

Règlement Européen sur l'I.A

Le règlement européen sur l'intelligence artificielle (IA) vise à fixer des règles pour encadrer les systèmes d'IA. Les objectifs visés par le règlement IA ou AI Act sont principalement l'harmonisation des normes au niveau de tous les Etats membres et des mesures visant la protection des droits afin de permettre à l'Europe de continuer à innover de manière responsable dans un cadre flexible et adapté.

Les principaux éléments à retenir visent les systèmes d'IA à haut risque qui doivent être surveillés via l'imposition d'exigences strictes. L'accent est également mis sur la transparence des systèmes IA, indépendamment de leur niveau de risque, afin de permettre aux utilisateurs d'être informés sur les systèmes avec lesquels ils interagissent.

Des bacs à sable ou sandbox réglementaires sous forme d'espaces de test encadrés font également partie des mesures prévues pour soutenir l'innovation.

Les étapes et le calendrier de l'AI Act

Le règlement européen sur l'IA a été adopté officiellement par le Conseil européen le **21 mai 2024** et est entré en vigueur le **1er août 2024**.

Une des premières étapes du calendrier d'application de l'AI Act a eu lieu le **2 février 2025** avec l'interdiction des systèmes d'IA présentant des risques jugés inacceptables en relation avec les chapitres 1 et 2 du règlement.

Les prochaines étapes sont attendues le **2 août 2026** avec l'application complète aux systèmes d'IA à haut risque et leurs obligations spécifiques. Les systèmes d'IA à haut risque déjà identifiés sont, entre autres, ceux relatifs à la biométrie, aux infrastructures critiques, à l'éducation, à l'emploi, à la justice. Les « bacs à sable réglementaires » et leur mise en place sont également attendus pour cette même date.

La dernière étape est planifiée pour le **2 août 2027** avec l'application aux systèmes d'IA à haut risque incorporés dans certains produits réglementés tels que les dispositifs médicaux ou les jouets.

Règlement Européen sur l'I.A

Classification des systèmes d'IA via une approche basée sur le risque

L'AI Act repose sur une approche basée sur le risque et catégorise les systèmes d'IA en fonction du niveau de risque qu'ils représentent.

Les principaux niveaux de risques sont les suivants :

Risque inacceptable (systèmes IA avec interdiction totale) : ce niveau de risque inclut principalement le scoring social, l'IA manipulatrice, l'exploitation des vulnérabilités des individus et la reconnaissance faciale en temps réel, sauf pour ce qui concerne des exceptions spécifiques.

Risque élevé (système IA réglementé avec des obligations réglementaires strictes) : ce niveau de risque cible les IA utilisées dans les infrastructures critiques telles que celles dans les domaines de l'éducation, l'emploi, la justice, l'application des lois et encore d'autres.

Risque limité : ce niveau de risque prévoit une obligation d'information et de transparence envers les utilisateurs. Les éléments ciblés par ce niveau de risque sont par exemple les chatbots ainsi que les deepfakes.

Risque minimal : il n'y a pas de réglementations ou d'obligations spécifiques pour ce niveau de risque, où sont principalement concernés les filtres anti-spam ou les éléments tels que l'IA dans les jeux vidéo. Les acteurs sont cependant encouragés à adopter des codes de conduite.

Obligations pour les acteurs

Les développeurs considérés comme **fournisseurs d'IA à haut risque** doivent garantir la conformité via différents éléments tels que des documents techniques, la gestion des risques, des tests et audits ainsi que la bonne gestion des éléments de cybersécurité.

Les déployeurs d'IA à haut risque considérés comme des utilisateurs professionnels, bien que soumis à des obligations moindres, doivent quand même garantir un usage conforme aux règles de l'UE.

Les fournisseurs d'IA à usage général ou GPAI sont soumis à des obligations de documentation et de conformité avec les lois sur le droit d'auteur. Ils doivent également publier un résumé des données d'entraînement. Concernant **les GPAI à risque systémique**, en plus de ces obligations, les fournisseurs doivent effectuer des évaluations de risques, des tests adversariaux et signaler les incidents graves.

Règlement Européen sur l'I.A

Interdictions spécifiques

Concernant les interdictions spécifiques introduites par l'Article 5 du Chapitre 2 de l'AI Act, de manière synthétique, les IA suivantes sont interdites :

IA déployant des techniques subliminales, manipulatrices ou trompeuses pour influencer le comportement de manière nuisible et entraver la prise de décision éclairée, causant ainsi un préjudice important.

- IA exploitant les vulnérabilités telles que celles liées à l'âge, au handicap ou à la situation socio-économique pour fausser le comportement, causant ainsi un préjudice important.
Le scoring social, c'est-à-dire l'évaluation ou la classification d'individus ou de groupes sur la base de leur comportement social ou de leurs traits personnels, entraînant un traitement préjudiciable ou défavorable de ces personnes.
- La prédiction criminelle basée sur l'évaluation du risque qu'une personne commette des infractions pénales en se fondant uniquement sur le profilage ou les traits de personnalité, sauf lorsqu'elle est utilisée pour compléter des évaluations humaines fondées sur des faits objectifs et vérifiables directement liés à l'activité criminelle.
- La constitution de bases de données de reconnaissance faciale par l'extraction non ciblée d'images faciales sur Internet ou via des séquences de vidéosurveillance.
La déduction des émotions sur le lieu de travail ou dans les établissements d'enseignement, sauf pour des raisons médicales ou de sécurité.
- Les systèmes de catégorisation biométrique déduisant des attributs sensibles (race, opinions politiques, appartenance syndicale, croyances religieuses ou philosophiques, la vie sexuelle ou l'orientation sexuelle, etc), à l'exception de l'étiquetage ou du filtrage d'ensembles de données biométriques acquis légalement ou lorsque les forces de l'ordre catégorisent des données biométriques.
- L'identification biométrique à distance (IBD) « en temps réel » dans des espaces accessibles au public pour les forces de l'ordre, sauf dans les cas suivants
 - la recherche ciblée de personnes disparues, de victimes d'enlèvement et de personnes victimes de la traite des êtres humains ou de l'exploitation sexuelle;
 - la prévention d'une menace spécifique, substantielle et imminente pour la vie ou la sécurité physique, ou d'une attaque terroriste prévisible ; ou
 - l'identification de suspects de crimes graves (par exemple, meurtre, viol, vol à main armée, trafic de stupéfiants et d'armes illégales, criminalité organisée, et autres), crimes contre l'environnement, etc).

Règlement Européen sur l'I.A

Les systèmes d'IA à haut risque

Les développeurs de systèmes d'IA à haut risque doivent respecter des obligations strictes en ce qui concerne, entre autres, la gestion des risques, la transparence et la cybersécurité.

Les systèmes d'IA considérés comme étant à haut risque sont ceux utilisés dans les secteurs et environnements dont, parmi les principaux :

- La biométrie avec des éléments tels que la reconnaissance et la classification faciale ou encore les systèmes d'émotion ;
- Les infrastructures dites critiques comme l'énergie, l'eau et les transports ;
- L'accès aux services essentiels tels que le domaine des assurances, l'accès au crédit ou aux services d'urgence ;
- Les domaines de l'éducation et de l'emploi ;
- L'application des lois et de la justice (profilage criminel, évaluation de preuves, etc) ;
- La migration et le contrôle des frontières ;
- Les processus démocratiques.

Les fournisseurs d'IA à haut risque, doivent, parmi leurs principales obligations :

- mettre en place un système de gestion des risques tout au long du cycle de vie du système, assurer la gouvernance des données,
- établir une documentation technique démontrant la conformité et fournissant aux autorités toutes les informations nécessaires liées à cette évaluation de la conformité,
- concevoir le système pour qu'il enregistre automatiquement les événements pertinents relatifs à l'identification des risques au niveau national ainsi que les modifications substantielles tout au long du cycle de vie,
- fournir des instructions d'utilisation,
- prévoir la mise en place d'une surveillance humaine, prévoir un système de gestion de la qualité pour garantir la conformité et veiller à ce que la conception du système atteigne les niveaux attendus de précision, robustesse et cybersécurité.

Règlement Européen sur l'I.A

Les GPAI ou IA à usage général

Les fournisseurs de GPAI doivent documenter et publier les données d'entraînement. Des obligations allégées ont été mises en place pour les modèles sous licence libre ou ouverte, sauf s'ils sont considérés comme « systémiques » selon un seuil défini lorsque la quantité cumulée de calcul utilisée pour la formation est supérieure à 1025 opérations en virgule flottante (FLOP). Dans ce cas, les modèles considérés comme présentant un risque systémique ont des obligations supplémentaires en matière de sécurité, de test et de rapports d'incidents. Ils doivent ainsi effectuer des évaluations de modèles avec la mise en place et la documentation des tests contradictoires dans l'objectif d'identifier et atténuer le risque systémique. Les incidents graves doivent être repérés, documentés et signalés à l'Office AI et aux autorités nationales compétentes dans les meilleurs délais avec les éventuelles mesures correctrices. La cybersécurité doit également être assurée de manière à atteindre un niveau adéquat attendu.

Gouvernance

Un bureau de l'IA ou AI Office va être mis en place sous l'autorité de la Commission européenne pour superviser l'application de la réglementation et coordonner les efforts des Etats membres dans l'objectif d'une mise en œuvre uniforme.

Ce bureau aura un rôle d'interlocuteur clé et les entreprises auront la possibilité de déposer des plaintes contre les fournisseurs en cas de non-conformité. Des audits seront également possibles en cas d'identification de risques systémiques.

Les systèmes d'IA implémentés au sein de l'UE mais provenant de pays tiers seront soumis aux mêmes règles afin de garantir la conformité aux normes européennes.

Sources :

<https://artificialintelligenceact.eu/fr/>

<https://www.entreprises.gouv.fr/la-dge/actualites/le-reglement-europeen-sur-lintelligence-artificielle-publics-concernes-dates-cles>

<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32024R1689>

European regulation on artificial intelligence

The European Regulation on Artificial Intelligence (AI) aims to establish rules to govern AI systems. The main aims of the AI Regulation or AI Act are to harmonise standards across all Member States and to introduce measures to protect rights so that Europe can continue to innovate responsibly within a flexible and appropriate framework.

The main elements are aimed at high-risk AI systems, which must be monitored by imposing strict requirements. Emphasis is also placed on the transparency of AI systems, regardless of their level of risk, to enable users to be informed about the systems with which they interact.

Regulatory sandboxes in the form of supervised test areas are also among the measures planned to support innovation.

The stages and timetable of the AI Act

The European AI Regulation was formally adopted by the European Council on **21 May 2024** and entered into force on **1st August 2024**.

One of the first steps in the AI Act timetable took place on **2 February 2025** with the banning of AI systems with risks deemed unacceptable in relation to Chapters 1 and 2 of the Regulation.

The next steps are expected on **2 August 2026** with the full application to high-risk AI systems and their specific obligations. The high-risk AI systems already identified include those relating to biometrics, critical infrastructures, education, employment and justice. The 'regulatory sandboxes' and their implementation are also expected by the same date.

The final stage is planned for **2 August 2027**, with application to high-risk AI systems incorporated in certain regulated products such as medical devices or toys.

European regulation on artificial intelligence

Classification of AI systems using a risk-based approach

The AI Act is based on a risk-based approach and categorises AI systems according to the level of risk they represent.

The main levels of risk are as follows:

Unacceptable risk (AI systems with a total ban): this level of risk mainly includes social scoring, manipulative AI, exploitation of individuals' vulnerabilities and real-time facial recognition, with specific exceptions.

High risk (regulated AI system with strict regulatory requirements): this risk level targets AI used in critical infrastructures such as education, employment, justice, law enforcement and others.

Limited risk: this level of risk includes an obligation to provide information and transparency to users. The elements targeted by this level of risk are chatbots and deepfakes, for example.

Minimal risk: there are no specific regulations or obligations for this level of risk, which mainly concerns spam filters or elements such as AI in video games. However, players are encouraged to adopt codes of conduct.

Obligations for players

Developers considered to be suppliers of **high-risk AI** must ensure compliance through various elements such as technical documents, risk management, tests and audits, as well as the proper management of cybersecurity elements.

Deployers of high-risk AI considered to be business users, although subject to lesser obligations, must still ensure compliance with EU rules.

General-purpose AI providers, or GPAIs, are subject to documentation and copyright compliance obligations. They must also publish a summary of the training data. **For systemically risky GPAIs**, in addition to these obligations, providers must carry out risk assessments and adversarial tests and report serious incidents.

European regulation on artificial intelligence

Specific prohibitions

With regard to the specific prohibitions introduced by Article 5 of Chapter 2 of the AI Act, in summary, the following AI are prohibited:

- AI deploying subliminal, manipulative or deceptive techniques to influence behaviour in a harmful way and hinder informed decision-making, thereby causing significant harm.
AI exploiting vulnerabilities such as age, disability or socio-economic status to distort behaviour, causing significant harm.
- Social scoring, i.e. the evaluation or classification of individuals or groups based on their social behaviour or personal traits, resulting in prejudicial or unfavourable treatment of these individuals.
- Criminal prediction based on the assessment of the risk of a person committing criminal offences based solely on profiling or personality traits, except where it is used to supplement human assessments based on objective and verifiable facts directly related to criminal activity.
- The creation of facial recognition databases through the untargeted extraction of facial images from the Internet or video surveillance footage.
- The deduction of emotions in the workplace or educational establishments, except for medical or security reasons.
- Biometric categorisation systems inferring sensitive attributes (race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation, etc.), with the exception of the labelling or filtering of legally acquired biometric data sets or when law enforcement agencies categorise biometric data.
- Real-time remote biometric identification (RBI) in publicly accessible areas for law enforcement, except in the following cases :
 - Targeted searches for missing persons, kidnap victims and victims of human trafficking or sexual exploitation;
 - The prevention of a specific, substantial and imminent threat to life or physical safety or a foreseeable terrorist attack; or
 - The identification of suspects in serious crimes (e.g. murder, rape, armed robbery, drug and illegal weapons trafficking, organised crime, environmental crime, etc.).

European regulation on artificial intelligence

High-risk AI systems

Developers of **high-risk AI** systems must comply with strict obligations regarding risk management, transparency, and cybersecurity, among other things.

AI systems considered to be high-risk are those used in sectors and environments including, among the main ones:

- Biometrics, with elements such as facial recognition and classification and emotion systems;
- Critical infrastructures such as energy, water and transport;
- Access to essential services such as insurance, credit and emergency services;
- Education and employment;
- Law enforcement and justice (criminal profiling, evidence evaluation, etc);
- Migration and border control;
- Democratic processes.

Among their main obligations, high-risk AI providers must implement:

- A risk management system throughout the system's lifecycle,
- ensure data governance,
- draw up technical documentation demonstrating compliance and provide the authorities with all the necessary information relating to this compliance assessment,
- design the system so that it automatically records events relevant to the identification of risks at a national level and substantial changes throughout the lifecycle,
- provide instructions for use,
- provide for human oversight,
- provide a quality management system to ensure compliance and ensure that the system's design achieves the expected levels of accuracy, robustness, and cyber security.

European regulation on artificial intelligence

GPAI or general-purpose AI

GPAI providers must document and publish training data. Reduced obligations have been introduced for models under a free or open licence unless considered 'systemic' according to a threshold defined when the cumulative amount of computation used for training is greater than 1025 floating point operations (FLOP). In this case, models considered to present a systemic risk have additional obligations regarding security, testing and incident reporting. They must, therefore, carry out model evaluations with the implementation and documentation of contradictory tests with the aim of identifying and mitigating systemic risk. Serious incidents must be identified, documented and reported to the AI Office and the relevant national authorities as soon as possible, along with any corrective measures that may be required. Cybersecurity must also be ensured to the expected adequate level.

Governance

An AI Office will be established under the European Commission's authority to oversee the regulations' application and coordinate the efforts of the Member States to achieve uniform implementation.

This office will act as a key point of contact, and companies will be able to lodge complaints against suppliers in the event of non-compliance. If systemic risks are identified, audits will also be possible.

AI systems implemented within the EU but originating from third countries will be subject to the same rules to ensure compliance with European standards.

Source:

<https://artificialintelligenceact.eu/fr/>

<https://www.entreprises.gouv.fr/la-dge/actualites/le-reglement-europeen-sur-lintelligence-artificielle-publics-concernes-dates-cles>

<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32024R1689>