WHITEPAPER

A guide to the EU AI Act: Regulations, compliance and best practices

protiviti®
Global Business Consulting

Introduction

As artificial intelligence (AI) continues its explosive growth within organisations around the world, with virtually every business function exploring opportunities to increase productivity, efficiency and revenue growth, a growing collection of regulations, standards and frameworks around the world is beginning to emerge. Among the most notable of these regulations is the European Union Artificial Intelligence Act, which went into effect in August 2024.

In this paper, we provide answers to some of the questions we are hearing most frequently in the market today in connection with the act. For more information and Protiviti's latest perspectives on AI, visit our AI Intelligence Hub at www.protiviti.com/us-en/artificial-intelligence-services.

Quick overview of the EU AI Act

The EU AI Act is a comprehensive regulatory framework designed to govern the development, deployment and use of AI across the EU. The act defines AI in broad terms, encompassing any system that can operate with elements of autonomy and achieve a given set of objectives using machine learning, logic-based and knowledge-based approaches. This broad definition ensures that a wide range of automated systems fall under its regulatory scope. The EU AI Act aims to address various aspects of AI, including accountability, risk management, data governance, robustness, security and transparency.

What are the requirements in the EU AI Act about explainability and transparency?

The EU AI Act emphasises the importance of transparency and explainability throughout its text. It states, "This regulation aims to strengthen the effectiveness of such existing rights and remedies by establishing specific requirements and obligations, including in respect of the transparency, technical documentation and record-keeping of AI systems." Additionally, the act defines transparency as ensuring that AI systems are developed and used in a manner that allows for appropriate traceability and explainability. This includes making humans aware when they are communicating or interacting with an AI system, as well as informing developers and deployers about the capabilities and limitations of the AI system and educating affected persons about their rights. Both the letter and the spirit of the act reinforce the concepts of transparency and explainability, placing the responsibility on designers and deployers to inform customers and other stakeholders.

How are AI systems classified and/or identified as being subject to the requirements of the EU AI Act?

The act uses a risk classification approach to identify the severity levels of risk associated with AI systems. This classification system consists of four distinct groups:

- The first group is unacceptable risk, which is explicitly prohibited by the act due to the high likelihood of overt violations of fundamental rights such as citizen profiling and scoring mechanisms.
- The second group is high risk, which requires specific compliance measures to be implemented, as these risks pose threats to individual health, safety and fundamental rights.
- The third group is limited risk, which mandates transparency requirements to be put in place to prevent manipulation of preferences and behavioral bias.
- The fourth group is minimal risk, which is subject to nominal restrictions.

Most violations of the act will cost companies €15 million or 3% of annual global turnover, but can go as high as €35 million or 7% of annual global turnover for violations related to AI systems that the act prohibits.

How does an organisation determine which risk group is applicable to their AI implementation?

To determine which EU AI Act risk category applies to an AI use case, one must consider the potential impact of the AI system on human rights, safety and fundamental values. Here are some key points to help identify the appropriate category:

- **Unacceptable risk:** These AI systems are banned in the EU due to their high likelihood of violating fundamental rights. The EU AI Act offers examples, which include AI systems that manipulate human behavior, exploit vulnerabilities, engage in social scoring, use subliminal techniques, exploit children, or employ biometric data for mass surveillance or social categorisation.
- **High risk:** AI systems that pose significant threats to health, safety and fundamental rights fall into this category. They must comply with stringent requirements, including data quality, technical robustness, human oversight and transparency. Examples include AI systems used for biometric identification, recruitment, credit scoring, law enforcement, migration, education, health, transport, and essential public and private services.
- **Limited risk:** These AI systems require transparency measures to prevent manipulation and bias. They are subject to lighter regulatory requirements compared to high-risk systems.
- **Minimal risk:** These include AI systems that pose low risks and are subject to minimal regulatory requirements. Examples include AI-enabled video games or spam filters.

How do we harmonise all the requirements and standards of the EU AI Act for a global company?

Harmonising the requirements and standards for a global company is a complex task. However, many organisations are already familiar with the basics due to other regulatory requirements, which presents a significant opportunity to leverage existing frameworks to meet future needs. Achieving harmonisation requires a detailed understanding of emerging and existing laws and regulations in various jurisdictions worldwide. This involves not only knowing the specific requirements of the EU AI Act but also understanding the AI systems used throughout the enterprise and the data feeding into those systems. It is crucial to understand the demographics and potential imbalanced treatment of certain groups to ensure that AI does not amplify biases. Additionally, it is essential to involve leaders from across the organisation and third-party experts in AI and compliance. This collaborative approach ensures that all aspects of the AI systems and their compliance with the EU AI Act are thoroughly addressed. By leveraging the expertise of these stakeholders, organisations can create a comprehensive strategy that aligns with the act's standards and requirements.

What are the GDPR principles relevant to AI systems and, by extension, the EU AI Act?

The EU's General Data Protection Regulation (GDPR), which went into effect on May 25, 2018, regulates how personal data of EU residents is processed and transferred. The GDPR contains numerous principles that are relevant to the EU AI Act and includes several key concepts that ensure the protection of personal data when using AI systems. These principles are essential for compliance with both the GDPR and the EU AI Act, especially when AI systems process personal data. Here are the main GDPR principles that are relevant:

- **Lawfulness:** AI systems must process personal data in a lawful manner, adhering to the legal grounds for processing as outlined in the GDPR.
- **Fairness:** The processing of personal data by AI systems must be fair, ensuring that individuals are not subjected to unjust or discriminatory outcomes.
- **Transparency:** AI systems must operate transparently, providing clear and accessible information about how personal data is being processed, including the logic behind automated decision-making.
- **Purpose limitation:** Personal data must be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes.
- **Data minimisation:** AI systems should only process the minimum amount of personal data necessary to achieve their intended purpose.
- **Accuracy:** Personal data processed by AI systems must be accurate and kept up to date, with measures in place to rectify inaccuracies.
- **Storage limitation:** Personal data should be kept in a form that permits identification of individuals for no longer than is necessary for the purposes for which the data is processed.
- **Integrity and confidentiality:** AI systems must ensure the security of personal data, protecting it against unauthorised or unlawful processing and against accidental loss, destruction, or damage.

In the first years after GDPR implementation, 81% of respondents felt that they had little or no control over the customer data collected by organisations. (Pew Research)

In what ways does the GDPR overlap with EU AI Act?

While the GDPR focuses on personal data protection, the EU AI Act aims to address the broader risks associated with AI systems. However, there are areas of overlap. For example, both regulations emphasise the importance of transparency and human oversight in automated decision-making. Compliance with the GDPR does not automatically ensure compliance with the EU AI Act, and vice versa.

Considering both the EU AI Act and GDPR, what are the implications for businesses?

Organisations developing or using AI systems in the EU need to navigate both the GDPR and the EU AI Act. This needs to involve, among other actions:

- Ensuring AI systems that process personal data comply with GDPR principles
- Identifying the risk classification categories, as defined in the act, under which the AI system and/or specific processes fall
- Implementing the necessary technical and organisational measures to meet the requirements of both regulations
- Providing clear information to individuals about the use of AI systems and their rights under both regulations.

Organisations need to start aligning their AI practices with both the GDPR and the act, if they have not done so already.

In regard to compliance with the act, what are the notable challenges related to data collection and processing?

The act states the following: "Privacy and data governance means that AI systems are developed and used in accordance with privacy and data protection rules, while processing data that meets high standards in terms of quality and integrity. Transparency means that AI systems are developed and used in a way that allows appropriate traceability and explainability, while making humans aware that they communicate or interact with an AI system, as well as duly informing deployers of the capabilities and limitations of that AI system and affected persons about their rights."

In light of this and understanding that data curation and access is a critical component of any AI project, directly influencing an AI model's performance and output quality, key challenges include:

- **Data availability and relevance:** Ensuring the data is not only accurate but also pertinent to the specific AI task, such as text generation or image creation.
- **Bias and diversity:** Human biases can seep into data, leading to biased outputs from AI models. A diverse dataset is crucial for generating unbiased and complete content.
- **Data imbalance and completeness:** Balancing the dataset to avoid underrepresentation of certain groups or topics and ensuring it lacks no critical information needed for the AI task.

How do we harmonise all the requirements and standards of the EU AI act for a global company?

Under the act, as with many data protection laws, individuals can access, rectify, or halt the processing of their data (subject to certain exceptions). A significant challenge arises with generative AI when someone requests data deletion. The challenges include:

- Organisations may not have direct control over the training data used by third-party generative AI solutions
- Even if an AI system has been trained using an individual's data, it might not truly be able to "forget" it, as AI identifies and uses patterns from its training data, possibly retaining traces of personal data
- There is the potential for the system to reproduce remnants, patterns, or full components sourced from its training data
- If sensitive data makes its way into the training data, it may be necessary to retrain the entire model, which can be a complex and resource-intensive process.

This is an evolving area of the law, and regulatory bodies like the Federal Trade Commission in the United States have used algorithmic disgorgement as a remedy in some cases. All providers of general-purpose AI (GPAI) models must:

- Draw up technical documentation, including training and testing processes and evaluation results
- Draw up information and documentation to supply to downstream providers that intend to integrate the GPAI model into their own AI system in order that the latter understands capabilities and limitations and is enabled to comply
- Establish a policy to respect the copyright directive
- Publish a sufficiently detailed summary about the content used for training the GPAI model.

What are the implications for GPAI models that pose systemic risks?

If a GPAI can potentially create systemic risks, providers must additionally:

- Perform model evaluations, including conducting and documenting adversarial testing to identify and mitigate systemic risk
- Assess and mitigate possible systemic risks, including their sources
- Track, document and report serious incidents and possible corrective measures to the European AI Office and relevant national competent authorities without undue delay
- Ensure an adequate level of cybersecurity protection.

Under the act, who is responsible for AI accuracy?

Unlike in other countries and regions where responsibility for AI accuracy can vary depending on the jurisdiction and specific AI regulation, the EU AI Act assigns the broadest obligations to the provider of the AI system.

What are the best practices for mitigating risks associated with AI adoption in regard to the EU AI Act?

Best practices include:

- Maintaining an inventory of AI use cases and systems (including third parties), classifying these based on level of risk as defined earlier (unacceptable, high, limited and minimal)
- Recognising and comprehending all obligations under data protection laws
- Conducting thorough impact assessments
- Evaluating and monitoring transparency
- Integrating privacy and ethics by design
- Forging clear contractual terms
- Regularly evaluating AI systems for precision and biases.

[1] The EU AI Act defines a "provider" as a natural or legal person, public authority, agency, or other body that develops an AI system or a general-purpose AI model. This definition also includes entities that have an AI system or a general-purpose AI model developed and place it on the market or put the AI system into service under their own name or trademark, whether for payment or free of charge.

Call to Action

To ensure compliance with the EU AI Act and to manage AI systems effectively within your organisation, it is crucial to take the following actions:

1. **Understand the EU AI Act:** Familiarise yourself with the comprehensive regulatory framework designed to govern the development, deployment and use of AI across the EU. The act defines AI broadly, encompassing any system that can operate with elements of autonomy and achieve a given set of objectives using machine learning, logic-based and knowledge-based approaches.

2. **Classify AI systems:** Determine which risk group your AI implementation falls into by considering the potential impact on human rights, safety and fundamental values. The act uses a risk classification approach with four distinct groups: unacceptable risk, high risk, limited risk and minimal risk.

3. **Ensure GDPR compliance:** Align your AI practices with the GDPR principles relevant to AI systems, such as lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, and integrity and confidentiality.

4. **Implement best practices:** Adopt best practices for mitigating risks associated with AI adoption, including conducting thorough impact assessments, maintaining transparency and explainability, integrating privacy and ethics by design, and regularly evaluating AI systems for precision and biases.

5. **Stay informed:** Keep up-to-date with the latest perspectives on AI and the EU AI Act by visiting resources such as Protiviti's AI Intelligence Hub.

By taking these steps, your organisation can effectively navigate the complexities of the EU AI Act and ensure the responsible use of AI systems.

About Protiviti

Protiviti (www.protiviti.com) is a global consulting firm that delivers deep expertise, objective insights, a tailored approach and unparalleled collaboration to help leaders confidently face the future. Protiviti and its independent and locally owned member firms provide clients with consulting and managed solutions in finance, technology, operations, data, digital, legal, HR, risk and internal audit through a network of more than 90 offices in over 25 countries.

Named to the Fortune 100 Best Companies to Work For® list for the 10th consecutive year, Protiviti has served more than 80 percent of Fortune 100 and nearly 80 percent of Fortune 500 companies. The firm also works with government agencies and smaller, growing companies, including those looking to go public. Protiviti is a wholly owned subsidiary of Robert Half Inc. (NYSE: RHI).

About the author

Christine Livingston
Managing Director, Protiviti

Christine is responsible for Protiviti's AI/ML capabilities and solutions. With over a decade of experience in AI/ML deployment, she has delivered hundreds of successful solutions, including many first-in-class AI-enabled applications. She has helped several Fortune 500 clients develop practical strategies for value-driven enterprise adoption of artificial intelligence, including the creation of capability-based AI-enabled technology road maps. She focuses on identifying emerging technology opportunities, incorporating AI/ML capabilities into enterprise solutions, and delivering tangible business value and outcomes through AI.

Contact Christine at christine.livingston@protiviti.com.

Face the Future with Confidence®

protiviti®
Global Business Consulting