EU AI Act: 101

By IAPP Westin Fellow Andrew Folks

Purpose of the AI Act

→ To lay down a comprehensive legal framework for the development, marketing and use of AI in the EU in conformity with EU values.

→ To promote the uptake of human-centric and trustworthy AI while ensuring a high level of protection of health, safety and fundamental rights, including democracy, the rule of law and environmental protections.

→ To support innovation while mitigating harmful effects of AI systems in the EU.

Key changes the AI Act will bring

→ Classifies AI systems by level of risk and mandate development, deployment, and use requirements, depending on the risk classification.

→ Establishes the AI Office to oversee general-purpose AI models, contribute to fostering standards and testing practices, and enforce rules across member states; the AI Board to advise and assist the European Commission and member state competent authorities; the Advisory Forum to advise and provide technical expertise to the board and the Commission; and Scientific Panel of independent experts to support implementation and enforcement of the act.

→ Prohibits unacceptable risk AI.

→ Introduces heightened technical and documentary requirements for high-risk AI systems, including fundamental rights impact assessments, and requires conformity assessments.

→ Requires human oversight and data governance.

Key challenges posed by the AI Act

→ Protecting the fundamental rights to the protection of personal data, private life and confidentiality of communications through sustainable and responsible data processing in the development and use of AI systems.

→ Fostering innovation and competitiveness in the AI ecosystem, and facilitating its development.

→ Understanding the interplay between the AI Act and existing rules applicable to AI, including on data protection, intellectual property and data governance.

→ Navigating the complex supervision and enforcement stakeholder map that is forming.

→ Designing and implementing appropriate multidisciplinary governance structures within organizations.

Important upcoming dates

→ The AI Act shall enter into force 1 Aug. 2024, following its publication in the Official Journal of the European Union 12 July 2024. It will be fully applicable 24 months after entry into force, with a graduated approach as follows:

- 2 Feb. 2025: Prohibitions on unacceptable risk AI become applicable.
- 2 Aug. 2025: Obligations for general-purpose AI governance become applicable.
- 2 Aug. 2026: All rules of the AI Act become applicable, including obligations for high-risk systems.
- 2 Aug. 2027: Obligations for all other high-risk systems become applicable.

Additional resources

→ IAPP AI Governance Center → EU AI Act: Next Steps for Implementation → EU AI Act Cheat Sheet → European Commission's AI – Questions and Answers

Updated July 2024.
iapp.org

FOCUS AREAS	AI ACT				
ORGANIZATIONS WITHIN SCOPE	Applies to: → Providers, importers and distributors of AI systems or general-purpose AI models that are placed on the EU market, put into service or used in the EU, even if they were established in a third country.				
DEFINITION OF AN AI SYSTEM	A machine-based system designed to operate with varying levels of autonomy 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.				
RISK LEVELS/ CATEGORIES	**Prohibited AI practices**	**High-risk AI system**	**General-purpose AI models**	**General-purpose AI models with systemic risks**	**Transparency requirements**
	A limited set of particularly harmful uses of AI that contravene EU values because they violate fundamental rights. These include: → Social behavioral scoring systems. → Emotion-recognition systems at work and in education. → AI used to exploit people's vulnerabilities, such as their ages or disabilities. → Behavioural manipulation and circumvention of free will. → Untargeted scraping of facial images for facial recognition. → Biometric categorization systems that use certain sensitive characteristics. → Specific predictive policing applications. → Law enforcement use of real-time biometric identification in public, apart from in limited, authorized situations.	AI systems that pose a significant risk to health, safety or fundamental rights, including in the following categories: → Profiling. → Medical devices. → Vehicles. → Emotion-recognition systems. → Law enforcement. → Credit scoring AI systems that are used as a safety component of a product, or are a product themselves, and are regulated by relevant EU sectoral product safety laws.	AI models that display significant generality, are capable of competently performing a wide range of distinct tasks, regardless of how they are placed on the market, and can be integrated into a variety of downstream systems or applications.	General-purpose AI models that have "high-impact capabilities" that match or exceed the capabilities recorded in the most advanced general-purpose AI models. These models must either: → Have high-impact capabilities evaluated against appropriate technical tools, indicators and benchmarks. This is presumed where the cumulative amount of compute used for training is greater than 10^{25} floating point operations. → Be designated as general-purpose AI with systemic risks by the Commission or scientific panel.	Specific transparency obligations are imposed on providers and deployers of certain AI systems. An AI system may be subject to both the transparency requirements and the requirements for high-risk AI systems or general-purpose AI models.

	Prohibited AI practices	High-risk AI system	General-purpose AI models	General-purpose AI models with systemic risks	Transparency requirements
KEY REQUIREMENTS	Such AI systems are prohibited under the AI Act.	These systems require providers to ensure: → Risk management → Data quality and governance. → Documentation and traceability. → Transparency. → Human oversight. → Accuracy, cybersecurity and robustness. → Demonstrated compliance via conformity assessments. → If deployed by public authorities, registration in a public EU database, unless used for law enforcement or migration.	Require providers to: → Perform FRIAs and conformity assessments. → Implement risk management and quality management systems to continually assess and mitigate systemic risks. → Inform individuals when they interact with AI. AI content must be labelled and detectable. → Test and monitor for accuracy, robustness and cybersecurity. General-purpose AI models with systemic risk are subject to greater testing and reporting requirements.	Specific obligations for general-purpose AI models with systemic risks include: → Standardized model evaluation and adversarial testing. → Assessment and mitigation of possible systemic risks. → Documentation of serious incidents and corrective measures. → Adequate level of cybersecurity and physical infrastructure of general-purpose AI models with systemic risks. → Adherence to codes of practice. → Compliance with confidentiality requirements laid under Article 78 for any obtained information or documentation, as well as trade secrets.	Providers of AI systems that interact directly with natural persons must design and develop AI systems to guarantee individual users are aware they are interacting with an AI system. Providers of AI systems that generate synthetic audio, image, video or text content and general-purpose AI systems must ensure outputs are marked in a machine-readable format and detectable as AI generated or manipulated. Deployers of emotion recognition or biometric categorization systems must inform natural persons of the systems' operations. Deployers of AI systems that produce deepfakes or manipulate text meant to inform the public on matters of public interest must abide by disclosure requirements.

SIGNIFICANT PROVISIONS	**Transparency.** Requirements are imposed on certain AI systems, for example when there is a clear risk of manipulation such as via the use of chatbots. Users should be aware that they are interacting with a machine. **Conformity assessments.** For AI systems that are safety components of products or are products themselves under listed EU product safety laws, conformity assessments must be performed prior to the systems' placement EU market or when a high-risk AI system is substantially modified. Importers of AI systems must ensure the foreign provider has already carried out the appropriate conformity assessment procedure. **FRIAs.** Before deployment, public law entity deployers, private operators who provide a public service and private operators who deploy AI systems to evaluate creditworthiness, emergency calls, or risk and pricing for life and health insurance must assess the systems' impact on fundamental rights. If a data protection impact assessment is required, the FRIA should be conducted in conjunction with that DPIA. **Generative AI.** Providers that generate synthetic audio, image, video or text content must ensure that content is marked in a machine-readable format and is detectable as artificially generated or manipulated.
ENFORCEMENT AND PENALTIES	The AI Office, housed within the European Commission, will supervise AI systems based on general-purpose models when the model and system are provided by the same provider. It will have the powers of a market surveillance authority. National market surveillance authorities are responsible for the supervision of all other AI systems. The AI Office will work to coordinate governance among member countries and supervise enforcement of rules related to general-purpose AI. Member state authorities will lay down rules on penalties and other enforcement measures, including warnings and nonmonetary measures. Individuals can lodge a complaint of infringement with a national competent authority, which can then launch market surveillance activities. The act does not provide for individual damages. There are penalties for: → Prohibited AI violations, up to 7% of global annual turnover or 35 million euros. → Most other violations, up to 3% of global annual turnover or 15 million euros. → Supplying incorrect information to authorities, up to 1% of global annual turnover or 7.5 million euros. The AI Board will advise on the act's implementation, coordinate between national authorities and issue recommendations and opinions.
FURTHER RULEMAKING	The Commission can issue delegated acts on: → Definitions of an AI system. → Criteria and use cases for high-risk AI. → Thresholds for general-purpose AI models with systemic risk. → Technical documentation requirements for general-purpose AI. → Conformity assessments. → EU declarations of conformity. The AI Office is to draw up codes of practice to cover, but not necessarily limited to, obligations for providers of general-purpose AI models. Codes of practice should be ready nine months after the act enters into force at the latest and should provide at least a three-month period before taking effect.