The Global AI Regulatory Environment

Alfredo M. Ronchi

Abstract: This chapter provides an overview of the key questions posed to policy makers and regulators as they define a sound regulatory framework that will not unintentionally limit the evolution of this technology. How to reduce or even eliminate the range of concerns raised by AI? A selection of the most relevant approaches to the problem and already developed regulatory frameworks are included as examples to stimulate discussion and action.

The challenges of AI systems

Artificial intelligence has returned to the forefront after a long period of underground development, in an easily accessible and ready-to-use format. Mistakes and hallucinations generated using AI, particularly generative AI, can have very serious consequences. Apart from the nightmares often generated by Sci Fiction movies like War Games, Eagle Eye or even Lucy[1], this set of technologies has generated the need to regulate both the development and the deployment of its applications. Both experts in the field and governmental bodies are asking to issue specific regulations and laws, one of the most relevant voices in this field was due to San Altman in "OpenAI's Sam Altman Urges A.I. Regulation in Senate Hearing" [May 2023] and a newspaper article on "The world wants to regulate AI but does not quite know how" [The Economist - Oct 24th, 2023]. More recently in October 2025 over 850 people signed a statement calling for a pause in the development of superintelligence[2], among them tech leaders like Virgin Group founder Richard Branson and Apple cofounder Steve Wozniak.

The key questions concerning similar problems are:

- How to achieve precautionary/protective regulatory goals on AI design. Use & autonomy without stifling a jurisdiction's competitive advantage?
- At which level should AI be treated as a morally[3] considerate agent and how could this be regulated?
- Recognizing that assigning 'legal personality' does not necessarily mean 'morally considerate'. Should moral considerations be integrated into regulatory approaches?
- Can artificial intelligence be accountable and liable?

To better understand the range of concerns and the willingness to regulate a short list of shared and contrasting underlying concerns we need to consider:

- Security (homeland/national & defence secrets, military, space, weapons, infrastructure, financial systems, personal information, political processes/elections, commercial secretes, biosecurity).
- Discrimination and BIAS in health, selection processes, procurement, justice, policy, regulation, and anything where AI is involved.

- AI moral responsibility: in case a trusted AI system will be used to pose ethical dilemmas and feel released from a personal ethical analysis and related responsibilities.
- Information oversight, misinformation, disinformation, disapproved information filtering, nudging, opinion formation and manipulation.
- Privacy, human rights & civil liberties protections.
- Commercial advantage e.g. promoting national industry – via different methods.
- Protection of system and information integrity programs from being hijacked.
- Consumer protection including liability for harmful products or non-regulation compliant e-services[4].

We have seen more formal and also ad hoc approaches to these concerns. Mixed frameworks for national security, commercial advantage, justice, and biosecurity. In terms of national security and commercial advantage justice systems are already at work assigning accountability and liability in cases of AI and Gen AI managed systems. This is due to the proliferation of AI managed systems and devices that have gained a key role in systemic liability allocation models and this ensures accountability without stifling innovation. However, there are other areas evolving such as "cyberbiosecurity", an emerging field that combines cybersecurity, biosecurity, and cyber-physical security to protect biological and healthcare systems and data from digital attacks and malicious manipulation. Safeguarding sectors such as healthcare, agriculture, renewable energy, and the environment are part of the cyberbiosecurity agenda. Defending the "bioeconomy" from intrusions that exploit the growing connection between biotechnology and the digital world, protecting information, processes, and materials in this interconnected space is a critical agenda item. Shared preventative solution generally do not include precautionary measures as they focus too late with 'after the fact liability' to be effective. It should involve the evaluation of potential risks and threats and related management within regulatory frameworks and, if needed, mitigation actions.

Among some more integrative frameworks that we might consider:

- Traditional commercial liability for programme creators (downside v/s preventative effectiveness), how to behave in case of neural networks "black boxes" or Actionable - Generative AI?
- Environmental law preventative frameworks (downsides v/s competitive or innovative advantage). Regulations and jurisdiction are usually followers in this space, and need to be careful not to shape or limit a priori innovation, as stated in the US AI executive order about AI leadership "excessive State regulation thwarts this imperative"[5] or UNESCO's "Ethics of Artificial Intelligence" [6].
- Graduated "AI" focused liability as with other independent actors, this is one of the basic approaches in the EU AI Act[7].

Some noteworthy global AI regulatory frameworks

Among the international organizations involved are UNESCO who launched the AI Initiative[8] and published different reports on this topic including 'UNESCO Generative AI', 'UNESCO Generative

AI in Education'[9], and others[10]. IEEE published Ethically Aligned Design[11], a vision for prioritizing human well-being with autonomous and intelligent systems. Similar initiatives were launched by The Global Partnership on Artificial Intelligence (GPAI), by OECD in on Artificial Intelligence and Robotics, and by the United Nations Interregional Crime (UNICRI) Justice Research publication "Toolkit for Responsible AI Innovation in Law Enforcement[12]" as well as "AI for Safer Children[13]".

OECD

The OECD in 2019 published one of the first document providing some basic guidelines to regulate AI technology "Recommendation of the Council on Artificial Intelligence"[14] which became a key reference globally, including the definition of an AI system: "*Artificial Intelligence (AI) is a general-purpose technology that has the potential to: improve the welfare and well-being of people, contribute to positive sustainable global economic activity, increase innovation and productivity, and help respond to key global challenges. It is deployed in many sectors ranging from production, education, finance and transport to healthcare and security.*"

At the end of 2023, the OECD expanded its role, leveraging its unique expertise, cooperation infrastructure, and convening power to support and foster a positive and proactive message about AI and privacy. The key role of OECD is to act as an observer, and it has been historically a partner of the Global Privacy Assembly (GPA) - a global network of Privacy Enforcement Authorities (PEAs) working with the Council of Europe (CoE).

In early 2024, the OECD[15] formally launched the OECD AI Expert Group on AI, Data, and Privacy, bringing together leading AI and privacy experts worldwide (data protection authorities, policymakers, industry, civil society, and academia). The OECD policy observatory published the 2024 OECD AI principles update[16]. The OECD Ministerial Council Meeting held the same year offered the opportunity to update the principles on AI governance established for the first time in 2019, the first intergovernmental standard. The 2024 update considers new technological and policy developments, ensuring they remain robust and fit for purpose. The updated principles now address emerging challenges with an enhanced focus on safety, privacy, intellectual property rights and information integrity.

These principles defined the basis for innovative, trustworthy AI respectfully considering human rights and EU democratic values[17]. Up to now we count 47 countries that adhere to these principles that provide recommendations and guidelines to policy makers. This process allows the creation of AI risk frameworks, building a foundation for global interoperability between jurisdictions.

The principles are based on shared values and can be summarized:

- Inclusive growth, sustainable development and well-being (Principle 1.1)
- Human-centred values and fairness (Principle 1.2)
- Transparency and explainability (Principle 1.3)
- Robustness, security and safety (Principle 1.4)
- Accountability (Principle 1.5)

In the implementation of these principles the recommendations for policy makers include investing in AI research and development (Principle 2.1); fostering a digital ecosystem for AI (Principle 2.2); shaping an enabling policy environment for AI (Principle 2.3); building human capacity and preparing for labour market transformation (Principle 2.4); and international co-operation for trustworthy AI (Principle 2.5).

As stated by the OECD "*The Principles serve as a benchmark for responsible AI development and a critical checklist to address these rapid changes effectively, ensuring that AI continues to benefit society without compromising standards and safety.*[18]" In addition, as already mentioned, there are some core aspects such as interoperability of AI and the definition of AI systems and its lifecycle that must be considered.

Australia

Governments, at international level, are introducing new regulations to address the risks of AI, with a focus on creating preventative, risk-based guardrails that apply across the AI supply chain and throughout the AI lifecycle. The Australian government is no different and launched consultations on safe and responsible artificial intelligence, recognizing that the regulatory system at the time failed to address the specific risks that artificial intelligence would pose.

The Australian Government (AGA) uses the OECD's definition of an AI system: "*An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.*[19]"

To develop a regulatory environment that builds community trust and promotes AI adoption, the Australian Government opened consultations in September 2024. Proposed mandatory guardrails for AI throughout their lifecycles, and particularly in high-risk settings include:

- AI systems should benefit individuals, society and the environment.
- AI systems should respect human rights, diversity & the autonomy of individuals (termed human-centred vision).
- AI systems should be inclusive & accessible & should not involve or result in unfair discrimination against individuals, communities or groups (termed Fairness)
- AI systems should respect & uphold privacy rights and data protection & ensure the security of data.
- AI systems should reliably operate in accordance with their intended purpose.

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. 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 outcome of the AI system. Those responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI system, and human oversight should be enabled.

Canada

Canada has positioned itself as an early leader in AI governance, launching the Pan-Canadian Artificial Intelligence Strategy in 2017 with an initial investment of $125 million, making it one of the first national AI strategies globally. The strategy has since been expanded to support AI research excellence and responsible development.

The Department of National Defence (DND) and Canadian Armed Forces (CAF) define AI as "*the capability of a computer to do things that are normally associated with human cognition, such as reasoning, learning, and self-improvement*.[20]" More recently, Canada's regulatory framework defines an artificial intelligence system as any technological system that processes data related to human activities to generate content or make decisions, recommendations, or predictions.

Regulatory Framework

In September 2023, the Federal Minister of Innovation, Science and Industry announced the Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems[21]. This voluntary framework encourages industry self-regulation while the government develops binding legislation. In 2024, the government committed $2.4 billion in funding to support AI computing infrastructure, startup development, and the expansion of research capabilities.

The cornerstone of Canada's regulatory approach is the Artificial Intelligence and Data Act (AIDA), introduced as part of Bill C-27 (the Digital Charter Implementation Act, 2022). Currently proceeding through the legislative process, AIDA will introduce mandatory requirements for businesses developing or deploying high-impact AI systems, defined as those that may cause serious harm to health, safety, or human rights, or have significant adverse impacts on economic interests. The Act would establish an AI and Data Commissioner to oversee compliance and enforcement.

Under AIDA, businesses will face obligations across three key phases:

• **Design**: Businesses will be required to identify and assess risks related to harm and bias in their AI systems, implement appropriate mitigation measures, and maintain comprehensive documentation of their design processes and risk assessments.

• **Development**: Businesses will be required to evaluate the intended uses and limitations of their AI systems, ensure adequate testing and validation, and provide clear information to users about system capabilities and constraints.

• **Deployment**: Businesses will be required to implement ongoing risk mitigation strategies, establish monitoring mechanisms to detect issues during operation, and maintain incident response procedures. Continuous monitoring requirements ensure that systems remain safe throughout their lifecycle.

AIDA includes substantial enforcement mechanisms, with penalties of up to CAD $25 million or 5% of global revenue for serious violations, demonstrating Canada's commitment to meaningful accountability.

Regulatory Philosophy

The Canadian government has adopted a risk-based, flexible policy approach where obligations are proportional to potential harms. Higher-risk AI systems face more stringent requirements, while lower-risk applications have lighter compliance burdens. This framework builds on existing best practices and emphasizes interoperability with international regulatory approaches, particularly the EU AI Act. By aligning with common standards and risk assessment methodologies, Canada aims to facilitate cross-border innovation while maintaining robust protections.
Canada's approach also recognizes the role of sector-specific regulators in overseeing AI applications within their domains, including healthcare, financial services, and transportation, creating a coordinated governance ecosystem that balances innovation with safety and rights protection.

China

China invested significant resources to lead the competition in the AI domain. As a result, China is leading the way in AI regulation releasing new strategies to govern algorithms, chatbots and more.

One of the definitions of AI stemming from Chinese research is that “Artificial intelligence (AI) aims to mimic human cognitive functions and execute intellectual activities like that performed by humans dealing with an uncertain environment”[22]. The Chinese approach to this technology is in line with Chinese culture and foresees, in recent rules, that AI must always be under human control (since the advent of Generative AI as a precautionary approach). To promote social cohesion, additional regulation prohibits algorithms that create division.

Early in 2021/2022 China developed and implemented detailed/binding AI regulation on companies with regards to content and user rights. A specific part of regulation is devoted to worker protections against gig algorithmic scheduling[23].

The Cyberspace Administration of China, the National Development and Reform Commission, the Ministry of Education, the Ministry of Science and Technology, the Ministry of Industry and Information Technology, the Ministry of Public Security, and the National Radio and Television Administration jointly released the Interim Measures for the Management of Generative Artificial Intelligence Services[24] (the "AI Measures"), which is the first administrative regulation on the management of Generative AI services, and it came into effect on August 15, 2023.

On September 2024 China proposed new regulations that would make it mandatory labelling of AI-Generated Content to clamp down on a surge in AI-related fraud. Rules requiring watermarks identifying deep fakes, protecting people’s “likeness rights” or harm the nation’s image became protected via an Algorithm registry[25]. Regulation was created in consultation with academia, government officials, journalists, private researchers, public debate, advocacy efforts, workshops, and more. Draft guidelines were published on September 14 and open for public comment one month in the “Notice on soliciting opinions on the mandatory national standard (draft for comments) of "Network Security Technology Artificial Intelligence Generation Synthetic Content Identification Method"”. This marked the first time that the Cyberspace Administration of China had proposed specific rules regarding the labelling of AI-generated content.

On November 29, 2025, Xi Jinping, general secretary of the Communist Party of China (CPC) Central Committee, has emphasized the importance of improving long-term mechanisms for cyberspace governance, Xi called for sustained efforts to cultivate a clean, healthy and sound online environment.

On October 30, 2025, at the Second International Conference on Cooperation and Development for Young Entrepreneurs, Chen Wenling, former chief economist of the China International Economic Exchange Centre, presented its ongoing plans to become the world's first power-producing country to excel in artificial intelligence systems.

On December 5, 2025, China's innovation is open and open-source, and the country is willing to share indigenous technologies and innovation scenarios with the world based on a briefing from Chinese foreign ministry spokesperson Lin Jian - "*Innovation' has indeed become a key word in China's economic and social development.*" He outlined the difference between competition in innovation and leading position; China is promoting fair competition among players in a mutual cooperation framework.

European Commission & Council of Europe

The EU definition: "*an 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.* [26]"

The European Commission clearly distinguish between "*simpler traditional software systems or programming approaches ... that are based on the rules defined solely by natural persons to automatically execute operations. A key characteristic of AI systems is their capability to infer*[27]".

The EC joint Research Centre explored the opportunity AI-generated synthetic data presents for privacy-safe data use. In such specific field EU Regulations adopt a risk-based approach to regulation. AI should be as neutral as possible to cover techniques that are not yet known or developed.

On 17 May 2024, "*the Council of Europe adopted the first-ever international legally binding treaty*[28] *aimed at ensuring the respect of human rights, the rule of law and democracy legal standards in the use of artificial intelligence systems. The treaty, which is also open to non-European countries, sets out a legal framework that covers the entire lifecycle of AI systems and addresses the risks they may pose, while promoting responsible innovation. This document was the outcome of two years of work due to an intergovernmental body, the Committee on Artificial Intelligence (CAI) composed by 46 CoE members states, 11 non-member states*[29]*, the European union plus the private sector, civil society and academia, who participated as observers. The convention adopts a risk-based approach to the design, development, use, and decommissioning of AI systems, which requires carefully considering any potential negative consequences of using AI systems.*"

Few days later, on 21 May, the Council of the European Union adopted the EU AI Act, which once published in the EU Official Journal in June, became the first set of AI regulations that has undergone a full legislative approval process. The EU AI Act is structured in 113 articles and counts 13 annexes, its holistic risk-based approach is suitable for any player in the field of AI from developers to deployers.

The EU vision on AI can be summarised as “beyond making our lives easier, AI is helping us to solve some of the world's biggest challenges: from treating chronic diseases or reducing fatality rates in traffic accidents to fighting climate change or anticipating cybersecurity threats.[30]” Dealing with a fast evolving technology the key aspect in defining a framework is to find a golden balance between shaping the evolution and leaving to technology the freedom to evolve naturally.

On September 23, 2025, Politico[31] issued an article stating that “*just one year after the European Union adopted a landmark plan to cut risks of artificial intelligence, it’s already preparing to put the brakes on.*” This decision was mainly due to big company CEOs lobbying and the unavailability of the EU standards defining high risks applications and the consequent impossibility to comply with such standards within the foreseen enforcement date in August 2025. Some EU governments supported a delay of at least twelve months to enforce the AI Act. At the United Nations General Assembly on September 2025, the need to regulate AI was raised.

United States

There has been a significant shift in AI policy in the U.S. since President Trump took office in 2025. President Trump rescinded President Biden's executive order within hours of assuming office on January 20, 2025, labeling it among "unpopular, inflationary, illegal, and radical practices". Three days later, Trump issued his own executive order titled "Removing Barriers to American Leadership in Artificial Intelligence[32]" signalling a shift from oversight and risk mitigation toward deregulation and promotion of AI innovation. Trump's approach emphasizes removing barriers to US AI leadership and frames AI development as essential to national and economic security without safeguards for the public. Going a step further, in December 2025, President Trump signed an additional executive order blocking states from enforcing their own AI regulations, directing the Attorney General to establish an AI Litigation Task Force to challenge state AI laws and calling for a single national framework[33].

In October 2023, former President Biden issued an Executive Order establishing comprehensive AI governance focused on safety, security, and trustworthy development. This order directed more than 50 federal entities to undertake more than 100 specific actions across eight policy areas, including safety and security, civil rights protection, privacy safeguards, and international leadership. The order required companies developing high-risk AI models to conduct safety tests and share results with the government, established AI safety standards, and promoted measures to address algorithmic discrimination in areas like housing, healthcare, and employment.

The most significant policy difference lies in regulatory philosophy and equity considerations. The Biden order explicitly sought to address discrimination and bias in AI applications, incorporating equity and civil rights protection throughout its framework. In contrast, the Trump order did not address these concerns, marking a departure from government intervention in AI ethics and fairness. The competing visions represent fundamentally different approaches: Biden's comprehensive regulatory framework, which prioritizes safety, equity, and risk mitigation, versus Trump's deregulatory stance, which emphasizes innovation, global competitiveness, and minimal government intervention.

India

In India, Artificial Intelligence (AI) refers to the development of computer systems that emulate human intelligence, performing tasks like thinking, learning, and decision-making to enhance human life and work. This definition aligns with global understanding but is often contextualized within

India's national strategy, which focuses on applying AI to societal challenges in areas such as healthcare, agriculture, and education to drive inclusive growth and societal transformation, as outlined by NITI Aayog[34] and the Ministry of Electronics and Information Technology (MeitY)[35]. Key Aspects of India's Perspective on AI include the Emulation of Human Intelligence, Complementing Human Capabilities, Focus on Societal Impact, Strategic National Approach, Data-Driven Learning, and Interdisciplinary Foundation.

India has no specific AI law but regulates the technology through existing frameworks like the Digital Personal Data Protection Act (DPDP) 2023[36] and the Information Technology (IT) Act, 2000[37]. The government also issues advisories, such as the MeitY advisory[38], requiring permission for deploying certain AI models and mandating steps against algorithmic discrimination and deepfakes. The broader strategy emphasizes a pro-innovation approach, balancing rapid development with ethical considerations like fairness, accountability, and transparency for AI systems.

Israel

Israel refers to the updated OECD definition "*An artificial intelligence system is a machine-based system that, for explicit or implicit objectives, infers from the input it receives how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. AI systems vary in their level of autonomy and adaptability after deployment.*"

Israel unveiled its first comprehensive AI policy, led by the Ministry of Innovation, Science and Technology and the Ministry of Justice to foster responsible AI innovation while addressing concerns related to bias, transparency, safety, accountability, and privacy. Minister of Innovation, Science and Technology Ofir Akunis ascerts " *The revolutionary impact of AI technology is yielding constant improvements in the quality of life of Israel's citizens in countless fields of activity, including in health care, science, technology, agriculture, education, and security. However, together with its many benefits, there are also many risks. These principles we have published facilitate development and responsible innovation, enabling the use of AI, while safeguarding basic rights and the public interest.*"

Key Highlights of Israel's AI Policy: Israel's AI Policy addresses seven challenges of private sector AI use, including discrimination and privacy and uses collaboration with stakeholders and alignment with OECD recommendations. The policy emphasizes "Responsible Innovation," balancing innovation and ethics.

Japan

The official definition of Artificial Intelligence in Japan[39]: "*artificial intelligence-related technology refers to the technology necessary to realize the function of replacing the intellectual ability related to human cognition, reasoning and judgment by artificial methods, and the information entered is processed using the technology, and It refers to the technology related to the information processing system to realize the function of outputting the results.*[40]"

The Japanese approach to AI governance doesn't impose rigid obligations and heavy fines, AI governance model encourages AI adoption while managing risks through existing frameworks, this is rooted in the innovation first model. This framework establishes a multi-level responsibility model:

- Government sets the vision and national policies.
- Local organizations adapt and apply AI to regional contexts.
- Universities and R&D institutions drive innovation and talent development.
- Businesses apply AI to create value and comply with regulations.

Together, these responsibilities ensure AI is promoted strategically, inclusively, and sustainably across society.

Saudi Arabia

Saudi Arabia was one of the first countries to establish the National Centre for AI (NCAI). Their definition of AI is: "*Systems that employ methods to gather data and use it to predict, recommend, or make decisions with varying degrees of autonomy, selecting the best course of action to achieve specific objectives*[41]".

The regulatory arrangement of the Saudi Data and AI Authority's Laws and Regulations (SDAIA) were issued by Council of Ministers Resolution No. (292) and amended by the Council of Ministers Resolution No. (195). Saudi Arabia aims to implement global best practices for data management and governance to improve performance, productivity and public services. The Kingdom seeks to leverage data as an economic resource to foster innovation, support economic transformation, and enhance national competitiveness. The Saudi Data and AI Authority and the National Data Management Office have created a national data governance framework, approved by the SDAIA Board of Directors.

The National Data Management and Personal Data Protection Standards are based on a directive from the Saudi Authority for Data and Artificial Intelligence. These standards, covering 15 domains, are intended for adoption by all Public Entities in the Kingdom. To support the development of the Data Management and Personal Data Protection standards[42], a set of international references, internal relevant policies and regulations, and guiding principles were defined. Government Entities must implement the standards, and compliance will be measured yearly to monitor progress and drive efforts towards a successful implementation. The Data Management and Personal Data Protection Standards[43] cover details for prioritizing and implementing:

1. Data Governance Domain
2. Data Catalog and Metadata Domain
3. Data Quality Domain
4. Data Operations Domain
5. Document and Content Management Domain
6. Data Architecture and Modelling Domain
7. Data Sharing and Interoperability Domain
8. Reference and Master Data Management Domain
9. Business Intelligence and Analytics Domain
10. Data Value Realization Domain
11. Open Data Domain

12. Freedom of Information Domain
13. Data Classification Domain
14. Personal Data Protection Domain
15. Data Security and Protection Domain

Do we need a global regulatory framework for AI?

The willingness to create a regulatory framework devoted to Artificial Intelligence is already a reality - several conferences or spokespersons from various sectors are concerned about the potential impact of AI on their activities and businesses. Some of them are concerned over the impact on jobs and careers, others are concerned about human rights and ethics, or the possible control over human lives. Among additional concerns, this chapter did not mention AI in cyberwarfare and intelligent autonomous weapons. This topic deserves an entire chapter including how AI influences opinions and could manipulate behaviour. Also, the use of AI in social engineering, or nudging and opinion manipulation boosted by AI. The use of AI as a filter in "Chat Control Regulation[44]" will be completely approved by the EU or a similar use of AI in the eServices Act[45] , adapting services to AI logic standards.

Regulatory bodies are faced with questions about emerging technologies that are not yet fully developed, and consider their own impact in influencing the development of AI such thus potentially biasing the evolution of this technology. This has already occurred in digital global domains such as intellectual property rights (IPRs) where privacy or cybersecurity regulations are attempting to deal with global "borderless" technologies. The hope is to release or promulgate a global regulatory framework limiting as much as possible unregulated "oasis". Currently, each country or institution is promoting or enacting its own solution. Sometimes approaches differ, other times they use similar principles.

One of the typical taxonomies in describing the levels of AI autonomy includes:

- **Level one** – basic automation, driven by fixed rules and scripts. At this stage, an AI or system follows predetermined instructions to execute simple, repetitive tasks, with minimal intelligence. Pros: Quick execution of repetitive tasks without errors, consistency, and reduced labour requirements for repetitive and tedious work. Cons: Rigidity of rule-based systems (lack of adaptability) and inability to handle exceptions or complexity.
- **Level two** – partial autonomy which incorporates some machine learning or adaptive capabilities where AI can make limited decisions on its own within a narrow scope but still requires human guidance or validation for most outputs. Pros: Better decision support, AI handles the analysis, greater accuracy than human labour, better results over time thanks to adaptive learning. Humans remain in control, so quality and ethical oversight are robust. Cons: If humans overly rely on AI, errors and hallucinations can happen. Human review is required because AI may provide incorrect recommendations outside of its training. To integrate machine learning systems into workflow it is a complex task.
- **Level three** - conditional autonomy, a more advanced AI that can make conditional decisions and act independently within well-defined circumstances. Pro: AI autonomously handles routine cases, improved efficiency, increased throughput. Humans fucus the attention only on exceptions. Contra: in case of exceptions errors can occur if the AI doesn't correctly alert humans. Systems require well-defined limits and real-time monitoring. Human supervisors must ensure that AI's decisions in its autonomous zone are correct and auditable. on the social side the duty transfer from humans to AI can cause some frictions.

- **Level four** - high autonomy where AI is capable of making independent decisions in complex scenarios with only minimal human oversight. Often foreseen in supply chain and logistics realms. Pros: Level 4 autonomy are game-changing, a high-autonomy operation can run continuously and react in real time to data. Cons: human oversight is minimal, so any mistake the AI makes can propagate widely before a human notice. AI must be trained to handle a wide range of scenarios to win the challenge of the "unknown unknowns.". Ex-post amendments to the system in case of hallucinations are hard to amend (the "black box" problem). This raises the risks concerning governance, ethics and accountability. Data governance is also paramount since the AI is largely on its own, feeding it high-quality, up-to-date data and preventing bias or drift is essential to keep decisions reliable. Another consideration is security – highly autonomous systems could be targets for cyberattacks, as malicious actors might try to manipulate an AI that controls critical operations warns of risks like "Agentic AI-driven cyberattacks" if guardrails are not in place). Companies may reshape the job roles to better cope with these new technologies.
- **Level five** - full autonomy where AI systems operate completely independently of human intervention. Pros: A Level 5 AI can handle any task, decision, or scenario within its defined domain just as a human expert would. Adequately training AI can learn and adapt but this is limited to a specific domain of knowledge not extended to Artificial General Intelligence (AGI). No need to interact with humans in daily operations, humans set high level directions AI agent does the rest. Full autonomy if implemented in AI driven factory as an example, can offer improved flexibility as ramping production up or down immediately. Cons: Level 5 comes with serious risks and responsibilities and raises profound issues of trust, ethics, and governance. Users must have absolute confidence in AI's reliability because having no human oversight there is no "safety net".

If ethics, moral principles and human rights are some of the key concerns in AI design, development and deployment these aspects are even more concerning in case of level five systems. Human responsible must be sure that the AI's decision aligns with human cultural identity, values and ethics. UNESCO promotes the idea that we should develop local AI platforms that reflect local culture and values. In addition, there is a potential societal impact to consider - the autonomous AI system following a logic of optimisation can displace large segments of the existing workforce while also creating new business opportunities.

Some research groups propose to create a graduated liability-based AI Agency with oversight over various levels of legal responsibility, accountability, and regulatory burden based on the risk, autonomy and impact of AI. The following levels of AI scenario could be considered:

- **Level Zer**o – non generative AI – Preventative programming restrictions along the lines of environmental, discrimination, privacy regulation, penalties based on harm/level of programming malfeasance.
- **Level One** – Generative or semi-independent AI, or AI as an independent actor in the sense of other nonhuman independent actors with less "autonomous agency leeway". Animal law, preventive programming restrictions and responsiveness (//training' responsibilities) (NZ Animal Welfare Act 1999 has provision for the Governor General to declare no listed entries as "animals"). Responsibility for harm lies with the creator or custodian of "animals" – and penalties lie with both AI and creator/custodians, confinement (redesign) or destruction for the former, financial or custodial penalties for the latter, up through manslaughter and beyond, depending on programming malfeasance.
- **Level Two** – AI as independent actor in the sense of human independent actors – AI creating new AI. Among the questions at this level is should this trigger morally based responsibilities and processes around control/penalties?

Conclusions

The shared willingness among private and public subject to create a regulatory framework devoted to Artificial Intelligence is almost a reality, as reported above several countries and even organisations developed their guidelines or even regulations. Several conferences or spokespersons pertaining completely different sectors are concerned, as it happened on the dawn of new technologies, about the potential impact of AI on their activities and businesses.

The initial trend from humans to AI systems, usually associated with "black boxes," considered human supervision to be 'not mandatory' and even unnecessary, as it created a significant time lag compared to decisions and actions that AI could execute on the fly[46]. This results in critical decisions being entrusted to automated decision-making software, with no room for human intervention.

Following the actual trend to embed AI at the decision level in a broad range of fields the legal system will soon face the problem to assign responsibility when AI systems will cause harm. How will jurisprudence define a proper methodology to evaluate AI decision-making case? If an algorithmic decision-making system can be analysed and evaluated or amended, the dynamic outcome of the AI system will make the case more complicated.

The intrinsic nature of "digital" immaterial, cloneable, and borderless suggests finding an agreement on a global regulatory framework addressing AI in its different phases, design, development and deployment, in any case regulations will carefully consider the different cultural models to do not flatten the outcome to a single cultural model.

References

1. Science fiction movies already proposed similar scenarios e.g. Wargames (1983 American techno-thriller film directed by John Badham) or Eagle Eye (2008 American action-thriller film directed by D. J. Caruso), Lucy (2014 French Sci Fi movie written and directed by Luc Besson)
2. https://superintelligence-statement.org/
3. Ethics as Moral Philosophy finds is root / etymology in the ancient Greek term "ἦθος," (ethos), this term refers to the characteristic spirit of a culture, era, or community as manifested in its beliefs and aspirations, so it is strictly related to cultural identity.
4. https://commission.europa.eu/news-and-media/news/digital-services-act-keeping-us-safe-online-2025-09-22_en
5. https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/
6. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics?hub=32618
7. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
8. https://www.unesco.org/en/artificial-intelligence
9. https://www.unesco.org/en/articles/generative-artificial-intelligence-education-think-piece-stefania-giannini
10. https://www.govtech.com/education/k-12/unesco-reassert-public-control-over-generative-ai
11. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9398613
12. https://unicri.it/topics/Toolkit-Responsible-AI-for-Law-Enforcement-INTERPOL-UNICRI
13. https://unicri.org/topics/AI-for-Safer-Children
14. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
15. OECD's definition of an AI system, OECD Artificial Intelligence Papers (2024), OECD March 2024 No. 8

16. Care of Juraj Čorba, Audrey Plonk, Karine Perset, Yoichi Iida
17. https://european-union.europa.eu/principles-countries-history/principles-and-values/aims-and-values_en
18. https://oecd.ai/en/wonk/evolving-with-innovation-the-2024-oecd-ai-principles-update
19. OECD's definition of an AI system, OECD Artificial Intelligence Papers (2024), OECD March 2024 No. 8
20. https://www.canada.ca/en/department-national-defence/corporate/reports-publications/dnd-caf-artificial-intelligence-strategy/what-is-ai.html
21. https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems
22. Chen Shuang - School of Chinese Medicine, University of Hong Kong, Hong Kong 999077, China
23. https://www.ctol.digital/news/china-algorithm-regulations-protect-workers-users-fairness-online/ - - https://www.chinajusticeobserver.com/a/spc-issues-guiding-cases-on-gig-worker-protection
24. https://www.chinalawtranslate.com/en/generative-ai-interim/
25. https://www.chinalawtranslate.com/en/proposed-ai-content-labeling-guidelines/ https://www.chinalawtranslate.com/en/ai-labeling/
26. Source: Artificial Intelligence Act https://artificialintelligenceact.eu/article/3/
27. https://ai-act-service-desk.ec.europa.eu/en/ai-act/recital-12
28. https://search.coe.int/cm#{%22CoEIdentifier%22:[%220900001680afb11f%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}
29. Argentina, Australia, Canada, Costa Rica, the Holy See, Israel, Japan, Mexico, Peru, the United States of America, and Uruguay
30. Artificial Intelligence for Europe
31. https://www.politico.com
32. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/
33. “Highlights of the 2023 Executive Order on Artificial Intelligence for ...https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/
34. https://niti.gov.in/
35. https://www.meity.gov.in/
36. https://www.meity.gov.in/static/uploads/2024/06/2bf1f0e9f04e6fb4f8fef35e82c42aa5.pdf
37. https://dhsgsu.edu.in/images/Reading-Material/Law/UNIT-IV-Second.pdf
38. https://www.meity.gov.in/static/uploads/2024/02/9f6e99572739a3024c9cdaec53a0a()ef.pdf
39. https://laws.e-gov.go.jp/law/507AC0000000053
40. https://laws.e-gov.go.jp/law/507AC0000000053#Mp-Ch_1-At_2
41. https://my.gov.sa/en/content/109729#section-1
42. The term “Sector Risk Management Agency” has the meaning set forth in 6 U.S.C. 650(23).
43. E.g., high pace decision making environments like stock exchange market, auto pilots, critical infrastructures real time management.
44. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52022PC0209
45. https://digital-strategy.ec.europa.eu/en/policies/digital-services-act
46. E.g., stock exchange market, real time control systems, etc.