Digital Policy Alert

Refined Handbook for the Mapping of AI Rules

Tommaso Giardini and Nils Deeg

November 2026

An initiative of the
St.Gallen Endowment
for Prosperity through Trade

Refined Handbook for the Mapping of AI Rules

This handbook explains the methodology developed by the Digital Policy Alert (DPA) team to systematise domestic AI rules and international AI frameworks.

- The Approach section explains how we developed our taxonomy of AI concerns and regulatory requirements, using the DPA database and Clairk AI library.
- The Taxonomy section provides an in-depth explanation of our taxonomy of AI concerns and regulatory requirements, with concrete examples from across the world.
- The Scope section outlines the texts covered by our analyses and tools.

To understand AI rules in depth, you can interact with the dataset on our AI Governance Hub and read our flagship report on global AI rules. This handbook refines previous work by our team on global AI rules and the interface between the domestic and international level. It was made possible with the generous support of the Patrick J. McGovern Foundation, for which we are grateful.

Approach

The DPA team specialises in systematising government regulation of the digital economy: We track policy and enforcement developments across 60 jurisdictions on a daily basis, totalling over 27000 entries since January 2020. To make sense of this data, we consistently apply the DPA taxonomy, which structures the policy instruments governments use and digital business models governments target, among other elements of digital policy.

The pace and importance of AI regulation keeps growing worldwide, creating demand for a dedicated taxonomy focusing specifically on how governments regulate AI. Currently, governments are experimenting with different approaches to tackle novel policy concerns raised by AI. Our dedicated AI taxonomy reveals that governments share these concerns and that the regulatory requirements to address them are diverse, but also recurring. Specifically, our taxonomy covers:

- ten AI concerns that governments seek to address: transparency, safety, consumer protection, data governance, human oversight, discrimination and bias, information integrity, labour, intellectual property, and accountability; and
- 50 granular regulatory requirements that governments use to address said concerns.

To identify the concerns, we analysed the DPA database, including over 2000 policy and enforcement developments targeting the AI sector, and the text of nine international AI frameworks. The ten shared concerns are the starting point of the taxonomy.[1] We tag these concerns into the DPA AI dataset and our Clairk library of international AI frameworks.

To systematise the regulatory requirements, we investigated the text of 84 domestic AI laws and regulations, including current legislative proposals.[2] Since we developed the requirements using the text of existing AI regulations, the taxonomy is more extensive in salient policy areas, such as safety. We tag requirements into the provisions of the Clairk library of domestic AI regulations.

This approach reveals which concerns governments prioritise, by showing how concerns appear in domestic AI regulation and international AI frameworks, and how governments tackle each concern, by demonstrating the range of regulatory requirements. To enable readers to conduct their own analyses, the sections below explain the taxonomy and scope of this project in detail. Our reports provide an example of how to draw on the DPA data, taxonomy, and technology for analysis.

Finally, we empower users to engage with our data in real time with unprecedented sophistication:

- The DPA's AI Governance Hub provides a dedicated page for each concern, with daily updates. Each page shows the relevant DPA entries and the provisions of domestic AI regulations and international AI frameworks in full text, structured using our taxonomy.
- The St.Gallen Endowment's technology suite empowers users to interact with our data, including our text analysis tool Clairk, our chat system (trained exclusively on verified data), as well as our API and MCP integrations, which feed our data directly into users' systems.

[1] Less salient concerns, including competition and sustainability, are not currently covered by our scope.
[2] We focus on regulatory requirements, meaning compliance obligations for private sector actors. Government action relating to the development of the AI sector and government use of AI are not in scope.

Taxonomy

This section explains each concern, defines the corresponding regulatory requirements with examples, and sets out how we distinguish and address overlaps between regulatory requirements.

- Transparency
- Safety
- Consumer Protection
- Data Governance
- Human Oversight
- Discrimination and Bias
- Intellectual Property
- Labour
- Information Integrity
- Accountability

Transparency

Regulatory Requirements

Governments are concerned about the opacity of AI systems, whose lack of transparency prevents users from making informed choices and regulators from overseeing AI systems. By demanding that AI firms disclose information about the use and functioning of AI systems, as well as the input and output of their systems, governments aim to start closing this transparency gap.

Our taxonomy includes eight distinct regulatory requirements regarding AI transparency:

- **Content marking**: Provisions requiring AI-generated content to be marked in order to be identified as such. This includes visible labelling of AI-generated content and watermarking of AI-generated content for identification through technical means. *Example: Article 11 of Viet Nam's AI Law requires providers to mark AI-generated audio, images, and video in a machine-readable format and requires deployers to visibly label AI-generated or manipulated audio, images, and video that simulate a real person or recreate a real event.*
- **Disclosure**: Provisions requiring prominent public disclosure of AI system use and functioning in ways that clearly draw attention to the element being disclosed. These include requirements to clearly disclose to users that they are interacting with an AI system (*Example: Article 50(1) of the EU AI Act requires providers to design and develop AI systems designed to interact directly with natural persons in a way that allows affected persons to be informed of the fact that they are interacting with an AI system*) and requirements to prominently display certain information about an AI system (*Example: Article 26 of China's Provisions on the Management of Algorithmic Recommendations in Internet Information Services requires service providers to prominently display filing numbers on public facing websites and applications*).
- **Documentation**: Provisions demanding the documentation and storing of relevant information about AI systems and procedures. *Example: Article 20 of Brazil's draft AI Bill would require providers or operators of high-risk AI systems to document the operation of the system and the decisions involved in its construction, implementation and use across relevant lifecycle stages, including design, development, evaluation, operation and discontinuation.*
- **Information sharing**: Provisions requiring AI firms to share information on AI systems, including system functioning and proper use. *Example: Article 16 of China's Provisions on the Management of Algorithmic Recommendations in Internet Information Services requires algorithm recommendation service providers to provide information about the basic principles, purposes, and operating mechanisms of such services.*
- **Information rights**: Provisions requiring AI firms to provide meaningful explanations of AI involvement, such as decisions made by AI systems, to those affected by AI systems upon request. *Example: Article 86 of the EU AI Act gives persons subject to a decision taken on the basis of a high-risk AI system that produces legal effects having an adverse impact on them a right to a clear and meaningful explanation of the role of AI systems in the decision process and the main elements of the decision.*
- **Logging**: Provisions requiring AI firms to maintain systematic internal records of system input, output, and relevant events. *Example: Article 9 of Chile's Law regulating AI systems,*

robotics and related technologies would require a high-risk AI system, in order to be authorised, to have an automatic system for recording events while it is operating.

- **Notification**: Provisions demanding notification of specific parties, such as regulators or affected parties, reactively upon a defined event, such as a safety incident. *Example: Section 1421(4) of New York's RAISE Act requires large developers to disclose each safety incident affecting a frontier model to the Attorney General and the Division of Homeland Security and Emergency Services within seventy-two hours of learning of the incident or facts supporting a reasonable belief that one occurred.*
- **Reporting**: Provisions demanding proactive, periodical reporting of information about AI systems to the public or a regulator. *Example: Section 22757.12. of California's Transparency in Frontier AI Act requires large frontier developers to send a summary of assessments of catastrophic risks from internal use of frontier models every three months.*

Delineations

Disclosure and information sharing are distinct insofar as the function of disclosure requirements is to draw attention to the information being displayed. A disclosure requirement will therefore refer to the prominence of information. The paradigmatic example is a disclosure that AI is in use, or has been used in reaching a decision. Information sharing requires information to be published, but without the need for it to be immediately prominent to users. It can include requirements to publish system details on a company website or providing instructions to users.

Transparency requirements are often related to each other. For instance, both disclosure and content marking requirements aim to inform users about AI use. Content marking is a specific form of disclosure, focusing specifically on the AI output, whereas disclosure broadly covers the use of AI. Logging is a specific, automated form of documentation. We always tag specific requirements, where applicable, and use the broader requirements only if the specific ones do not apply.

Notification requirements are distinct from other transparency requirements since they are triggered by specific events, such as the discovery of a serious incident, whereas disclosure and reporting requirements are continuous. When such an event occurs, a notification must be made regardless of a request from a user (information rights) or authority (accountability).

Transparency requirements can also interact with requirements from other concerns. Logs, for instance, can inform monitoring and other safety requirements. Documents kept on record may have to be provided to authorities upon request, which is an aspect of accountability.

Safety

Regulatory Requirements

AI systems can pose risks to people, property and critical infrastructure, whether through malfunction, misuse or adversarial attack. While cyberattacks dominated the headlines, governments have built safety rules that span a system's whole lifecycle, from design and testing requirements during development, risk management, monitoring and corrective measures post-deployment, and outright prohibitions of certain systems that pose unacceptable risks.

Our taxonomy includes thirteen distinct regulatory requirements regarding the safety of AI:

- **Authorisation**: Obligation for AI providers to obtain approval for a specific AI system before developing it, offering it on the market, or putting it into service. *Example: Article 17 of Mexico's draft AI Law would require AI developers or providers to request authorisation from the competent authority before making an AI system available on the market.*
- **Corrective measures**: Provisions obligating AI firms to address problems that arise after deployment, including fixing issues or recalling AI systems. *Example: Article 20(1) of the EU AI Act requires providers of high-risk AI systems to take corrective actions when they consider that a system is not in conformity with the Act's requirements.*
- **External assessment**: Provisions requiring an external body to assess, certify or audit AI systems. This includes assessments conducted by government bodies (*Example: Section 26 of Kenya's draft AI Bill would require the AI Commissioner to review risk assessments for systems used in the public sector*) and by independent third-parties (*Example: Article 13 of Viet Nam's AI Law requires high-risk AI systems to undergo conformity assessments by recognised organisations.*)
- **Licensing:** Provisions mandating a licence in order for AI firms to operate in a given sector. *Example: Article 13 of China's Provisions on the Management of Algorithmic Recommendations in Internet Information Services requires algorithm recommendation service providers who provide internet news information services to obtain licences.*
- **Local operations**: Provisions demanding AI firms to establish or maintain a legal presence within a jurisdiction or designate a local representative. *Example: Article 14(6) of Viet Nam's AI Law requires foreign suppliers of high-risk AI systems to have a local legal point of contact, while suppliers of high-risk AI systems subject to mandatory conformity certification must also have a commercial presence or authorised representative in Viet Nam.*
- **Monitoring**: Provisions to continuously monitor and review AI system behaviour. This includes monitoring by AI firms (*Example: Article 32 of the Republic of Korea's Framework Act on AI requires business operators of systems above the computing threshold to monitor and respond to safety accidents*) and measures to enable monitoring by government bodies (*Example: Article 15 of Mexico's draft AI Law would require developers and providers of AI systems to ensure that systems can be monitored by the competent authorities while in use*).
- **Prohibition**: Provisions prohibiting certain AI systems or uses. This includes prohibitions of AI systems that pose a specific danger (*Example: Article 5 of Colombia's draft Law regulating AI would prohibit AI systems that do not allow humans to correct failures and exercise control over a systems' lifecycle*) or violate fundamental rights (*Example: Article 5 of Chile's draft Law regulating AI systems would prohibit AI systems that are incompatible with fundamental rights*

of people), as well as the use of AI for illegal activities (*Example: Article 7(2)(a) of Viet Nam's AI Law prohibits the use of AI to carry out acts prohibited by law*).

- **Registration**: Provisions requiring AI systems to be registered with a regulatory authority or on a public database, without requiring regulatory approval. *Example: Section 27 of Kenya's draft AI Bill would require the AI Commissioner to maintain a public register of high-risk AI systems.*
- **Risk and security assessment**: Provisions requiring AI firms and other actors to assess the risks, harms, impacts and security implications of their own AI systems proactively, often prior to deployment. *Example: Article 32 of the Republic of Korea's Framework Act on AI requires AI business operators to assess risks of AI systems in which the cumulative amount of computation for learning meets the threshold set by Presidential Decree.*
- **Risk management**: Provisions demanding measures to address the risks resulting from the development and deployment of AI systems. *Example: Article 25 of China's Interim Measures for Administration of Anthropomorphic AI Interaction Services requires app stores and distribution platforms to fulfil security management responsibilities with regard to covered services.*
- **System safety**: Provisions requiring cybersecurity and system resilience measures in AI system design and deployment. *Example: Article 15 of the EU AI Act requires high-risk AI systems to be designed to achieve certain levels of accuracy, robustness, and cybersecurity.*
- **Testing**: Provisions mandating tests for AI systems, including on their safe functioning and conformity with defined safety standards. *Example: Article 11 of Brazil's draft Bill on AI development, implementation, and application would require providers and operators of high-risk AI systems to conduct reliability testing, including tests of robustness, accuracy, precision and coverage.*
- **User identification**: Provisions requiring AI firms to identify users who access an AI system. *Example: Article 12 of China's Interim Measures for Administration of Anthropomorphic AI Interaction Services requires service providers to register users and collect identifying information about them as a condition to providing access to the service.*

Delineations

Safety requirements can be similar to each other but are distinct, as evidenced by registration, authorisation and licensing requirements. Registration requires no action by the regulator, authorisation requires the approval of a specific AI system, while licences are given by government bodies to AI providers. Similarly, risk assessment refers to the evaluation of risks, while risk management covers how said risks are to be addressed. When a provision includes both requirements, we tag both.

Overlaps between safety requirements and requirements under other concerns are also common. Prohibitions under safety cover all prohibitions that do not fall under another concern. For example, a prohibition of subliminal techniques falls under the prohibition of manipulative/deceptive practices (consumer protection), a prohibition of discriminatory social scoring falls under non-discrimination (discrimination and bias), a prohibition on systems that scrape biometric data falls under prohibition of certain data use (data governance), and the prohibition of certain AI systems at the workplace falls under worker protection (labour). Another example of similar but distinct requirements is system safety, which governs AI systems' overall operational integrity, whereas data security (data governance) focuses specifically on the data component.

Consumer Protection

Regulatory Requirements

Governments are concerned that AI can deceive users, manipulate their behaviour, and lead to addiction. Protecting consumers in an era where AI is becoming ubiquitous is a shared goal worldwide. This includes regulating AI firms' and AI systems' behaviours, as well as establishing safeguards to protect vulnerable groups, such as minors.

Our taxonomy covers seven distinct regulatory requirements for consumer protection on AI:

- **Fair marketing and advertising**: Provisions prohibiting misrepresentations about the capacity, knowledge, or expertise of AI systems. This includes prohibitions on false or misleading advertisements about AI systems (*Example: Section 4999.9(b) of California's Assembly Bill 489 on Health Advice from AI prohibits advertisements or functionality descriptions of AI systems from implying that the services are provided by a licensed healthcare professional*) and requirements to ensure that AI systems do not misrepresent their own capabilities. (*Example: Section 5 of Iowa's SF 2417 Conversational AI Services prohibits operators from causing AI chatbots to represent themselves as designed to provide licensed psychological or behavioural health services*).
- **Human/fundamental rights impact assessment**: Provisions requiring an assessment of the broader societal or fundamental rights impacts of an AI system. *Example: Article 27 of the EU AI Act requires deployers of high-risk AI systems in public interest contexts to conduct a fundamental rights impact assessment before deployment.*
- **Interaction rights**: Provisions giving consumers and users rights with regard to their interaction with AI systems. These include the right to opt in or out of the use of AI (*Example: Article 17 of China's Provisions on the Management of Algorithmic Recommendations in Internet Information Services gives users the right to turn off personalised recommendations*); *the* right to file complaints (*Example: Article 15 of China's Interim Measures for the Management of Generative Artificial Intelligence Services requires providers to establish a convenient complaint mechanism*), and the right to request human review of an AI-assisted decision (*Example: Articles 9-10 of Brazil's draft AI Bill (2338/2023) gives a person significantly impacted by an AI system's decision the right to human intervention or review*).
- **Minor protection**: Provisions demanding the explicit protection of minors, including usage restrictions and parental controls. *Example: Article 14 of China's Interim Measures for Administration of Anthropomorphic AI Interaction Services mandates minor protection measures such as identification of minors, implementation of a minor mode, as well as parental consent measures.*
- **Prohibition of manipulative/deceptive practices**: Provisions prohibiting certain AI systems, functions, and applications based on manipulation of consumer or user behaviour resulting in harm. This includes prohibitions of manipulative AI systems (*Example: Article 5(1)(a)-(b) of the EU AI Act prohibits the deployment of subliminal techniques and exploitation of vulnerabilities with the objective or effect of distorting behaviour)* and prohibitions of specific practices, such as dark patterns, personalised pricing, and manipulating rankings. (*Example: Article 14 of China's Provisions on the Management of Algorithmic Recommendations in*

Internet Information Services prohibits the use of algorithms to manipulate rankings or search results.)

- **Quality of service**: Provisions requiring AI services to be continuous and stable. *Example: Article 13 of China's Interim Measures for the Management of Generative Artificial Intelligence Services requires providers to ensure secure, stable, and continuous services.*
- **User self-protection**: Provisions requiring technical or organisational safeguards to ensure that users of an AI system are protected from engaging in behaviour that is harmful to themselves, including addiction, emotional dependency, and self-harm. *Example: Oregon's AI Companions Act (SB 1546) requires AI service providers to establish a protocol to detect suicidal or self-harm ideation and provide referrals and contact information for suicide and crisis lifelines.*

Delineations

Prohibitions of manipulative/deceptive practices and user self-protection requirements often overlap with other requirements in US state-level chatbot laws. For example, Section 1 of Oregon's AI Companions Act requires operators to take measures to prevent chatbots from misrepresenting their capabilities (overlap with fair marketing and advertising), presenting themselves as humans to minors (overlap with minor protection), and creating romantic interest by the user (overlap with content generation restriction, see information integrity).

Fair marketing and advertising refers to misrepresentations regarding the capabilities of an AI system, including those made by AI systems. Requirements regarding firms' misrepresentations are a special form of manipulative and deceptive practices. Requirements regarding misrepresentations by AI systems are a special form of content generation restrictions (see information integrity). Fair marketing and advertising requirements are closely related to disclosure and information sharing requirements (see transparency): Fair marketing and advertising requirements generally prohibit firms from providing false information while disclosure and information sharing requirements require firms to provide accurate information.

Minor protection requirements frequently involve other requirements, such as a disclosure requirement (see transparency) stating that a chatbot is not suited for minors.

Interaction rights are tied to user requests, to which firms must react. The right to request human review is distinct from a general requirement to ensure human-in-the-loop requirements, which firms have to proactively comply with (see human oversight). Information rights (see transparency) are also "reactive" but focus on the gathering of information as opposed to shaping AI outputs. Brazil's draft AI Bill showcases this difference: Article 8 would establish the right of an affected person to receive an explanation of a decision, while Articles 9-11 would set out the right to contest decisions made by an AI system and receive human review.

Human/fundamental rights impact assessments are distinct from risk and security assessments (see safety) by the scope of the assessment. Both requirements can be tagged on the same provision if a single assessment procedure requires consideration of both safety and the impact on human rights (e.g. in Brazil's draft AI Bill).

Data Governance

Regulatory Requirements

AI models are trained on enormous volumes of data, which in turn shapes AI outputs, raising concerns about how that data is collected, stored, and processed. Governments regulate the composition of training data, the security of data during usage, as well as the rights of individuals with regard to their personal information when interacting with AI systems.

Our taxonomy covers five distinct regulatory requirements for data governance in AI:

- **Data security**: Provisions demanding technical or organisational measures to prevent loss, damage or unauthorised access to data related to AI systems. *Example: Article 14 of China's Regulations on the Management of Deep Synthesis Internet Information Services requires service providers to take technical measures to ensure the security of training data.*
- **Data subject rights**: Provisions establishing individual rights relating to their personal information in connection to AI systems. *Example: Article 16 of China's Interim Measures for Administration of Anthropomorphic AI Interaction Services give users the ability to choose to delete or copy interactive data, such as chat logs.*
- **General data protection**: Provisions establishing rules for personal data collection and processing in the context of AI. *Example: Section 26 of Kenya's draft AI Bill would require providers of high-risk AI systems to comply with existing data protection laws, including by conducting data protection impact assessments where required*.
- **Prohibition of certain data use**: Provisions prohibiting the use of certain data to train AI systems (*Example: Article 16 of China's Interim Measures for Administration of Anthropomorphic AI Interaction Services prohibit the use of interactive data containing sensitive personal information for model training*) and certain AI data gathering uses (*Example: Article 6 of Chile's draft Law regulating AI systems would consider AI systems that create biometric classifications based on sensitive personal data and public real-time remote biometric identification systems to pose unacceptable risk*).
- **Training data quality**: Provisions regarding the quality and composition of datasets used for training AI models. *Example: Article 10 of the EU AI Act requires training data to be relevant, sufficiently representative, error-free, and complete.*

Delineations

Data governance requirements can overlap with each other. Training data quality focuses on requirements relating solely to the input for AI model training, while general data protection requirements cover all other aspects of how data is processed, stored, and otherwise used during an AI system's lifecycle. Data subject rights are a specific form of general data protection and accordingly prevail whenever a provision involves individual rights.

Data governance requirements can also overlap with safety requirements (see safety). When an AI system is prohibited because of the data it was trained on, we categorise it as a prohibition of certain data use, rather than a general prohibition. For a provision focused on the security of data used by AI systems, we use data security. The system safety requirement covers cybersecurity measures beyond the data component (including to prevent unauthorised access to AI systems).

Human Oversight

Regulatory Requirements

As AI systems become increasingly capable, a core concern for governments and firms is ensuring that humans retain control over AI decisions, particularly in high-risk contexts. The design of effective human oversight mechanisms is imperative, but complicated by the speed, scale, and opacity at which cutting-edge AI systems operate.

Our taxonomy includes four distinct regulatory requirements regarding human oversight of AI:

- **Human intervention and suspension**: Provisions requiring AI systems to be designed for authorised humans to intervene in systems and override individual outputs or suspend systems entirely. *Example: Article 14(4)(d) and (e) of the EU AI Act requires the capability to suspend AI systems.*
- **Human-in-the-loop**: Provisions demanding that human beings supervise AI systems and actively participate in AI-assisted decisions. This includes requirements to design AI systems in a fashion that enables effective human oversight (*Example: Article 14(1) of the EU AI Act requires high-risk AI systems to be designed and developed in a way that enables natural persons to effectively oversee them*) and provisions determining the required timing and extent of human participation (*Example: Singapore's Model AI Governance Framework for Agentic AI recommends a human oversight checkpoint before high-risk or irreversible decisions.)*
- **Organisational governance**: Provisions mandating organisations to implement operational structures, roles, policies, and procedures to ensure adequate human oversight. This includes requirements to establish specific internal oversight mechanisms and governance structures. *Example: Article 28 of the Republic of Korea's Framework Act on AI allows AI business operators to establish private autonomous AI ethics committees to check compliance with ethical principles*.
- **Training and staffing**: Provisions requiring AI firms to ensure appropriate levels of qualifications and AI literacy of staff within the organisation. *Example: Article 11 of Brazil's draft Bill on AI development, implementation, and application (1465/2024) would require providers and operators of high-risk AI systems to implement continuous training for all personnel involved in their supervision, operation, or use.*

Delineations

Human oversight requirements can overlap with each other, but are distinct. Human-in-the-loop concerns human oversight during the regular functioning of an AI system, while intervention and suspension requirements focus on the active intervention capacity of humans to intervene in and override a specific decision or bring a process to a halt altogether. Both requirements can be established in the same or adjacent provisions, as in Article 14 of the EU AI Act.

Discrimination and Bias

Regulatory Requirements

Governments worldwide are concerned with the risk that AI systems reproduce or amplify existing bias and lead to discrimination. This risk is particularly acute where AI informs high-stakes decisions in areas such as hiring, credit, healthcare, or criminal justice.

Our taxonomy covers three distinct regulatory requirements regarding discrimination and bias in AI:

- **Bias detection**: Provisions demanding active testing for and detection of possible biases in AI training datasets and AI system outputs. *Example: New York City Local Law 144 on Automated Employment Decision Tools requires employers to conduct bias audits before using automated employment decision tools.*
- **Bias mitigation**: Provisions requiring firms to prevent or mitigate AI bias. *Example: Canada's AI and Data Act would require persons responsible for high-impact AI systems to establish measures to mitigate risks of biased output resulting from the system's use.*
- **Non-discrimination**: Requirements for AI systems to be fair and non-discriminatory, as well as prohibitions of discriminatory AI systems. This includes obligations for AI providers to ensure that AI systems are non-discriminatory and fair. *Example: Article 7 of Viet Nam's AI Law prohibits the use of AI to exploit the weaknesses of vulnerable groups.*

Delineations

Bias detection and bias mitigation requirements often overlap. For example, Article 10(2)(g) of the EU AI Act requires both the adoption of appropriate measures to detect possible biases as well as the subsequent mitigation of those biases.

Information Integrity

Regulatory Requirements

AI can shape public discourse, including through AI-driven feeds and generative AI systems that produce and amplify information at scale. Governments have responded by walking a thin line between setting guardrails around AI-generated content, including the moderation of certain content, and maintaining freedom of expression, by protecting users' lawful speech.

Our taxonomy includes four regulatory requirements regarding AI and information integrity:

- **Content generation restrictions**: Provisions restricting the generation of certain content using AI systems and the use of certain AI-generated content. This includes prohibitions on using AI systems to generate specific content (*Example: Article 4(j) of China's Interim Measures for the Management of Generative Artificial Intelligence Services prohibit, among other things, the generation of content which endangers national security, or content which promotes terrorism*), as well as outright prohibitions of AI systems capable of generating or modifying certain types of content (*Example: Article 9 of Mexico's draft AI Law would prohibit AI systems that are intended to be used to alter the image or voice content of a person without consent from the affected person or owner of the content*).
- **Content moderation**: Provisions requiring AI firms and platforms to moderate prohibited types of AI-generated content. *Example: Article 14 of China's Interim Measures for the Management of Generative Artificial Intelligence Services obliges providers of generative AI services to cease transmitting and delete illegal content.*
- **Content reporting**: Provisions empowering users to report illegal or harmful AI-generated content. *Example: Section 3 of the TAKE IT DOWN Act requires platforms to establish reporting channels to submit notices of non-consensual AI-generated deepfakes, among others.*
- **User speech protection**: Provisions protecting the lawful speech by users from removal and requiring justification thereof. *Example: We have not yet identified binding regulation on enshrining user speech rights with regard to AI. However, in 2025, the Judiciary Committee of the US House of Representatives sent letters to a number of AI companies, including Alphabet, Anthropic, OpenAI and Meta, requesting information regarding their interaction with the previous Federal Administration with respect to content moderation of AI models.*

Delineations

Content generation restrictions can overlap with copyright requirements. If a restriction concerns the violation of personal rights or consent requirements, it is a content generation restriction, whereas a provision restricting certain content generation to avoid copyright violations is considered a copyright requirement. Provisions addressing both are tagged with both requirements.

User speech protection can overlap with interaction rights and human review (consumer protection) insofar as they allow a user to challenge decisions made with regard to them. When such a provision specifically focuses on user-generated content, it is categorised as user speech protection.

Labour

Regulatory Requirements

AI is set to reshape labour markets worldwide, with governments concerned about both the transformation and especially the displacement of roles in their workforce. At the same time, AI raises concerns about worker surveillance and unwanted forms of AI use in management.

Our taxonomy includes two regulatory requirements regarding AI and labour:

- **Worker monitoring/management**: Provisions restricting the use of AI to surveil, score, rank or manage workers. *Example: Article 5(1)(f) of the EU AI Act prohibits the use of AI to infer emotions of workers for non-medical or safety reasons*.
- **Worker protection**: Provisions aiming to protect workers against AI-related harm in the context of employment, such as protection against certain AI-assisted hiring and termination practices, information obligations regarding the use of AI at the workplace, and rights with regard to AI usage in employment decisions. *Example: Section 33 of Kenya's draft AI Bill would require providers or deployers of AI systems likely to impact employment to conduct a workforce impact assessment and implement mitigation measures.*

Delineations

Provisions creating private-sector obligations with regard to unemployment are, as of yet, relatively uncommon in AI-specific legal frameworks. Notable exceptions include Kenya's recent draft AI Bill and the Philippines' draft Protection of Labor Against AI Automation Act, which would require job displacement mitigation and obligate employers who use AI resulting in permissible job displacements to provide transition assistance to affected human workers.

To ensure extensive coverage of this novel and growing concern, we always tag labour-related provisions as one of the two labour requirements, while adding additional tags using regulatory requirements from other concerns, if applicable. A future extension of our labour taxonomy is probable.

Intellectual Property

Regulatory Requirements

How intellectual property rights evolve in the age of AI is essential for the business model of AI firms, for the creative sector, and for AI users. Courts and legislators in many jurisdictions are actively discussing whether existing intellectual property frameworks are adequate, or whether new rules are needed to address the distinctive features of AI-generated and AI-assisted creation. This includes concerns regarding the training of AI models using copyrighted material, the appropriate compensation of creators of AI training material, and ownership rights of AI-generated outputs.

Our taxonomy includes three regulatory requirements regarding AI and intellectual property:

- **Copyright**: Provisions governing the relation between copyright protection and AI systems. This includes conditions for the use of copyrighted works in the training of AI systems (*Example: Article 21 of Mexico's draft AI Law would require prior agreement with intellectual property rightholders for using their content to train AI models*) and rules determining copyright ownership of user prompts and AI-generated content (*Example: Arkansas' Act on ownership of model training and generated content holds that the person supplying input to a generative AI tool owns the generated content provided that no copyright or intellectual property infringement occurs*).
- **Patent**: Provisions regulating the status of patent inventor when AI systems are used to create new inventions and mentioned in patent applications. *Example: We have not yet identified binding regulation regarding inventorship and AI systems. However, the US Patent and Trademark Office has released non-binding guidance on the matter (Revised Inventorship Guidance for AI-Assisted Inventions, adopted 26 November 2025), stating in its section IV, that AI systems cannot be named as inventors or joint inventors in US patent applications.*
- **Remuneration**: Provisions demanding the compensation for copyright owners when their protected works are used for AI training or appear in AI output. *Example: We have not yet identified binding regulation regarding the obligation to compensate copyright-holders for training use. However, the European Parliament has adopted recommendations which would ensure fair remuneration for copyright holders whose content is used in AI training (European Parliament Recommendations on copyright and generative AI, adopted by Committee on 28 January 2026).*

Delineations

Provisions regulating intellectual property and AI in AI-specific legal frameworks are, as of yet, relatively uncommon. This may be because governments are waiting for their judicial branches to rule on the application of existing intellectual property provisions to AI, or amending existing intellectual property-specific rules rather than new AI-specific frameworks.

Copyright requirements can overlap with each other and with requirements from other concerns. Remuneration is downstream from copyright, meaning overlaps are possible, but we tag remuneration when a provision focuses only on compensation. Provisions regulating training data from an intellectual property perspective fall under copyright, whereas data governance requirements apply when training data is regulated from a data protection perspective.

Accountability

Our concept of accountability is distinct from all other concerns: Accountability does not encompass distinct regulatory requirements, but rather sets out the responsibility for compliance, the consequences of non-compliance, and the oversight mechanisms to ensure compliance. These three elements of accountability are relevant for every regulatory requirement in the taxonomy, meaning accountability cuts across all other concerns.

The responsibility for compliance is a core element of accountability, since it describes who governments target with regulatory requirements. Targets are typically determined based on AI systems' risk level, the AI systems' technological specificities, and AI firms' role in the AI value chain.

- Risk-based frameworks tend to classify AI systems into risk tiers and apply different regulatory requirements to each tier. For example, Viet Nam's AI Law and its implementing decree set out criteria for low-, medium- and high-risk AI systems and establish different requirements for each tier (mostly for high-risk AI systems).
- Technology-based frameworks establish regulatory requirements based on different types of AI systems and capabilities. China's frameworks, for instance, each apply to one type of AI system, namely algorithmic recommendation technology, deep synthesis technology, generative AI technology, and anthropomorphic AI services. In California, some rules apply based on a training compute threshold, typically 10^{26} floating point operations (FLOPs).[3]
- Frameworks can also determine which firms must comply with regulatory requirements based on their position in the AI value chain. For example, the EU AI Act establishes requirements for AI providers (Article 16), importers (Article 23), distributors (Article 24), and deployers (Article 26).

As AI's risk factors, technical capabilities, and value chain evolve continuously, we do not yet tag the responsibility for compliance systematically. We plan to cover targets in future iterations, including liability for AI from a practical litigation standpoint, such as a private right of action (e.g. in Oregon) and a narrowing of legal defences (e.g. in California) in relation to AI.

The consequences of non-compliance with a regulatory requirement are a strong indicator of how a government creates accountability. Governments are currently experimenting with a spectrum of provision types, from high-level principles, to non-binding or voluntary commitments, to binding requirements that carry a sanction threat. We currently focus on the latter, tagging provisions that determine sanctions (including penalties, administrative sanctions and fines), such as Articles 99-101 of the EU AI Act. We plan to compare sanctions regimes in detail in future iterations.

Oversight mechanisms determine how governments create accountability in practice: Which existing or new government authorities are in charge of overseeing AI firms and what powers do these authorities have to enforce regulatory requirements? We currently tag provisions that cover both elements, oversight structures and powers, such as Section 14 of India's draft Artificial Intelligence (Human Health and Medical Education) Regulation Bill, 2026. We plan to compare these structures and powers and leverage the DPA's enforcement tracking to provide more insight into oversight mechanisms in practice in future iterations.

[3] Elements can be combined: The EU AI Act sets 10^{25} FLOPs as systemic risk threshold for general-purpose AI models.

Scope

This section outlines the scope and depth of our tagging at the domestic and international level.

Domestic Frameworks

Our tagging of domestic frameworks occurs on two levels: We tag concerns into the 2000+ policy and enforcement developments targeting AI on the DPA database and tag regulatory requirements, at the provision level, into the text of 84 domestic AI laws, regulations, and proposals.

How we tag domestic frameworks

To tag concerns into the DPA database, we create a bridge between the DPA's taxonomy of digital policy instruments and the concerns (see DPA Instrument - Concern Matching). For most DPA policy instruments, there is a 1-1 match to a concern: A data protection development relates to data governance, for example. We automatise such matches. Some DPA instruments, however, can address multiple concerns. For these cases, we have developed an automated workflow to identify which concerns are mentioned in the DPA description, which we doublecheck manually. Finally, some DPA instruments are not related to AI, meaning we do not tag concerns. To date, we tag DPA entries that target the AI sector, but we plan to expand this matching to the entire DPA database.

To tag regulatory requirements into the provisions of domestic AI frameworks, we again use our Clairk tool. Our team manually tagged regulatory requirements into the Clairk library of domestic AI frameworks, which spans 153 documents.

Scope of domestic frameworks

We applied our taxonomy of regulatory requirements to 84 laws, regulations, and proposals, across 19 jurisdictions (see DPA Tagged Domestic AI Frameworks). In total, we have tagged 454 provisions and added 935 regulatory requirement tags (since one provision can include multiple regulatory requirements).

Rather than tagging the entire AI library, we focus on frameworks that include requirements for the private sector. This includes:

- adopted AI laws, including the EU, Viet Nam, and the Republic of Korea, and US states (including California, Colorado, Oregon);
- adopted AI regulations, including China's AI regulations and relevant US Executive Orders, as well as the implementing regulations of AI laws;
- select draft AI laws currently under deliberation in parliaments, e.g. in Brazil and Kenya; and
- select historic draft AI laws, e.g. Canada's AI and Data Act.

The project excludes domestic laws focused on developing (rather than regulating) the AI ecosystem or on how government entities are to use AI. Non-binding documents and guidelines are also not covered to date, with some few exceptions (e.g. Singapore's Model AI Governance Framework).

International Frameworks

Our tagging of international frameworks covers nine texts and occurs at the provision level: We analyse the text of these nine frameworks, article by article, to identify which concerns they address. We additionally show whether the language of the provision is focused on a concern or only touches upon it tangentially, and whether the provision contains actionable language or does not give actionable guidance for governments and firms aiming to implement it. The full set of tagged international provisions is accessible here: International framework concern mapping

How we tag international frameworks

The goal of tagging the text of international frameworks is to reveal alignment between governments on AI concerns. Different international frameworks, with accordingly different participating governments, addressing the same concerns provide fertile ground for international cooperation at a broader level. Building bridges across borders begins by identifying commonalities.

There are inevitable challenges in applying the same taxonomy to nine distinct international frameworks: The frameworks differ in terms of their scope, bindingness, and addressees.[4] For instance, their language ranges from abstract principles, to concrete policy recommendations for governments, to suggestions for the private sector.

Our solution is to tag provisions in two steps: First, whether a provision mentions a concern, and second, how the provision covers that concern. Our threshold for determining whether a concern is mentioned is relatively low. Frameworks interpret concerns differently, with language that can stand in contrast to our own taxonomy and that often does not separate between concerns as systematically.[5] Our heuristic is to rely on our own taxonomy and tag a provision whenever it mentions what we consider to fall under a concern, as outlined above.

We then analyse whether the provision focuses on the concern and whether the provision contains actionable language for governments and private sector actors to address the concern. For "focus", we distinguish between "core" provisions, in which the concern is the provision's principal subject (be it in the title or substance) and "tangential" provisions, that only reference the concern in passing or clearly subordinate the concern to another, core, concern.[6] For "actionability", we distinguish between "actionable" provisions that contain a specific action to address the concern (e.g. a practical measure or oversight mechanism) and "non-actionable" provisions that only define, discuss, endorse, promote, or reference a concern, without formulating any concrete action that actors can undertake to address the concern.

Finally, since we consider domestic implementation to be the crux of international AI coordination, we note when international frameworks mention the regulatory requirements of our AI taxonomy. At the international level, this reveals whether international frameworks connect to the reality of AI regulation "on the ground" and whether they do so in similar fashion, or suggest different regulatory

[4] For example, the G7 Hiroshima Process Guiding Principles and Code of Conduct address organisations directly, while the Council of Europe's Convention addresses signatory states.
[5] For example, the ASEAN Guide on AI Governance establishes guiding principles that usually relate to a single concern, but then provides concrete examples of applications and measures that refer to another concern/requirement (e.g. the transparency principle provides examples of application to labour context and discrimination concerns).
[6] If a provision covers multiple concerns with similar prominence and without hierarchy, we consider each concern "core".

requirements to address the same concerns. Crucially, this creates a bridge between the international and domestic level: It reveals how governments implement international AI frameworks and whether they do so similarly or create granular differences across borders, despite international alignment on shared concerns and regulatory requirements to address them. Our newest report analyses this intersection between the international and domestic layer, while previous reports by our team highlight differences even when governments apply the same regulatory requirement.

To tag concerns into the provisions of international AI frameworks, we use our Clairk tool. This full-text analysis tool allows users to interact with policy text, including searching across texts, comparing texts side-by-side, and tagging provisions. Our team manually tagged the Clairk library of international AI frameworks, which spans 24 documents.

Scope of international frameworks

The Digital Policy Alert tracks over 200 international developments focused on AI, but our team tags nine international frameworks on AI in detail. Our focus lies on frameworks, declarations and treaties that contain language guiding countries in their regulation of AI or establishing principles for AI that could be codified in regulation (or directly by the private sector). To ensure a global perspective, we include frameworks by UN bodies as well as regional bodies.

Other international documents that contain important language, including documents issued following AI Summits in Bletchley, Seoul, Paris, and Delhi, are not covered by our tagging. This can be because their language either does not span across concerns or does not include specific language regarding regulation. We plan to expand the scope of our tagging in future iterations.

This section sets out the international frameworks we selected, noting the official source, participating jurisdictions, relevant dates, and structure of each framework, namely:

- OECD AI Principles
- UN General Assembly Resolution A/RES/78/265 - Seizing the opportunities of safe, secure and trustworthy AI systems for sustainable development
- Council of Europe Framework Convention on AI and Human Rights, Democracy, and the Rule of Law
- G7: Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems
- G7: Hiroshima Process International Code of Conduct for Advanced AI Systems
- ASEAN Guide on AI Governance and Ethics
- ASEAN Expanded Guide on AI Governance and Ethics – Generative AI
- UNESCO: Recommendation on the Ethics of AI
- Joint Declaration on AI, Freedom of Expression and Media Freedom

OECD AI Principles

Links: Original Source / Clairk

Jurisdictions: OECD (47 adherents, 8 non-OECD)

Date: Initial version adopted on 22 May 2019, updated on 3 May 2024.

Structure: The OECD AI Principles consist of ten principles across two sections, with each principle consisting of either a single or several paragraphs. We focus on Section 1 ("Values-Based Principles"), which addresses AI actors and stakeholders, while we exclude Section 2, which addresses investment, capacity building and interoperability.

UN General Assembly Resolution A/RES/78/265 - Seizing the opportunities of safe, secure and trustworthy AI systems for sustainable development

Links: Original Source / Clairk

Jurisdictions: UN Member States

Date: adopted on 21 March 2024

Structure: The General Assembly Resolution consists of 13 paragraphs addressing member states, private sector, and agencies. We focus on paragraph 6, which calls for member states to adopt regulatory and governance approaches that conform with its 17 sub-paragraphs outlining various governance measures.

Council of Europe Framework Convention on AI and Human Rights, Democracy, and the Rule of Law

Links: Original Source / Clairk

Jurisdictions: 21 signatories, counting the EU as a single signatory, and including 5 non-Council of Europe members: Canada, Israel, Japan, US, Uruguay. The treaty has been ratified by Bosnia and Herzegovina and the EU.

Date: Opened for signature on 5 September 2024. Enters into force once 5 countries have ratified it, including at least 3 Council of Europe states.

Structure: The Convention is currently the only binding treaty requiring parties to adopt regulatory measures on AI, including provisions on follow-up mechanisms and international cooperation. According to Article 3, the Convention applies to activities in the AI lifecycle undertaken by public authorities, or by private actors on their behalf. Parties are further required to address risks and impact from private sector AI activities "in a manner conforming with the object and purpose of this Convention". Parties must thus submit a declaration on how they intend to implement this private-sector-specific obligation, either by applying the principles and obligations of the Convention, or by other appropriate means. The chapters that are relevant to the AI taxonomy are Chapter 2 (general obligations to protect human rights), Chapter 3 (principles related to AI lifecycle), Chapter 4 (remedies), Chapter 5 (risk assessment/mitigation), and Chapter 6 (implementation, including ensuring non-discrimination and protection of vulnerable groups).

G7: Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems

Links: Guiding Principles (Original Source / Clairk)

Jurisdictions: G7, with 70 “friends” as of August 2026

Date: Published on 30 October 2023

Structure: The Guiding Principles are part of the Hiroshima Process, a G7 initiative aimed at promoting a comprehensive policy framework for AI building on the OECD AI Principles. The document sets out a non-exhaustive list of 11 guiding principles addressed to organisations developing and using AI systems in the private sector, public sector, and in civil society. Each guiding principle is accompanied by explanations and, in some cases, recommended actions. The Guiding Principles share the same basic structure as the Code of Conduct (below), which adds additional detail in several areas.

G7: Hiroshima Process International Code of Conduct for Advanced AI Systems

Links: Code of Conduct (Original Source / Clairk)

Jurisdictions: G7, with 70 “friends” as of August 2026

Date: Published on 30 October 2023

Structure: The Code of Conduct is part of the Hiroshima Process, a G7 initiative aimed at promoting a comprehensive policy framework for AI building on the OECD AI Principles. The Code of Conduct contains the same structure and text as the Guiding Principles, while adding additional detail in the form of concrete recommended actions across the board.

ASEAN Guide on AI Governance and Ethics

Links: Original Source / Clairk

Jurisdictions: 11 ASEAN Member States

Date: Published on 2 February 2024

Structure: The ASEAN Guide addresses organisations developing and deploying AI and further formulates recommendations for policymakers at the national and regional level. The parts that are relevant to the taxonomy are: Section B’s seven guiding principles, which include examples of potential concerns and actions; and Section C’s AI governance frameworks, which address organisations on four levels: internal governance, determining human involvement, operations management across the AI lifecycle, and stakeholder interaction. The national and regional recommendations focus on supporting the AI ecosystem and innovation, which is not in scope.

ASEAN Expanded Guide on AI Governance and Ethics – Generative AI

Links: Original Source / Clairk

Jurisdictions: 11 ASEAN Member States

Date: Published on 17 January 2025

Structure: The Expanded Guide addresses ASEAN policymakers, building on the AI Guide to address generative AI-specific risks. The sections that are relevant to the taxonomy are the modified seven guiding principles and the regional policy recommendations.

UNESCO: Recommendation on the Ethics of AI

Links: Original Source / Clairk

Jurisdictions: 193 UNESCO Member States

Date: Unanimously adopted on 23 November 2021

Structure: The Recommendation addresses member states as the responsible authorities for developing legal frameworks on AI, while also providing an ethical impact assessment framework to AI actors in all sectors. The Recommendations set out four values (basic motivations for desirable behaviour) and ten principles (concrete formulations of desirable actions and objectives). The Recommendation also provides eleven areas of policy action that include more specific pathways for operationalising the values and principles.

Joint Declaration on AI, Freedom of Expression and Media Freedom

Links: Original Source / Clairk

Jurisdictions: Jointly issued by the UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression, and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information in Africa.

Date: Adopted on 24 October 2025

Structure: The Joint Declaration outlines four AI-related challenges with regard to freedom of expression and seven general principles addressed to states and stakeholders, as well as a series of recommendations to specific stakeholders. The most relevant sections for this project are the recommendations to states, state regulation of companies, the private sector, and journalists.

Digital Policy Alert

Contact
Written by Tommaso Giardini and Nils Deeg
Please send questions and suggestions to nils.deeg@sgept.org

The independent Digital Policy Alert is a pillar of the Swiss-based St. Gallen Endowment for Prosperity Through Trade

St.Gallen Endowment
for Prosperity through Trade