov.ng/wp-content/uploads/2020/11/FGICTDTIfinal.pdf
59. NITDA (2019) FRAMEWORK AND GUIDELINES FOR THE USE OF SOCIAL MEDIA PLATFORMS IN PUBLIC INSTITUTIONS. https://nitda.gov.ng/wp-content/uploads/2020/11/FrameworkAndGuidelinesForTheUseOfSocialMediaPlatformsInPublicInstitutions1.pdf
60. Nigerian Communications Commission (2023) Ethical & Societal Impact of Artificial Intelligence (AI). Available: https://ncc.gov.ng/technical-regulation/research/919-ethical-societal-impact-of-artificial-intelligence-ai
61. Stahl, B. C.; Leach, T.; Oyeniji, O.; Ogoh, G. AI Policy as a Response to AI Ethics? Addressing Ethical Issues in the Development of AI Policies in North Africa. In *Responsible AI in Africa: Challenges and Opportunities*; Eke, D. O., Wakunuma, K., Akintoye, S., Eds.; Social and Cultural Studies of Robots and AI; Springer International Publishing: Cham, 2023; pp 141–167. https://doi.org/10.1007/978-3-031-08215-3_7.
62. 'Riyadh AI Call for Action (RAICA) Declaration', Digital Cooperation Organization, Riyadh, Saudi Arabia, Sep. 2022. Available: https://dco.org/wp-content/uploads/2024/06/Riyadh-AI-Call-for-Action-RAICA-Declaration.pdf
63. Purian, Ronit. (August 12, 2023). "Technical Landscape, Design Research, Social Action & Responsibility." Presentation at The Future of System Goals and Human Rights: Data Governance & AI Ethics [PDS]. AMCIS 2023. Panama City, Panama. Available: DOI 10.17605/OSF.IO/VRP8H
64. Ministry of Innovation, Science and Technology Israel (2022) Regulation and Ethics for the Field of Artificial Intelligence in Israel [Draft Policy]. *GOV.IL*, https://www.gov.il/he/departments/news/most-news20223110
65. Fridman, L. (2023, July 12). Benjamin Netanyahu: Israel, Palestine, Power, Corruption, Hate, and Peace. Lex Fridman Podcast, #389 Available: https://lexfridman.com/benjamin-netanyahu
66. TAU GDPR Guide: https://research-authority.tau.ac.il/sites/resauth.tau.ac.il/files/media_server/Research-Authority/instructions/TAU-guide-GDPR.PDF

67. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance)*. 2016. 119 OJ L http://data.europa.eu/eli/reg/2016/679/oj/eng (July 17, 2023).
68. CHARTER OF FUNDAMENTAL RIGHTS OF THE EUROPEAN UNION, 2000. https://www.europarl.europa.eu/charter/pdf/text_en.pdf (accessed 2023-10-16).
69. Chassang, G. The Impact of the EU General Data Protection Regulation on Scientific Research. 2017. https://doi.org/10.3332/ecancer.2017.709.
70. European Data Protection Board. Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, 2019 [adopted 20 October 2020].https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-42019-article-25-data-protection-design-and_en (accessed 2023-10-16).
71. *Adequacy decisions*. https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en (accessed 2023-10-16).
72. *Adequacy decisions*. https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en (accessed 2023-10-16).
73. Pollina E, Coulter M (2023) Italy bans U.S.-based AI chatbot Replika from using personal data. *Reuters*, https://www.reuters.com/technology/italy-bans-us-based-ai-chatbot-replika-using-personal-data-2023-02-03/
74. Shiona McCallum (2023) ChatGPT banned in Italy over privacy concerns. *BBC News*, https://www.bbc.com/news/technology-65139406
75. Vincent J (2023) Italian regulators order ChatGPT ban over alleged violation of data privacy laws. *The Verge*, https://www.theverge.com/2023/3/31/23664451/italy-bans-chatgpt-over-data-privacy-laws &
76. GPDP Garante Per La Protezione Dei Dati Personali (2023) Intelligenza artificiale: il Garante blocca ChatGPT. Raccolta illecita di dati personali. Assenza di sistemi per la verifica dell'età dei minori. https://www.garanteprivacy.it:443/home/docweb/-/docweb-display/docweb/9870847 (it) / (eng)Italian Data Protection Authority (Garante per la protezione dei dati personali) (2023) Provvedimento del 30 marzo 2023 [9870832]. https://www.garanteprivacy.it:443/home/docweb/-/docweb-display/docweb/9870832
77. (2023) ChatGPT accessible again in Italy. *BBC News*, https://www.bbc.com/news/technology-65431914
78. EU legislation in Progress: European Parliament (2023) Artificial Intelligence Act. https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf
79. 1. (2023) AI Act: a step closer to the first rules on Artificial Intelligence | News | European Parliament. https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence
80. European Commission (2010) Council Directive 2001/95/EC on general product safety. 95 http://data.europa.eu/eli/dir/2001/95/2010-01-01/eng
81. S. Chander and E. Jakubowska, 'Civil society calls on the EU to put fundamental rights first in the AI Act', *European Digital Rights (EDRi)*, Nov. 30, 2021. Available: https://edri.org/our-work/civil-society-calls-on-the-eu-to-put-fundamental-rights-first-in-the-ai-act/. [Accessed: Jul. 17, 2023]
82. Council of Europe, 'The Council of Europe and the European Union', *The Council of Europe: guardian of Human Rights, Democracy and the Rule of Law for 700 million citizens - Portal*, 22 2024. Available: https://www.coe.int/en/web/portal/european-union. [Accessed: Jul. 18, 2023]
83. Committee on Artificial Intelligence (CAI), *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. 2024. Available: https://rm.coe.int/1680afae3c. [Accessed: Oct. 14, 2024] and M. HERNANDEZ RAMOS, 'Committee on Artificial Intelligence (CAI) Word of the Chair', *Artificial Intelligence*, Sep. 05, 2024. Available: https://www.coe.int/en/web/artificial-intelligence/cai. [Accessed: Oct. 14, 2024] & 'Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law', Council of Europe, Vilinus, 225, May 2024. Available: https://rm.coe.int/1680afae67. [Accessed: Aug. 19, 2024]
84. Committee on Artificial Intelligence (CAHAI), 'Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law'. Council of Europe, Feb. 16, 2022. Available: https://rm.coe.int/possible-elements-of-a-legal-framework-on-artificial-intelligence/1680a5ae6b. [Accessed: Aug. 19, 2024]
85. European Commission for the Efficiency of Justice Working Group on quality of justice (CEPEJ-GT-QUAL) (2018) "European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. " https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c & Quality of justice - European Commission for the Efficiency of Justice (CEPEJ) - www.coe.int. *European Commission for the Efficiency of Justice (CEPEJ)*, https://www.coe.int/en/web/cepej/cepej-work/quality-of-justice

86. European Commission for the Efficiency of Justice (CEPEJ) CEPEJ Artificial Intelligence Advisory Body (AIAB). *European Commission for the Efficiency of Justice (CEPEJ)*, https://www.coe.int/en/web/cepej/ai-advisory-board
87. CEPEJ Resource Centre on Cyberjustice and AI - European Commission for the Efficiency of Justice. *European Commission for the Efficiency of Justice (CEPEJ)*, https://www.coe.int/en/web/cepej/resource-centre-on-cyberjustice-and-ai
88. CEPEJ Resource Centre on Cyberjustice and AI - European Commission for the Efficiency of Justice (CEPEJ) - www.coe.int. *European Commission for the Efficiency of Justice (CEPEJ)*, https://www.coe.int/en/web/cepej/resource-centre-on-cyberjustice-and-ai
89. Central Digital and Data Office, 'UK government publishes pioneering standard for algorithmic transparency', *GOV.UK*, Nov. 09, 2021. Available: https://www.gov.uk/government/news/uk-government-publishes-pioneering-standard-for-algorithmic-transparency. [Accessed: May 04, 2023]
90. Department for Science, Innovation and Technology, *AI regulation: a pro-innovation approach*.(CP 815) London, United Kingdom: The Stationery Office, 2023. Available: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach. [Accessed: Aug. 15, 2024]
91. Alan Turing Institute, 'AI Standards Hub - The New Home of the AI Standards Community', *AI Standards Hub*, Jun. 24, 2022. Available: https://aistandardshub.org/. [Accessed: Aug. 15, 2024]
92. 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 Electronic Journal*, https://doi.org/10.2139/ssrn.3817999
93. House TW (2021) U.S.-EU Trade and Technology Council Inaugural Joint Statement. *The White House*, https://www.whitehouse.gov/briefing-room/statements-releases/2021/09/29/u-s-eu-trade-and-technology-council-inaugural-joint-statement/
94. A. Engler, 'The EU and U.S. diverge on AI regulation: A transatlantic comparison and steps to alignment', *Brookings Institution*, Apr. 25, 2023. Available: https://www.brookings.edu/articles/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/. [Accessed: Aug. 19, 2024]
95. A. Engler, 'The EU and U.S. diverge on AI regulation: A transatlantic comparison and steps to alignment', *Brookings Institution*, Apr. 25, 2023. Available: https://www.brookings.edu/articles/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/. [Accessed: Aug. 19, 2024]
96. NIST AIRC, 'NIST Trustworthy & Responsible Artificial Intelligence Resource Center (AIRC).', *NIST Trustworthy & Responsible AI Resource Center*. Available: https://airc.nist.gov/Home. [Accessed: Aug. 15, 2024]
97. 'AI Risk Management Framework: AI RMF (1.0)', National Institute of Standards and Technology, Gaithersburg, MD, NIST AI 100-1, Jan. 2023. doi: 10.6028/NIST.AI.100-1. Available: https://doi.org/10.6028/NIST.AI.100-1. [Accessed: Jul. 19, 2023]
98. U.S. NIST AIRC Team, 'Roadmap for the NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0)', *NIST AI RMF Roadmap*. Available: https://airc.nist.gov/AI_RMF_Knowledge_Base/Roadmap. [Accessed: Aug. 16, 2024]
99. AIRMF v. 1 (2023), p 4 doi: 10.6028/NIST.AI.100-1
100. AIRMF v. 1 (2023), p 4 doi: 10.6028/NIST.AI.100-1
101. AIRMF v. 1 (2023), p 35-37. doi: 10.6028/NIST.AI.100-1
102. Ibid.
103. White House Office of Science and Technology Policy (2022) Blueprint for an AI Bill of Rights | OSTP. *The White House*, https://www.whitehouse.gov/ostp/ai-bill-of-rights/
104. Institute of Industrial and Systems Engineers (IISE) (2022) "AI Bill of Rights" aims to curb possible abuse. *IISE ; Industrial and Systems Engineering at Work*, 54(12):18–18.
105. Hine E, Floridi L (2023) The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction. *Minds and machines (Dordrecht)*, https://doi.org/10.1007/s11023-023-09625-1
106. Burke, Garance (2022) White House unveils artificial intelligence "Bill of Rights." *AP News*, https://apnews.com/article/technology-business-artificial-intelligence-7a39848340d210592aeea2478225f489
107. Hine E, Floridi L (2023) The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction. *Minds and machines (Dordrecht)*, https://doi.org/10.1007/s11023-023-09625-1
108. Loten A (2022) White House Issues "Blueprint for an AI Bill of Rights"; Some tech executives voice concerns that the nonbinding guidelines could lead to stifling regulation concerning artificial intelligence. *The Wall Street journal. Eastern edition*,

109. N. Sampath, 'Consumer Reports applauds White House's AI Bill of Rights', *Consumer Reports: Advocacy*, Oct. 04, 2022. Available: https://advocacy.consumerreports.org/press_release/consumer-reports-applauds-white-houses-ai-bill-of-rights/. [Accessed: Jul. 17, 2023]
110. Lilian Edwards, 'Expert explainer: The EU AI Act proposal', *Ada Lovelace Institute*, Apr. 08, 2022. Available: https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/. [Accessed: Jul. 17, 2023]
111. AI Index Steering Committee, 'The AI Index 2024 Annual Report', Institute for Human-Centered AI Stanford University, Stanford, CA, Apr. 2024. Available: https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf. [Accessed: Jul. 24, 2024]
112. Jerry McNerney, *Ai in Government Act of 2020*. 2020. Available: https://www.congress.gov/116/bills/hr133/BILLS-116hr133enr.pdf#page=1105. [Accessed: Aug. 15, 2024]
113. Technology Transformation Services U.S. General Services Administration (GSA), 'Artificial Intelligence (AI) Center of Excellence (CoE)', *Artificial Intelligence | GSA - IT Modernization Centers of Excellence*. Available: https://coe.gsa.gov/coe/artificial-intelligence.html. [Accessed: Aug. 15, 2024]
114. D. L. Parker, 'National Artificial Intelligence Initiative Act of 2020 (NAIIA)', presented at the Artificial Intelligence and Emerging Technology Inaugural Stakeholder Meeting, Jun. 29, 2022.
115. P. J. Phillips *et al.*, 'Four principles of explainable artificial intelligence', National Institute of Standards and Technology (U.S.), Gaithersburg, MD, NIST IR 8312, Sep. 2021. doi: 10.6028/NIST.IR.8312. Available: https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8312.pdf. [Accessed: Aug. 15, 2024]
116. U.S. Office of Management and Budget (OMB), 'Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence'. OMB, Mar. 28, 2024. Available: https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf. [Accessed: Aug. 15, 2024]
117. US EEOC. Select Issues: Assessing Adverse Impact in Software, Algorithms, and Artificial Intelligence Used in Employment Selection Procedures Under Title VII of the Civil Rights Act of 1964, 2023.https://www.eeoc.gov/select-issues-assessing-adverse-impact-software-algorithms-and-artificial-intelligence-used (accessed 2023-07-21).
118. National Conference of State Legislatures. *Artificial Intelligence 2023 Legislation*. https://www.ncsl.org/technology-and-communication/artificial-intelligence-2023-legislation (accessed 2023-07-21).
119. 'Artificial Intelligence 2024 Legislation'. Available: https://www.ncsl.org/technology-and-communication/artificial-intelligence-2024-legislation. [Accessed: Oct. 15, 2024]
120. *ACR-215 23 Asilomar AI Principles.* Available: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180ACR215. [Accessed: Oct. 21, 2024]
121. G. Newsom, 'EXECUTIVE ORDER N-12-23'. Executive Department State of California, Sep. 06, 2023. Available: https://www.gov.ca.gov/wp-content/uploads/2023/09/AI-EO-No.12-_-GGN-Signed.pdf. [Accessed: Sep. 18, 2024]
122. S. of California, 'Governor Newsom announces new initiatives to advance safe and responsible AI, protect Californians', *Governor of California*, Sep. 29, 2024. Available: https://www.gov.ca.gov/2024/09/29/governor-newsom-announces-new-initiatives-to-advance-safe-and-responsible-ai-protect-californians/. [Accessed: Oct. 15, 2024]
123. 'Colorado's Groundbreaking AI Law: A Model for National Regulation', *Av{ai}lable Law*. Available: https://www.availablelaw.com/blog/colorados-groundbreaking-ai-law-a-model-for-national-regulation. [Accessed: Sep. 23, 2024]
124. New York City Local Law 144 https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9&Options=ID%7CText%7C&Search= [Accessed: Oct. 19, 2024]
125. UNESCO. *CAF and UNESCO Will Create a Council to Review Ethical Criteria for Artificial Intelligence in Latin America and the Caribbean | UNESCO*. 23 June 2022, https://www.unesco.org/en/articles/caf-and-unesco-will-create-council-review-ethical-criteria-artificial-intelligence-latin-america-and
126. V. Authors, 'Declaración de Montevideo sobre Inteligencia Artificial y su impacto en América Latina / Declaração de Montevidéu sobre Inteligência Artificial e seu impacto na América Latina / Montevideo Declaration on Artificial Intelligence and its Impact in Latin America', Mar. 2023. Available: https://doi.org/10.5281/zenodo.8208793. [Accessed: Oct. 15, 2024]
127. 'Declaration of Cartagena de Indias for Governance, Building Artificial Intelligence (AI) Ecosystems and Promoting Ethical and Responsible AI Education in Latin America and the Caribbean'. Latin American and Caribbean Ministerial Summit, Aug. 09, 2024. Available: https://mintic.gov.co/portal/715/articles-383990_recurso_1.pdf. [Accessed: Oct. 15, 2024]

128. Centro Nacional de Inteligencia Artificial, 'Índice Latinoamericano de Inteligencia Artificial: Governance', *Índice Latinoamericano de Inteligencia Artificial*, 2024. Available: https://indicelatam.cl/gobernanza-3/. [Accessed: Oct. 15, 2024]
129. E. C. for L. A. and the Caribbean, 'ILIA 2024: Evaluating AI Readiness and Progress in Latin America', Economic Commission for Latin America and the Caribbean, Text. Available: https://www.cepal.org/en/notes/ilia-2024-evaluating-ai-readiness-and-progress-latin-america. [Accessed: Oct. 15, 2024]
130. 'Readiness Assessment Methodology | Global AI Ethics and Governance Observatory'. UNESCO, 2023. Available: https://www.unesco.org/ethics-ai/en/ram. [Accessed: Aug. 21, 2024]
131. Artificial Intelligence Task Force Presidency of the Republic of Colombia (2021) Ethical Framework For Artificial Intelligence In Colombia. https://inteligenciaartificial.gov.co/en/ethical-framework/
132. F. Giandana Gigena *et al.*, 'Regulatory Mapping on Artificial Intelligence in Latin America', Access Now, Jul. 2024. Available: https://www.accessnow.org/wp-content/uploads/2024/07/TRF-LAC-Reporte-Regional-IA-JUN-2024-V3.pdf
133. Government of the Autonomous City of Buenos Aires Artificial Intelligence Plan | Buenos Aires City -. https://buenosaires.gob.ar/jefaturadegabinete/innovacion/plan-de-inteligencia-artificial
134. The story behind Latin America's first AI strategy. *Apolitical*, https://apolitical.co/solution-articles/en/mexico-the-story-and-lessons-behind-latin-americas-first-ai-strategy
135. Emma Martinho-Truswell, Hannah Miller, Isak Nti Asare, André Petheram, Richard Stirling, Constanza Gómez Mont, Cristina Martinez (2018) Towards An Ai Strategy In Mexico: Harnessing the AI Revolution. *Oxford Insights*, https://www.oxfordinsights.com/mexico.
136. *Artificial Intelligence in Mexico: A National Agenda*; Del Pozo, C. M., Gómez Mont, C., Martínez Pinto, C., Eds.; IA2030Mx: Mexico, 2020.
137. ACS (2017) Data sharing frameworks: Technical White Paper. https://www.acs.org.au/insightsandpublications/reports-publications/data-sharing-frameworks.html
138. ACS (2019) Privacy-Preserving Data Sharing Frameworks. https://www.acs.org.au/insightsandpublications/reports-publications/privacy-preserving-data-sharing-frameworks.html
139. ACS (2019) Privacy-Preserving Data Sharing Frameworks: People Projects Data and Output. https://www.acs.org.au/insightsandpublications/reports-publications/privacy-preserving-data-sharing-frameworks.html
140. Department of Industry S and R (2023) Australia's Artificial Intelligence Ethics Framework | Department of Industry, Science and Resources. *https://www.industry.gov.au/node/75445*, https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework
141. Australian Government et al., 'National framework for the assurance of artificial intelligence in government', Commonwealth of Australia, Jun. 2024.
142. Commonwealth of Australia (Digital Transformation Agency), 'Policy for the responsible use of AI in government v 1.1'. Commonwealth of Australia, 2024. Available: https://www.digital.gov.au/sites/default/files/documents/2024-07/24-0033%20digital%20experience%20policy%20FINAL%20Acc%20r2_KC_UA.pdf
143. Australian Government Department of Industry, Science and Resources. Proposals Paper for Introducing Mandatory Guardrails for AI in High-Risk Settings. Commonwealth of Australia, Sep. 2024. Available: https://storage.googleapis.com/converlens-au-industry/industry/p/prj2f6f02ebfe6a8190c7bdc/page/proposals_paper_for_introducing_mandatory_guardrails_for_ai_in_high_risk_settings.pdf. [Accessed: Oct. 09, 2024]
144. A. Saeri, M. Noetel, and J. Graham, 'Survey Assessing Risks from Artificial Intelligence (Technical Report)'. Rochester, NY, Mar. 07, 2024. doi: 10.2139/ssrn.4750953. Available: https://papers.ssrn.com/abstract=4750953. [Accessed: Oct. 09, 2024]
145. The Artificial Intelligence Forum of New Zealand (AI Forum), 'AI Forum NZ: About', *AI Forum New Zealand*, 2017. Available: https://aiforum.org.nz/about/. [Accessed: Oct. 15, 2024]
146. New Zealand Government, 'Algorithm charter for Aotearoa New Zealand - (2020) data.govt.nz', *data.govt.nz*. Available: https://data.govt.nz/toolkit/data-ethics/government-algorithm-transparency-and-accountability/algorithm-charter/. [Accessed: Aug. 19, 2024]
147. Ibid.
148. The Government Chief Digital Officer (GCDO), NZ Digital Government, 'Initial advice on Generative Artificial Intelligence in the public service', *Interim Generative AI guidance for the public service | NZ Digital government*, July, 2023. Available: https://www.digital.govt.nz/assets/Standards-guidance/Technology-and-

architecture/Generative-AI/Joint-System-Leads-tactical-guidance-on-public-service-use-of-GenAI-September-2023.pdf
149. *Interim Generative AI guidance for the public service* [Website] | *NZ Digital government*, Sep. 2023. Available: https://www.dns.govt.nz/standards-and-guidance/technology-and-architec147147ture/artificial-intelligence/interim-generative-ai-guidance-for-the-public-service . [Accessed: Aug. 19, 2024]
150. Office of the Prime Minister's Chief Science Advisor, 'Regulating risks associated with AI', *Artificial intelligence | Prime Minister's Chief Science Advisor, New Zealand*, Jul. 30, 2023. Available: https://bpb-ap-se2.wpmucdn.com/blogs.auckland.ac.nz/dist/f/688/files/2023/07/Regulatory-landscape.pdf. [Accessed: Aug. 19, 2024]
151. J. Gerrard, I. Town, R. Benson, E. Brown, and C. Varughese, *Capturing the benefits of AI in healthcare for Aotearoa New Zealand*. Auckland, NZ: Office of the Prime Minister's Chief Science Advisor, 2023. Available: https://bpb-ap-se2.wpmucdn.conew zeland m/blogs.auckland.ac.nz/dist/f/688/files/2023/12/PMCSA-AI-healthcare-LONG-REPORT-FINAL-pdf-version-v3.pdf. [Accessed: Oct. 10, 2024]
152. Farrell, H., & Newman, A. (2015). The New Politics of Interdependence: Cross-National Layering in Trans-Atlantic Regulatory Disputes. *Comparative Political Studies*, *48*(4), 497-526.
153. renAIssance Foundation (2020) The call | Rome Call. *Rome Call | What is the Matter with AI Ethics?*, https://www.romecall.org/the-call/
154. Tech Ethics Lab University of Notre Dame, 'Looking Back at the Rome Call for AI Ethics: A Global University Summit', *Tech Ethics Lab*, Nov. 10, 2022. Available: https://techethicslab.nd.edu/news/looking-back-at-the-rome-call-for-ai-ethics-a-global-university-summit/. [Accessed: Jan. 23, 2023]
155. Tech Ethics Lab University of Notre Dame, 'Rome Call for AI Ethics: A Global University Summit', *Tech Ethics Lab*, Oct. 27, 2022. Available: https://techethicslab.nd.edu/news-and-events/rome-call-for-ai-ethics-a-global-university-summit/. [Accessed: Jan. 23, 2023]
156. U. of N. Dame, 'Statement on Generative AI and Academic Integrity (May 2023)', *Undergraduate Academic Code of Honor*. Available: https://honorcode.nd.edu/official-statement-regarding-generative-ai-may-2023/. [Accessed: Oct. 15, 2024]
157. U. of N. Dame, 'AI Recommendations for Instructors', *Undergraduate Academic Code of Honor*. Available: https://honorcode.nd.edu/ai-recommendations-for-instructors/. [Accessed: Oct. 15, 2024]
158. U. of N. Dame, 'Statement on Generative AI and Academic Integrity (May 2023)', *Undergraduate Academic Code of Honor*. Available: https://honorcode.nd.edu/official-statement-regarding-generative-ai-may-2023/. [Accessed: Oct. 15, 2024]
159. U. of N. Dame, 'Policies and Guidelines', *AI@ND*. Available: https://ai.nd.edu/policies-and-guidelines/. [Accessed: Oct. 15, 2024]
160. Future of Life Institute, Signatories (2017) Asilomar AI Principles. https://futureoflife.org/open-letter/ai-principles/
161. Sterling B (1018) The Asilomar AI Principles. *Wired*, https://www.wired.com/beyond-the-beyond/2018/06/asilomar-ai-principles
162. University of California AI Council, 'UC Responsible AI Principles'. UCOP, Oct. 18, 2021. Available: https://ai.universityofcalifornia.edu/_files/documents/ai-council-uc-responsible-ai-principles.pdf. [Accessed: Oct. 15, 2024]
163. 'Applicable Law and UC Policy'. Available: https://ai.universityofcalifornia.edu/governance-transparency/applicable-law-and-policy.html. [Accessed: Oct. 15, 2024]
164. M. Zeff, 'Here is what's illegal under California's 18 (and counting) new AI laws', *TechCrunch*, Sep. 29, 2024. Available: https://techcrunch.com/2024/09/29/here-is-whats-illegal-under-californias-18-and-counting-new-ai-laws/. [Accessed: Oct. 15, 2024]
165. HAI, 'The AI Index 2024 Annual Report', AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, Apr. 2024. Available: https://aiindex.stanford.edu/report/. [Accessed: Aug. 19, 2024]
166. Napolitano, A. *Artificial Intelligence: The New Frontier of the EU's Border Externalisation Strategy*; EuroMed Rights: Copenhagen, Denmark, 2023. https://euromedrights.org/wp-content/uploads/2023/07/Euromed_AI-Migration-Report_EN-1.pdf (accessed 2023-07-23).
167. Directorate-General for Migration and Home Affairs (European Commission), *Opportunities and challenges for the use of artificial intelligence in border control, migration and security. Volume 1, Main report*. LU: Publications Office of the European Union, 2020. Available: https://data.europa.eu/doi/10.2837/923610. [Accessed: Aug. 21, 2023]

168. Bircan, T.; Korkmaz, E. E. Big Data for Whose Sake? Governing Migration through Artificial Intelligence. *Humanit Soc Sci Commun* 2021, 8 (1), 1–5. https://doi.org/10.1057/s41599-021-00910-x.
169. U.S. Department of Education, Office of Educational Technology. *Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations*; Washington, DC, 2023. https://www2.ed.gov/documents/ai-report/ai-report.pdf (accessed 2023-07-22).