MetaCompliance®

# Introduction to the EU AI Act

Start your journey to compliance with the world's first legal framework for Artificial Intelligence.

# Contents

# AI and the Need for Regulation

**In 1950, Alan Turing published a paper titled "Computing Machinery and Intelligence", proposing the (now) widely known Turing Test. The test was simple: a human evaluator reads a transcript of a natural-language conversation between a human and a machine. If the evaluator can't reliably tell them apart, the machine passes.**

The test was simple: a human evaluator reads a transcript of a natural-language conversation between a human and a machine. If the evaluator can't reliably tell them apart, the machine passes.

It would be 64 years before a chatbot named **Eugene Goostman** would pass the test — albeit by pretending to be a 13-year-old boy. Less than a decade later, many AI systems can comfortably pass the test while masquerading as fully fledged adults. After decades of scientists and engineers quietly working away in the background, AI has finally exploded into the mainstream and seen widespread adoption by organisations and individuals all over the world.

**The last decade has seen the rise of:**

**Neural networks in the 2010s for image recognition and virtual assistants, e.g., Alexa.**

**Large Language Models (LLMs) that enable tools like ChatGPT in the early 2020s.**

**Use of AI for medical research and drug development, e.g., for COVID-19.**

**Widespread AI use for both business and personal endeavours.**

In a few short years, we've gone from 'interesting but functionally limited' to 'extremely valuable in a wide range of contexts'. However, this rate of development isn't without its risks and concerns. The EU Commission is so concerned about AI's potential risks and hazards that it has decided to implement the world's first legal framework governing AI systems.

This guide provides a detailed introduction to the EU AI Act, including:

- **The potential benefits, risks, and consequences of AI systems**
- **Why the EU believes AI regulation is necessary**
- **The stated goals and timeline for the EU AI Act**
- **Which organisations will be affected by the act**
- **The four risk categories for AI systems**
- **Transparency requirements for AI systems**
- **Penalties for non-compliance**
- **The regulatory and enforcement landscape**

Most importantly, we'll cover what the EU AI Act could mean for your company and how to begin your journey to compliance ahead of the relevant enforcement dates.

# The Good, the Bad, and the Unacceptable

**It's easy to see how today's crop of AI systems and tools could benefit individuals, organisations, and society as a whole. It's also plain that certain AI tools are already demonstrating a significant negative impact in some areas.**

What's not easy is forecasting how these benefits and consequences might evolve over the next decade. In this section, we'll look at how AI is being used for good — and the potential consequences the EU Council hopes to mitigate through regulation.

## Current and Anticipated Uses of AI

AI promises a wide range of business and personal benefits, including:

- **Better healthcare, including research and drug development**
- **Safer and cleaner transport**
- **More efficient manufacturing**
- **Crime prevention through prediction, risk assessment, and trend analysis**

AI's potential to uncover patterns in huge datasets that would be difficult for humans to identify and study will yield many more benefits. Examples include business analytics, scientific research, cyber security, and more.

## Potential Consequences of AI

AI development isn't without risk. The potential threats posed by different AI systems will depend on how they are designed and implemented and what data they use.

Widely raised concerns and risks include:

**Intentional and unintentional bias or discrimination**
(e.g., ethnicity, gender, and age bias)

AI systems are often a "black box," so it's tough to tell whether bias has been involved in individual cases. Naturally, the potential for bias and discrimination in automated and AI-informed processes is a serious concern — particularly when those processes directly affect individuals.

**Threat to democracy**

A combination of AI-powered bots, highly realistic (but fake) video, audio, and photo content, and the creation of so-called "echo chambers" has already **threatened democratic processes** in several countries. Unregulated use of AI systems could likely lead to further polarisation in the public sphere and be used to manipulate election results covertly.

AI systems and content also threaten individuals' reputations, and there is significant potential for intentional smear campaigns against political groups and candidates.

### Unacceptable levels of surveillance and curtailing of personal freedoms

AI systems can already track individuals across diverse online and offline activities. While there may be legitimate uses of this for law enforcement and national security, uncontrolled use would infringe on people's personal freedoms and right to privacy.

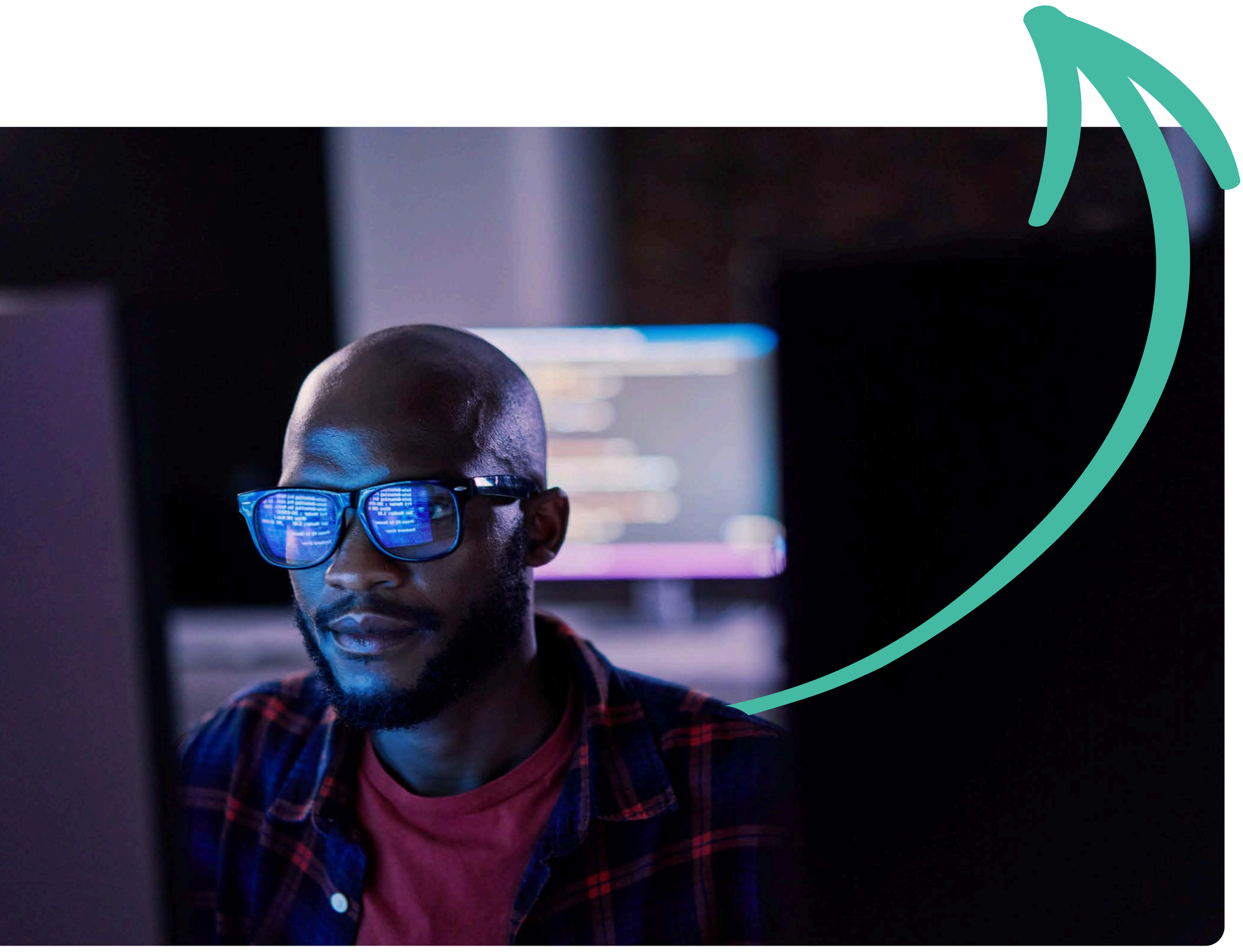
### Transparency issues

Even today, it's not always clear whether an individual is interacting with an AI system or a human. This will only grow more challenging over time, and these systems are already being used to perpetrate fraud and social engineering scams against individuals and businesses.

## Why Regulate AI in the EU?

The primary goal of regulation is to promote the development, sale, and use of human-centric and trustworthy AI systems. With the proper controls, AI technologies can potentially provide enormous benefits — for organisations, individuals, and humanity as a whole.

However, as we've just seen, unfettered development and use of AI pose an unacceptable risk. This is why the EU has decided that regulation is essential and that AI systems must support democracy, the rule of law, and respect for fundamental human rights.



# What is the EU AI Act?

**First proposed in April 2021, the EU AI Act (Regulation (EU) 2024/1689) is the world's first legal framework for AI. It aims to address and mitigate the risks of AI systems while positioning Europe to play a leading role in the space.**

The AI Act provides developers and deployers — vendors and other organisations — with clear requirements and obligations regarding the development, maintenance, and use of AI systems. The act aims to do this without placing undue burden on smaller organisations or stifling AI development across EU nations.

The act is part of a wider package of policies intended to support the development of trustworthy AI, which includes the **AI Innovation Package** and **Coordinated Plan on AI**. Together, these measures aim to guarantee the safety and fundamental rights of people and organisations as they relate to AI. They are also designed to strengthen investment and innovation in AI across the EU.

## Stated Purpose

The EU's stated goals for the AI Act are:

**Promote safety and accountability**
Some AI systems — particularly in critical areas like healthcare, transportation, and law enforcement — have the potential to cause great harm if they are misused or poorly designed. The act ensures accountability by establishing developer, vendor, and user responsibilities.

**Protect human rights**
Given the potential for bias, discrimination, privacy violations, and undue monitoring of private citizens, the act aims to place heavy restrictions on certain functionality of AI systems — and bans others altogether.

- **Support fair competition**
  With standardised rules across the EU, the act aims to ensure fair competition while encouraging companies to invest in AI development without fear of unpredictable legal challenges. The act also aims to ensure smaller companies aren't disproportionately burdened by its requirements.

- **Support innovation**
  The act aims to provide start-ups and smaller companies with opportunities to develop and train AI models before release. National authorities must provide companies with a testing environment that simulates real-world conditions.

**Support public trust and confidence**
AI is already being used to mislead and misinform individuals. By setting requirements and banning certain uses, the act aims to prevent further harm to public trust in institutions and public figures.

**Provide transparency**
Given the potential for AI systems to represent themselves as humans, the act aims to ensure transparency so individuals are always aware when they are interacting with an AI system. There will also be strict transparency requirements for the data used to train AI systems, particularly when it includes copyrighted works.

- **Trust.**
  So far, many AI systems have effectively been "black boxes," and it has been difficult to determine how results and outputs are reached. The act aims to ensure that Europeans can trust what AI provides.

## Timeline

The EU Parliament **adopted the AI** Act in March 2024, following which it was **approved by the Council** in May 2024, and came into force on 1st August 2024. It will be fully enforceable 24 months after this date, although some aspects of the act will apply sooner:

- **The ban on AI systems posing "Unacceptable Risk" applies six months after entry into force.**
- **Codes of practice apply nine months after entry into force.**
- **Transparency rules for general-purpose AI (GPAI) apply 12 months after entry into force.**

AI systems deemed "High Risk" will have more time to comply due to more stringent requirements. These requirements will apply 36 months after entry into force.

To facilitate the transition, the Commission launched the AI Pact — a voluntary initiative that invites AI developers to comply with the AI Act ahead of time.

## How is AI Defined?

The act defines an AI system as:

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

For a more specific understanding of the types of systems included, we can refer to **Annex I** of the act: "Artificial Intelligence Techniques and Approaches". Annex I identifies the following as characteristics of an AI system:

- **Machine Learning (ML) approaches, including supervised, unsupervised, and reinforcement learning, using a wide variety of methods including deep learning.**
- **Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems.**
- **Statistical approaches, Bayesian estimation, search and optimisation methods.**

Notably, these characteristics lean heavily into ML — a widely recognised subset of AI and the driving force behind most AI advancements in recent years. However, we also see various techniques and approaches that aren't necessarily components of what has traditionally been considered AI, but which contribute to a final product that could appear to be "intelligent". For example, while Bayesian estimation is commonly used in AI systems, it's not AI in its own right.

## Which Organisations Are Affected by the EU AI Act?

The Act has **"extra-territorial scope,"** meaning it will apply to all organisations operating inside the EU, even if they are based elsewhere. In other words, any company wanting to sell into the EU market must comply.

However, there are some exemptions. Most notably, for national security.

According to the act, national security remains the sole responsibility of Member States, in accordance with Article 4(2) TEU (Treaty on European Union). AI systems developed purely for national security uses — including military and defence — are exempt.

This exclusion applies to public and private entities developing AI systems solely for national security. However, it does not apply to dual-use technologies that are also used outside a national security context.

Other notable exemptions include:

**Scientific research**
AI systems developed purely for scientific research and development are considered to pose no significant threat, even if they would otherwise be regarded as High or Unacceptable Risk.

**Open source AI**
While there are no exemptions for open source AI systems considered to be High or Unacceptable Risk, other non-GPAI systems are not subject to the act's requirements so long as their source code is made public and there is no attempt to monetise them.

## Categorising AI Systems by Risk: The 4 Levels

The act divides AI systems into four categories based on the risk they pose to individuals, organisations, and society. The categories are:

**Minimal Risk**
These AI systems pose practically zero risk to safety, privacy, or individual rights — for example, AI-enabled video games or spam filters. According to the EU Parliament, most AI systems currently in use in the EU fall into this category.

The act allows free development, provision, and use of these systems without restrictions or requirements.

**Limited Risk**
This category covers AI systems that pose risks associated with a lack of transparency.

The act includes specific transparency obligations to ensure humans are informed when interacting with an AI system. For example, when using AI systems like chatbots, humans must be made aware of that fact. This aims to avoid situations where humans unknowingly interact with AI systems.

Providers must also ensure that AI-generated content is clearly labelled and identifiable. This includes AI-generated text published to inform the public on matters of public interest and any audio or video content that might be considered a "deep fake".

In short, if your company deploys, provides, or uses an AI system in this category, you must inform users when they are interacting with an AI system and label all audio, video and photo outputs as AI-generated.

## High Risk

This category covers AI systems that could harm health, safety, fundamental rights, environment, democracy, and the rule of law. The act divides High Risk AI systems into two categories:

1. **Systems used in products subject to the EU's product safety legislation, including toys, aviation, cars, medical devices, and lifts.**
2. **Systems that fall into areas that must be registered in an EU database, including:**

- **The potential benefits, risks, and consequences of AI systems**
- **Why the EU believes AI regulation is necessary**
- **The stated goals and timeline for the EU AI Act**
- **Which organisations will be affected by the act**
- **The four risk categories for AI systems**
- **Transparency requirements for AI systems**
- **Penalties for non-compliance**
- **The regulatory and enforcement landscape**

The High Risk category also includes certain systems involved with law enforcement, migration and border management, justice, and democratic processes.

All High Risk AI systems will be assessed before going to market, along with further assessments throughout their lifecycle. Individuals will have the right to file complaints about AI systems to designated national authorities.

The AI Act sets out a number of requirements for High Risk systems, including:

- **Adequate risk assessment and mitigation systems**
- **High data quality in datasets used to train and operate the system to minimise risks and discriminatory outcomes**
- **Activity logging to ensure traceability and analysis of results**
- **Detailed documentation of systems and their purpose for authorities to assess compliance**
- **Clear and adequate information must be provided to system deployers**
- **Appropriate human oversight measures throughout production and testing**
- **High levels of robustness, security, and accuracy**

## Unacceptable Risk

AI systems deemed to pose a serious threat to people are categorised as Unacceptable Risk.

They include:

- **"Cognitive behavioural manipulation" of individuals or specific vulnerable groups, e.g., voice-activated toys that encourage dangerous behaviour in children**
- **"Social scoring" systems that classify people based on their behaviour, socio-economic status, or personal characteristics, e.g., race, gender, sexual orientation, etc.**
- **Biometric identification and categorisation of people, including real-time and remote biometric identification systems such as facial recognition**
- **Untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases**
- **Emotion recognition in the workplace or schools**

AI Systems that fall into the Unacceptable Risk category are banned.

Some exceptions will be allowed for law enforcement purposes. For example:

- **Real-time remote biometric identification systems will be permitted in a limited number of serious cases.**
- **Remote biometric identification systems where identification occurs after a significant delay will be allowed to prosecute serious crimes if there is court approval.**

As noted earlier, there are also exemptions for scientific research and national security applications.

## Transparency Requirements

Generative AI systems, like ChatGPT and Claude, aren't classified as High Risk, but must comply with transparency requirements and EU copyright law. This includes:

**Disclosing that content (text, image, audio, and video) was generated by AI systems**

**Disclosing when content has been altered by AI systems**

**Designing models to prevent AI systems from generating illegal content**

**Publishing summaries of copyrighted data used for training**

High-impact GPAI models that could pose a systemic risk — for example, GPT-4 — will undergo thorough evaluations, and serious incidents must be reported to the European Commission.

## Penalties for Non-Compliance

The AI Act will be enforced with heavy fines:

- **Violations of prohibited AI restrictions will be subject to fines of up to €35m or 7% of global annual turnover for the preceding financial year.**
- **Violations of most other obligations will result in fines of up to €15m or 3% of global annual turnover for the preceding financial year.**
- **Supplying incorrect, incomplete, or misleading information to a competent authority in response to a request will be subject to fines of up to €7.5m or 1% of global annual turnover for the preceding financial year.**

The lower of the two amounts will be applied to small and medium-sized enterprises (SMEs), while the higher amounts will be applied to larger entities.

Exact fines will depend on the type of infringement and company size.

In line with **GDPR enforcement**, affected companies should expect these fines to be significant and act accordingly to ensure compliance with the new requirements.

## Who Will Enforce the EU AI Act?

This is a simple question, but the answer is a little complex. The AI Act sets out five — arguably six — different entities responsible for implementing and enforcing the AI Act.

## AI Office

Unveiled in May 2024, the AI Office sits within the European Commission. It will supervise the use of AI systems and GPAI models and take enforcement action in cases of non-compliance.

The AI Office's duties include:

- **Developing model contract terms, guidelines and templates**
- **Facilitating drawing up codes of practice (particularly for GPAI models)**
- **Receiving reports of serious incidents**
- **Working with competent authorities in Member States**
- **Providing advice on best practices and access to AI sandboxes**

The AI Office is responsible for monitoring and supervising providers of GPAI models and, where necessary, taking enforcement action against them.

## AI Board

The AI Board will include one representative from each Member State. The Board will advise and assist the Commission and Member States to ensure consistent and effective application of the AI Act and its requirements. To do this, the AI Board will:

- **Help coordinate between national competent authorities and market surveillance authorities**
- **Conduct joint investigations**
- **Collect and share technical and regulatory practices and expertise**
- **Advise on AI Act implementation and enforcement**
- **Promote AI literacy, public awareness, and understanding of AI systems**

## Advisory Forum and Scientific Panel

An EU-wide advisory forum will be established to provide technical expertise and advise the AI Board and the European Commission. In addition, the Commission will be supported by an independent scientific panel of experts to aid in enforcing the Act.

## National Competent Authorities

1. **Notifying authority**
   Responsible for setting up and conducting procedures for assessment, designation, and notification of "conformity assessment bodies" (see below).
2. **Market surveillance authority**
   Responsible for overseeing compliance with the AI Act and acting as a single point of contact for that Member State.

These bodies will serve as the NCA(s) for the purpose of the AI Act.

## Notified Bodies

Notified bodies are the conformity assessment bodies responsible for assessing high-risk AI systems, including testing, certification, and inspection. As noted above, they will be monitored by the notifying authorities.

These bodies must be independent of providers of any high-risk AI systems they are responsible for assessing and any other entity with an economic interest in those systems or their competitors.

# Next Steps

**The good news for companies in the AI space is that a high proportion of AI systems will be considered Minimal Risk, where no additional actions are required, or Limited Risk, where the new transparency requirements should be relatively straightforward to achieve.**

However, for systems that fall under the High Risk category, meeting the new requirements laid out by the AI Act will likely constitute a significant investment of time, energy, and money.

**If your company falls into the Limited or High Risk category, we recommend the following steps:**

1. Assess how the EU AI Act relates to your systems and products and determine which requirements are applicable.
2. Identify compliance partners or technical experts who can provide guidance and support as you begin and navigate your compliance journey.
3. Conduct a thorough gap analysis to identify areas where your AI systems currently fall short of the new requirements.
4. Educate your board and relevant stakeholders about the AI Act, develop a comprehensive plan for reaching compliance, and secure any additional budget needed to enact it.
5. **Implement the changes and measures needed to reach compliance. This may include revising policies, introducing additional monitoring and logging, and potentially updating development roadmaps and business strategies to reflect the act's requirements.**

Of course, some companies will find their AI systems falling into the Unacceptable Risk category. For these companies, there is likely to be a hard road ahead — either to adapt their systems to reduce the risk posed or to identify European entities that can purchase and use such systems due to their involvement in national security, scientific research, etc.

# Education is Key

**Wherever your systems fall on the AI risk scale, there is likely to be a need to educate relevant personnel to make them aware of the EU AI Act, its requirements, and its impact on your company.**

MetaCompliance's Cyber Security E-Learning platform contains modules on the EU AI Act, including:

**What is the EU AI Act?**
This module will help learners:

- Understand the purpose and scope of the EU AI Act
- Identify the key goals of the regulation in promoting trustworthy AI
- Recognise the importance of a harmonised legal framework for AI in the EU
- Learn how the Act balances innovation with public safety and fundamental rights
- Know the different risk levels applied to AI systems under the Act

**Compliance Requirements**
This module will help learners:

- Understand the steps needed to ensure compliance with the EU AI Act
- Learn how to classify AI systems based on their risk level
- Know mandatory requirements for high-risk AI systems
- Recognise the importance of ongoing monitoring and risk management for AI systems
- Understand the role of providers and deployers in ensuring compliance

**Understanding Risk Levels**
This module will help learners:

- Understand the EU AI Act's risk-based approach to regulating AI systems
- Identify the four risk categories for AI systems
- Learn what makes a system fall under the Unacceptable Risk category
- Know the requirements for high-risk AI systems
- Understand the obligations for low-risk and minimal-risk systems

Find out how MetaCompliance can help educate your stakeholders on the journey to EU AI Act Compliance.

Start learning