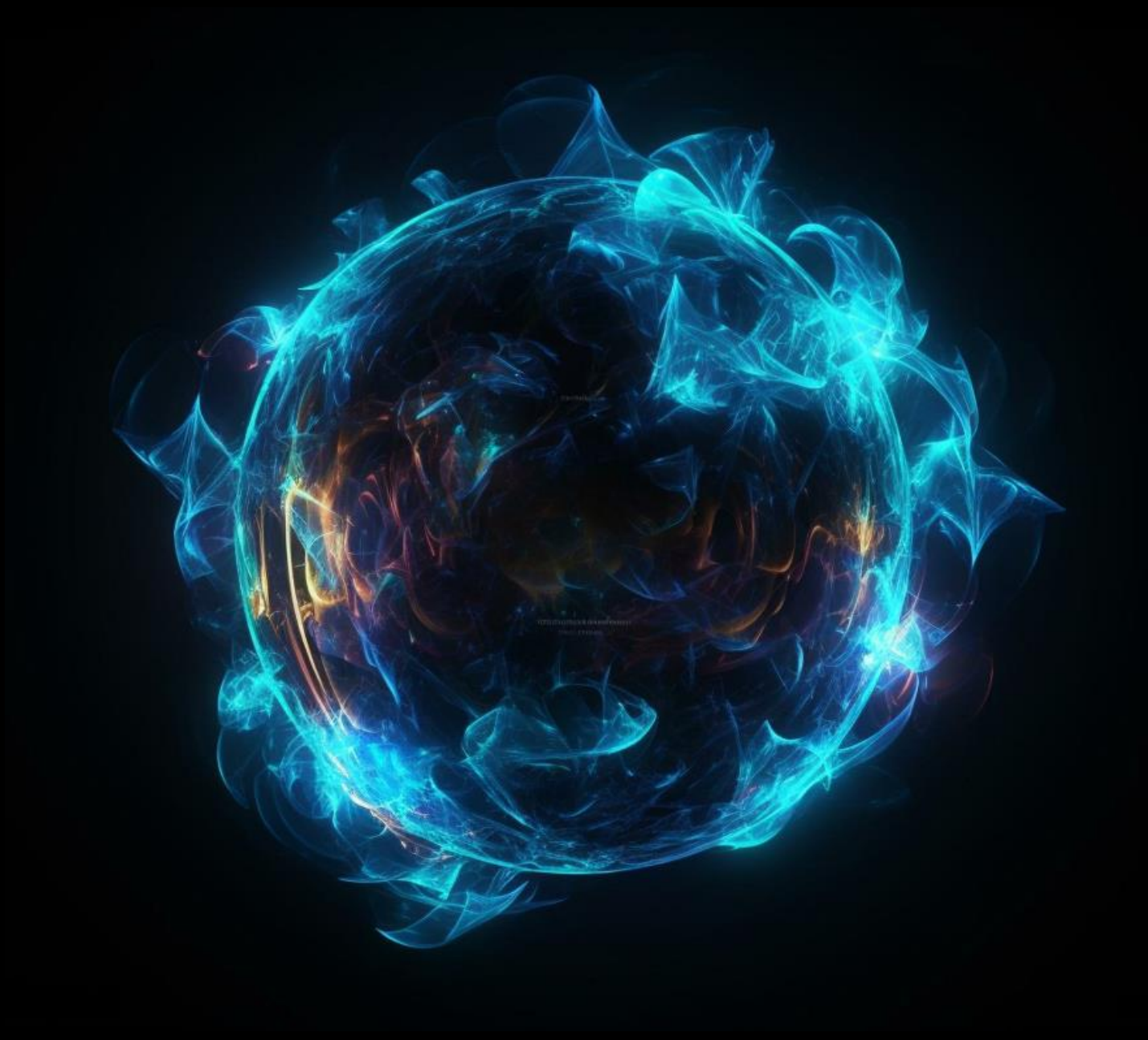
Deloitte.
Legal



Latest global news on Artificial Intelligence

March 2024

MAKING AN IMPACT THAT MATTERS since 1845

Index

The countdown is on: the AI Act will soon be law

On February 2nd, 2024 the Council of the **EU's Committee of Permanent Representatives (COREPER)** endorsed the compromise text of the Artificial Intelligence Act ("**AI ACT**" or "**Regulation**"), as agreed upon by the Council presidency and the negotiators from the European Parliament on December 8th 2023. After the approval of the compromise text by the European Parliament's **Internal Market and Civil Liberties Committees**, on March 13th, 2024 **MEPs endorsed the AI Act** in plenary session with a large majority. The Regulation is still subject to a final lawyer-linguist check and is expected to be finally adopted before the end of the legislature. The law also needs to be formally endorsed by the Council.

Once approved, the AI Act will become applicable 2 years after its entry into force, with some exceptions: for example, prohibitions on prohibited practices will already apply after 6 months; obligations on general-purpose AI models will apply after 12 months; and obligations for high-risk systems after 36 months.

This document explores the key innovations provided for in the AI Act, a pioneering framework for artificial intelligence.

Definition of AI system

The AI Act provides the following definition.

> *"AI system is a machine-based system 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".*

This definition highlights AI systems' key features, distinguishing them from simpler traditional software systems or programming approaches. The definition of AI system provided by the AI Act is closely aligned with the work of international organizations, among which the **OECD**. Indeed, the Regulation aims at facilitating international convergence and acceptance of what can be defined as AI, as well as legal certainty and capability to adapt to the rapid technological developments in the field of AI.

Risk-based approach

The AI Act's primary objective is to govern AI technologies based on their potential to pose risks to society, using a **risk-based approach**: the greater the risk, the more stringent the rules.

The AI Act establishes obligations based on AI potential risks and level of impact:

- **Unacceptable risk**: **prohibited systems**, among which:
  - **behavioral manipulation**;
  - **exploitation of personal vulnerabilities** (due to age, disability, social or economic situation);
  - **biometric categorization** systems based on biometric data;
  - **social scoring** based on social behavior, or personal or personality characteristics;
  - **real-time remote biometric identification** in publicly accessible spaces for law enforcement purposes (with narrow exemptions);
  - certain applications of **predictive policing**;
  - **untargeted scraping of facial images** from the internet or CCTV footage to create or expand facial recognition databases;
  - **emotion recognition** systems **in the workplace and education institutions** (with exemptions for safety and medical reasons).

- **High risk**: systems that must comply with **strict requirements** and pertain to the areas of:
  - **biometrics** (among which remote identification and categorization as well as emotion recognition);
  - **critical infrastructure**;
  - **education and vocational training** (determining access to training, evaluating learning outcomes, assessing appropriate education level to receive, and test monitoring);
  - **employment, workers management, and access to self-employment** (recruitment and decisions on work terms);
  - **essential private and public services and benefits** (including creditworthiness evaluation);
  - **law enforcement**, if permitted under union or national law;
  - **migration, asylum, and border control management**, if permitted under union or national law;
  - **administration of justice and democratic processes**.

- **Limited risk**: systems subject to light **transparency obligations** (e.g., chatbots), encouraged to adhere to codes of conduct.

- **Minimal risk**: systems with minimal or no risk, **exempt from obligations** (e.g., recommender systems or spam filters), encouraged to adhere to codes of conduct.

Main requirements for high-risk AI systems

The AI Act establishes **strict requirements** for high-risk AI systems:

- **Use of risk management systems**, consisting of a continuous process running through the entire lifecycle of the AI system and with a systematic review and update.
- **Development based on training, validation and testing data sets meeting quality criteria**, including:
  - Data governance and management practices.
  - Relevant, sufficiently representative, free of errors, and complete datasets.
  - Consideration of geographical, contextual, behavioral and functional settings.
  - Additional conditions applicable for special categories of personal data.
- **Technical documentation** to be drafted before the system is placed on the market or put into service and kept up-to-date.
- **Record-keeping** by automatically recording specific events ("logs") over the system's whole lifetime and ensuring an appropriate level of traceability of the system's functioning.
- **Transparency and provision of information to deployers**.
- **Design and development enabling human oversight** that should be commensurate to the risk, level of autonomy and context of the use of the AI system.
- **Appropriate levels of accuracy, robustness and cybersecurity**, throughout the whole lifecycle.

The AI Act also states that deployers of high-risk AI systems that are bodies governed by public law, or are private entities providing public services and deployers of high-risk systems used to evaluate creditworthiness or used for risk assessment and pricing in relation to life and health insurance shall perform a **fundamental rights impact assessments.** This assessment shall complement the data protection impact assessment and its results must be notified to the market surveillance authority.

Moreover, providers shall **register themselves and their systems** in the EU database for high-risk AI systems and carry out a **conformity assessment procedure** to demonstrate compliance with the requirements above. High-risk AI systems should bear the **CE marking** to indicate their conformity with this Regulation. For high-risk AI systems provided digitally, a **digital CE marking of conformity** shall be used.

Transparency obligations of certain AI systems

The AI Act provides a series of **transparency obligations** for providers and deployers (or "users") of certain AI systems and **general-purpose AI models** ("**GPAI**"), that must be provided to natural persons when they first interact or are exposed to such systems.

In particular, the AI Act lays down that:

- Providers of AI systems intended for **direct interaction with people** must ensure that the system is designed and developed to **inform persons they are engaging with an AI system** (unless this is obvious for a reasonably well-informed, observant and circumspect person and considering the circumstances of the case).
- Providers of AI systems (including GPAI) that **generate synthetic audio, image, video or text content** must ensure that they are **labeled in a machine-readable format** and are **identifiable as generated or manipulated by AI systems**.
- Deployers of **emotion recognition or biometric categorization systems** must **inform natural persons when they are exposed to such systems** and process personal data in accordance with **Regulation (EU) 2016/679 (GDPR)**, among others.
- Deployers of AI systems that **generate or manipulate image, audio or video content constituting a deep fake** must **disclose the content has been artificially generated or manipulated**.

  If the content is part of a clearly artistic, creative, satirical and fictional analogous work or program, providers only need to disclose the artificial nature of such content in a manner appropriate not to hamper its display or enjoyment.

  Deployers of AI systems that **generate or manipulate text intended to inform the public about matters of public interest** are subject to the same disclosure obligation.

General purpose AI models

The AI Act provides the following definition of general-purpose AI model:

> *"AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications".*

The AI Act lays down some **obligations for providers of GPAI**. In particular, it imposes to:

- ❑ Draft and keep up-to-date the GPAI's **technical documentation**.
- ❑ Draw up, keep up-to-date and **make available information and documentation to providers** of AI systems who intend to integrate the GPAI model into their AI systems.
- ❑ Adopt a policy to ensure **compliance with copyright law.**
- ❑ Draft and make publicly available a **detailed summary** of the content used for training.
- ❑ Appoint **authorized representatives** if providers are established outside the European Union market.

Additional requirements are provided for **GPAI models with systemic risks**, which are:

- GPAI having **high-impact capabilities**, namely capabilities that match or exceed the capabilities of the most advanced GPAI;
- **Based on a decision** of the Commission or following a qualified alert by the scientific panel that a general-purpose AI model has capabilities or impact equivalent to those of the point above.

Providers of GPAI with systemic risks are required to comply with the **obligations for GPAI** and, additionally, to;

- ❑ perform **model evaluations**;
- ❑ **assess and mitigate risks** at the EU level;
- ❑ keep track of, document and report to the AI Office and national competent authorities, relevant **information about serious incidents** and **possible corrective measures**;
- ❑ ensure an **adequate level of cybersecurity protection** of the model and its physical infrastructure;
- ❑ **adhere to a code of practice** or **a harmonized standard** for presumption of conformity.

Governance & enforcement

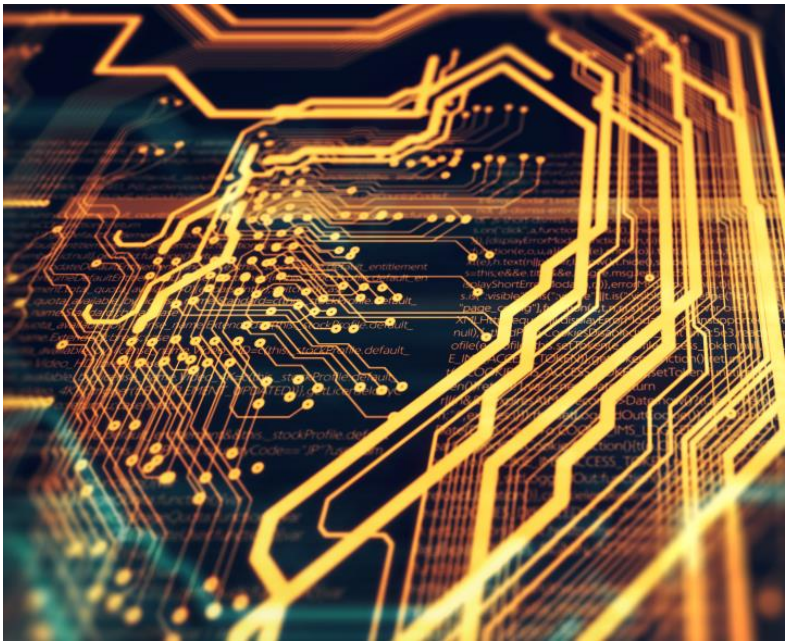
The AI Act also includes new rules on governance that provide a more **centralized system for oversight and enforcement**. In particular, the governance framework is made of:

- The **AI Office**, a new body established within the Commission, whose mission is to develop Union expertise and capabilities in the field of AI and to contribute to the implementation of Union law on AI.
- The **Scientific Panel of independent experts**, an advisory body with expert members selected by the Commission that provides technical advice and input to the AI Office.
- The **AI Board**, an advisory body to the Commission and Member States, that comprises member states' representatives, providing advice on the implementation of the Regulation, in particular as regards the rules on GPAI models.
- The **Advisory forum for stakeholders**, which provides stakeholders' input to the Commission (and the AI Office) and the AI Board.
- At least one **notifying authority** and at least one **market surveillance authority** designated or established by each Member State as **national competent authorities.**

Penalties

The AI Act has established **penalties for various types of infringements**. Indeed, the AI Act lays down administrative fines:

- **Up to € 35 million** or, if an undertaking, up to **7% of the total worldwide annual turnover** (whichever is higher) for non-compliance with prohibited AI practices.
- **Up to € 15 million** or, if an undertaking, up to **3% of the total worldwide annual turnover** (whichever is higher) for non-compliance with other obligations related to operators or notified bodies.
- **Up to € 7,5 million** or, if an undertaking, up to **1% of the total worldwide annual turnover** (whichever is higher) for supplying incorrect, incomplete, or misleading information to notified bodies and national competent authorities.



The AI Office

The **AI Office**, a new governance body established by the AI Act within the Directorate-General for Communication Networks, Content and Technology of the Commission, will be the **centre of AI expertise** across the EU.

The AI Office will play a pivotal role in **implementing the AI Act** and enforcing **general-purpose AI rules**. Indeed, the AI Office will:

- **Contribute to the coherent application** of the AI Act (including the set-up of advisory bodies at EU level).
- Develop tools and standards to **assess GPAI scope and capabilities** and **classify models with systemic risks.**
- **Draft state-of-the-art codes of practice** in cooperation with AI developers and other experts.
- **Investigate potential infringements of rules** and request corrective actions.
- **Draw up guidance and guidelines** and monitor compliance.

The AI Office will also **foster international cooperation** by promoting the **EU's approach to trustworthy AI**, fostering **international cooperation and governance on AI**, and supporting the creation of **international agreements on AI**.

The AI Office will also **strengthen the development and use of trustworthy AI** across the internal market by:

- **Promoting actions and policies** to harness the benefits of AI.
- **Offering guidance on best practices** and facilitating ready-access to AI sandboxes, real-world testing, and other support structures.
- **Fostering dynamic ecosystems of trustworthy AI.**
- Assisting the EU Commission in **leveraging the use of AI tools** and **reinforcing AI literacy**.

The AI Office will **collaborate with different institutions, experts, and stakeholders**. At the institutional level, the AI Office will cooperate with the **European Artificial Intelligence Board** and the **European Centre for Algorithmic Transparency** (ECAT). It will also work closely with the **Scientific Panel of independent experts**, which ensures a strong link to the scientific community, the **Advisory Forum**, and **individual experts and organizations**, and create fora for providers of AI models and systems and the open-source community. The AI Office will also oversee the **AI Pact**, facilitating businesses to engage with the Commission and other stakeholders.

AI innovation package of the EU Commission

On January 24th 2024 the EU Commission launched a package of initiatives to **assist European startups and small and medium-sized enterprises (SMEs)** in the development of **trustworthy Artificial Intelligence** that respects EU values and rules ("**Innovation package**"). This set of measures includes:

- ❑ An amendment of the **EuroHPC Regulation** to establish **AI Factories,** which will support the EuroHPC Joint Undertaking to develop European supercomputers and offer a one-stop shop for startups and innovators.
- ❑ The establishment of the **AI Office**, which contributes to developing and coordinating AI policy at the European level and supervises the application and enforcement of the AI Act.
- ❑ The **EU AI Start-Up and Innovation Communication** specifying additional key activities regarding:
  - **financial support** to generative AI;
  - **public and private investments** in AI start-ups and scale-ups, including through venture capital or equity support;
  - the acceleration of the development and deployment of **Common European Data Spaces** to the AI community;
  - the **GenAI4EU initiative** to support the development of new use cases and emerging applications in the European industrial ecosystems and the public sector (e.g., in robotics, health, biotech, manufacturing, mobility, climate and virtual worlds).

The ECJ strategy on AI

The Court of Justice of the European Union recently published its **strategy for integrating artificial intelligence** into its activities. In particular, the Court aims to use AI to improve:

- the **efficiency and effectiveness** of administrative and judicial processes;
- the **quality and consistency** of judicial decisions; and
- **citizens' access and transparency**.

Efficiency and effectiveness of processes

The Court of Justice seeks to **shape a more effective and efficient judicial system** by enhancing efficiency, decreasing its workload and fine-tuning resource optimization.

To achieve this, the Court is testing several solutions, including:

- the **implementation of SIGA**, a case management system that will provide end-to-end case management (e.g., assisting users by detecting references in the text);
- the use of **speech-to-text machines** for the automatic production of hearing transcripts;
- the **newest version of Eureka**, a search engine for judicial documents that will enable semantic search (instead of a simpler keyword search).

Quality and consistency of processes

By integrating AI to support decision-making processes, the Court's staff will expedite their tasks and enhance decisions' transparency and accuracy.

To this end, the Court aims to:

- **leverage automation** (e.g., automated processing of decisions and conclusions in SIGA);
- **enhance legal research** (e.g., visual representation of cases to facilitate connection between cases);
- **leverage standardization** (e.g., applying a visual filter to rearrange document layouts).

Citizen's access & transparency

The Court aims to **promote equal access to justice for all individuals** to ensure a fair and equitable legal system and benefit individuals and society.

For example, the Court plans on relying on **assistive technologies** for people with disabilities (e.g., screen readers and virtual assistants), employing **natural language processing** (e.g., offering real-time translations) and **automating briefing production** (e.g., through AI avatars).

To identify the best AI tools suitable for these purposes, the Court will adopt a specific governance model and harness the potential of the Innovation Lab.

Italian DPA notifies breaches of privacy laws to OpenAI

While the AI Act was in the final stage of its approval process, the Italian Data Protection Authority (Garante per la protezione dei dati personali) reported **violations of data protection rules** to OpenAI, the artificial intelligence research company that created and launched ChatGPT's AI platform.

After imposing a **temporary processing ban** on OpenAI on March 30, 2023, due to the lack of a privacy policy and an adequate legal basis, and following a preliminary investigation, the Italian DPA concluded that available evidence suggests one or more breaches of EU GDPR provisions.

OpenAI may now submit its **legal brief** regarding the alleged violations within 30 days.

The Garante will consider the ongoing work of the **dedicated task force**, established by the European Data Protection Board (EDPB) to foster cooperation and exchange information on possible enforcement actions, in reaching its final decision.

Committee on artificial intelligence - Draft Framework Convention

On the 18th December 2023, the Committee on Artificial Intelligence of the Council of Europe ("**CAI**") released the latest version of the **Draft Framework Convention on Artificial Intelligence, human rights, democracy and the rule of law** ("**Draft Framework Convention**"). This document, currently under development and set to serve as the basis for the final reading, aims to:

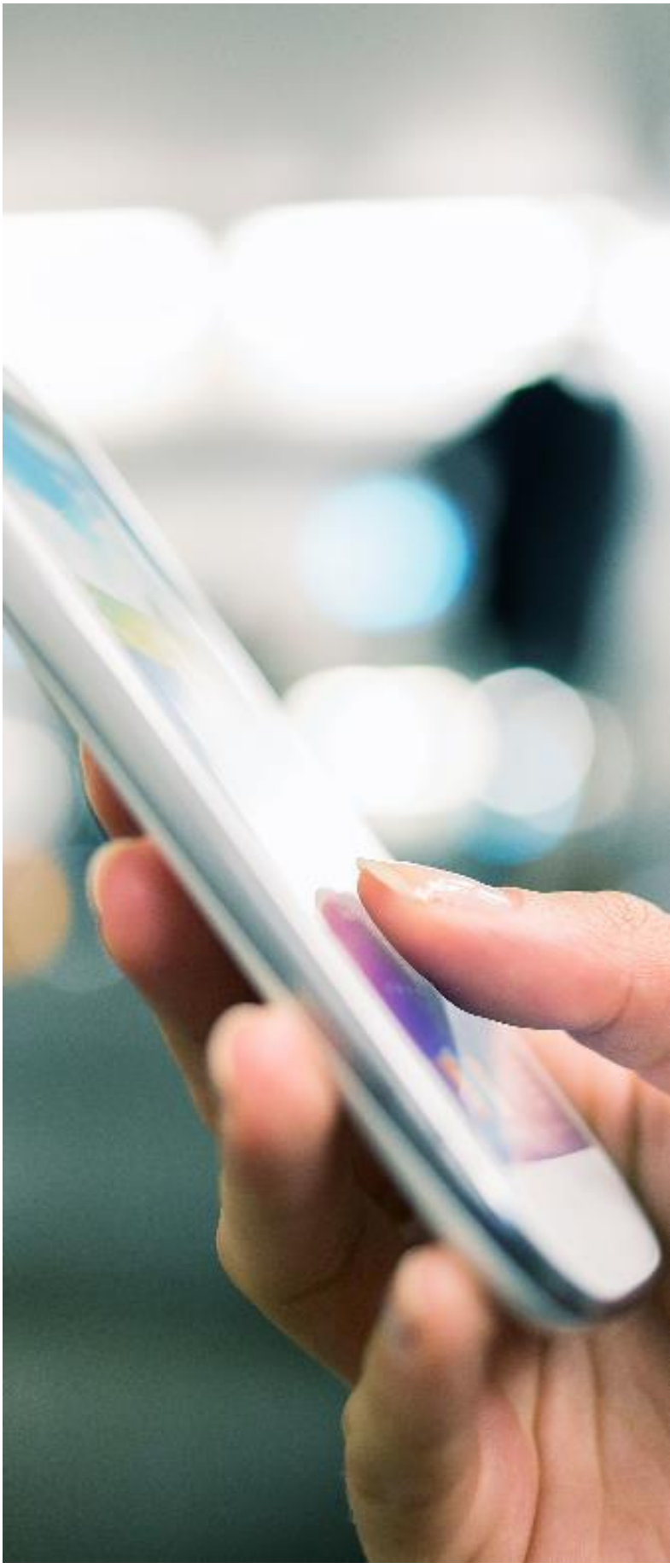
- Ensure that the **activities within the lifecycle of AI systems are compatible with obligations to protect human rights**, as enshrined in applicable international and domestic law.
- Ensure that AI systems **are not used to undermine the integrity, independence, and effectiveness of democratic institutions and processes**, including the principle of separation of powers, respect for judicial independence, and access to justice.
- **Protect individuals' participation in democratic processes, fair access to public debate, and the ability of individuals to reach decisions free from undue/harmful and malicious external influence or manipulation** in the context of activities within the lifecycle of AI systems.

The Draft Framework Convention also establishes a set of **principles** related to activities within the lifecycle of AI systems, including:

- **Human dignity** and **individual autonomy**.
- **Transparency** and **oversight** tailored to the specific contexts and risks.
- **Accountability** and **responsibility.**
- **Equality** and **non-discrimination.**
- **Privacy** and **personal data protection** (e.g. effective guarantees and safeguards shall be put in place for data subjects).
- **Preservation of health and the environment**.

- **Reliability, safety, validity, and trust** in AI systems, including accuracy, data quality, data integrity, data security, cybersecurity, and robustness.
- **Safe innovation**.

The CAI mandates the adoption of **measures regarding the identification, assessment, prevention, and mitigation of risks and impacts** to human rights, democracy, and the rule of law arising from the design, development, use, and decommissioning of AI systems.



ISO/IEC 42001:2023: the world's first standard for AI management systems.

The **adoption of an AI management system** to extend the existing management structures is a **strategic decision** for **any organization** (regardless of size, type and nature) that provide or deploy AI-based products or services.

The establishment and implementation of the AI management system should consider the **many use cases for AI** and the need to strike the appropriate **balance between governance mechanisms and innovation**.

On the 18th December 2023, the International Organization for Standardization (**ISO**) adopted the world's first standard for AI management systems, the **ISO/IEC 42001:2023**.

The standard outlines **requirements for establishing, implementing, maintaining and continually improving an AI management system** within the context of an organization.

It offers a structured framework for managing risks, balancing innovation and accountable governance, and ensuring compliance with legal and regulatory obligations.

Organizations can elect to apply the requirements of the standard using a **risk-based approach** to ensure that the **appropriate level of control** is applied for the **particular AI use cases, services or products** within the organization's scope.

World Health Organization- AI ethics and governance guidance for large multi-modal models.

On the 18th January 2024, the World Health Organization ("**WHO**") released a new guidance on the ethics and governance of large multi-modal models ("**LMMs**") expected to see widespread use in the healthcare sector.

The WHO outlines the main applications of LMMs, including their potential benefits and risks:

- **Diagnosis and clinical care** that assists, for example, in managing complex cases or reviewing routine diagnoses.
- **Patient-guided use**, such as virtual health assistants.
- **Clerical and administrative tasks**, such as drafting clinical notes after patients' visits.
- **Medical and nursing education**, including providing trainees with simulated patient encounters.
- **Scientific research and drug development**.

Despite the significant benefits offered by LMMs, their use can carry several **risks**, including the production of hallucinated, inaccurate, biased, or incomplete statements, which could cause serious harm to patients.

The guidance also addresses broader **regulatory and systemic risks** that may arise from the use of LMMs, including **cybersecurity risks**, **accessibility**, and **affordability** concerns (considering the "digital divide" and subscription fees to access LMMs), impacts on **labour**, and system-wide **biases**.

Finally, the WHO recommends that governments, technology companies, and healthcare providers ensure the appropriate development, provision, and deployment of LMMs for healthcare. For example:

- Governments should require **post-release auditing and impact assessments**, including for data protection and human rights.
- Governments should ensure **international governance** that encourages companies to develop and deploy LMMs meeting adequate international standards of safety and efficacy.
- Developers should ensure **transparency** in training data and **involve diverse stakeholders** in design.

Stanford University - Rethinking Privacy in the AI Era

In February 2024, Stanford University published a whitepaper titled "*Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World*" ("**Whitepaper**") that explores how existing and future privacy and data protection frameworks regulate the development and adoption of AI, providing solutions to protect and preserve privacy.

The Whitepaper starts from the observation that **the connective tissue between privacy and AI** is **data**, which forms the foundations of all AI systems, requiring massive amounts of data for training. This unlimited information collection entails unique **individual and societal privacy risks** that the current legislative landscape fails to regulate appropriately. Current privacy and data protection laws - which indirectly govern AI development - are insufficient to address individual and systemic privacy harms posed by AI systems. On the other hand, laws explicitly addressing algorithmic decision-making and other AI aspects do not provide appropriate rules for governing the use of large amounts of data in AI systems.

This creates a need for **new policy approaches** to safeguard privacy in an increasingly AI-dominant world.

The Authors of the Whitepaper propose three provocations to reduce the risks to privacy posed by the development and adoption of AI.

- **Denormalize data collection by default by shifting away from opt-out to opt-in data collection**. Data collectors should operationalize the data minimization principle by adopting privacy by default strategies, standards and infrastructure for meaningful consent mechanisms.
- **Focus on the AI data supply chain to improve privacy and data protection**. Ensuring dataset transparency and accountability across the entire life cycle must be a focus of any regulatory system that addresses data privacy (e.g. contracts, policies, etc.). The quality of data should be ensured in all the phases. A strategy that looks beyond individual rights is necessary to address adequately the complexities connected with the data supply chain.

- **Flip the script on the creation and management of personal data**. To ensure and automate the exercise of individual data rights and preferences, policymakers should promote the development of innovative governance mechanisms and technical infrastructure, such as data intermediaries and data permissioning infrastructure.

Contacts

Ida Palombella

Partner - ipalombella@deloitte.it

Pietro Boccaccini

Director - pboccaccini@deloitte.it

Research & Editorial

Simone Prelati

Alessandro Amoroso

Camilla Torresan

Benedetta Antonelli

Margherita Demattè

Deloitte.
Legal