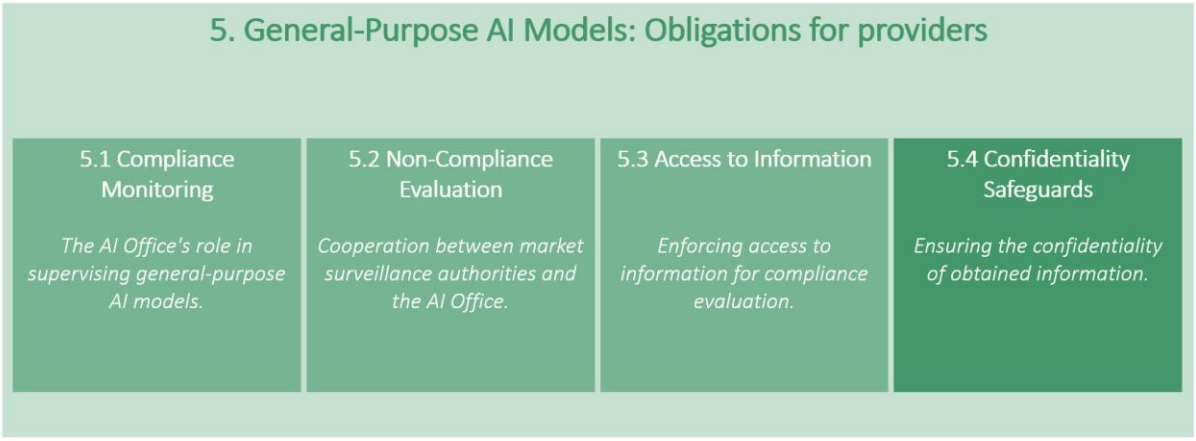
AI & Partners
Amsterdam – London - Singapore

Safeguarding Secrets: Confidentiality Measures for AI Providers Under the EU AI Act

Co-authored with Patricia Shaw, **Beyond Reach Consulting Limited**, *CEO and Founder*

28 October 2024



Introduction

In the realm of artificial intelligence ("AI"), the advent of the European Union ("EU") AI Act (the "EU AI Act") marks a significant stride towards regulating the risk of AI technologies. The main emphasis on general-purpose AI models concerns transparency for the purposes of accountability and by so doing, this legislation attempts to balance fostering innovation with ensuring the confidentiality of sensitive information. Providers of general-purpose AI models are mandated to navigate the complex landscape of compliance, where they must disclose technical documentation and operational details of their AI models to regulatory bodies, including the AI Office.

The EU legislators were acutely aware of the importance of protecting trade secrets, confidential business information, and intellectual property rights whilst making the EU AI Act. For this reason, explicitly mandates that while providers are required to share information to demonstrate compliance, this does not extend to disclosing sensitive proprietary information that could compromise competitive advantage. Furthermore, the EU AI Act establishes a framework where the confidentiality of the information obtained in regulatory processes is safeguarded, ensuring that intellectual property rights and trade secrets are protected.

This delicate equilibrium highlights the EU's nuanced approach: it aims to promote transparency and accountability in AI development and deployment while rigorously protecting the confidentiality of critical business information. This ensures that the AI ecosystem within the EU remains both innovative and secure, fostering trust among stakeholders and the public alike.

Overview of the EU AI Act

The EU AI Act is a groundbreaking piece of legislation designed to regulate the risks associated with AI technologies across its member states and/or where it is used in respect of EU citizens to preserve EU values. Its primary aim is to foster an environment where AI can be developed and deployed in a manner that is safe, compliant with regulations, and conducive to innovation by being trustworthy (both for use and adoption). The EU AI Act sets forth a comprehensive framework that addresses the lifecycle of AI systems and the complexity of the value chain, from development to deployment, ensuring that these technologies enhance rather than endanger health, safety, and fundamental rights, in that also recognising the important role of environmental protection to all of the above.

A significant focus of the EU AI Act is on general-purpose AI models, which are versatile and capable of being used across a variety of applications. Providers of these models are required to adhere to specific obligations, including the maintenance and disclosure of technical documentation that details the model's training, testing, and evaluation processes. This ensures that the AI models are transparent and their operations can be scrutinized for compliance and safety.

Moreover, the EU AI Act underscores the importance of confidentiality safeguards. It mandates that while providers must share certain information to demonstrate compliance, this does not extend to disclosing proprietary or sensitive information that could compromise trade secrets or intellectual property. This delicate balance between transparency and confidentiality is crucial for maintaining competitive advantages while ensuring that AI technologies are developed and used responsibly within the EU. Through these measures, the EU AI Act aims to create a trustworthy AI ecosystem that supports innovation while protecting the interests of all stakeholders involved.

Confidentiality Obligations for Providers

Under the EU AI Act, providers of general-purpose AI models are entrusted with significant responsibilities to ensure the confidentiality of information obtained during the compliance process. Article 53 outlines the obligations for these providers to meticulously document the technical aspects of their AI models, including training, testing processes, and evaluation results. This documentation is crucial for demonstrating compliance with the Act's requirements and must be made available to the AI Office and national competent authorities upon request.

Article 54 firmly grounds providers who are based outside the EU with a presence in the EU through the appointment of an authorised representative to ensure accessibility and accountability of the general-purpose AI model providers through their proxy to the AI Office.

Additionally, Article 55 imposes further obligations on providers of general-purpose AI models identified with systemic risks. Whilst systemic risks in and of themselves are left undefined – this is something to be identified by the EU AI Office- it is for these providers of GPAI models trained above the 10-25 FLOPs threshold to conduct and document model evaluations, assess and mitigate such systemic risks, and report serious incidents to the AI Office and competent authorities. The crucial part for any general-purpose AI model provider whose model does present a systemic risk will be to seek conformity with the soon to be established codes of practice made in collaboration with general-purpose AI model providers, national competent authorities, civil society, independent experts, and the EU AI Office. The timeline for having such codes of practice ready however is short: 9 months after the entry into force date, leaving it to be in place by 1st May 2025.

The Act explicitly mandates that any information or documentation acquired pursuant to Articles 53 and 55, including trade secrets, must be treated in accordance with the confidentiality obligations set out in Article 78. Whether it is the Commission, market surveillance authorities, notified bodies or any other person involved in the application of the EU AI Act, confidentiality is to be observed and is of paramount importance.

This ensures that while providers are required to share critical information to demonstrate compliance, the integrity and confidentiality of their proprietary and sensitive information are safeguarded. This framework underlines the EU's commitment to fostering an environment where innovation can thrive, supported by a foundation of trust and security.

Mechanisms for Protecting Confidential Information

The EU AI Act establishes stringent mechanisms to protect the confidentiality of sensitive information, including trade secrets, within the realm of AI development and deployment. Article 78 of the Act delineates the confidentiality obligations that the AI Office, national competent authorities, and notified bodies must adhere to. These entities are required to respect the confidentiality of information and data obtained in carrying out their tasks, ensuring the protection of intellectual property rights, confidential business information, or trade secrets. This provision is critical in fostering a trustful environment for AI providers, as it guarantees that the disclosure of technical documentation and operational details to regulatory bodies does not compromise their competitive edge or expose proprietary technologies.

The Act specifies that these confidentiality obligations cover a broad spectrum, including the effective implementation of the regulation, public and national security interests, and the integrity of criminal and administrative proceedings. It seeks to preserve public and national security interests in this way because of the nature and Importance being ascribed to the potential harm that systemic risks could have on EU values.

Furthermore, Article 78 mandates that the authorities involved only request data strictly necessary for assessing the risk posed by AI systems and for exercising their powers in accordance with the regulation. Adequate and effective cybersecurity measures must be put in place to protect the security and confidentiality of the information obtained, and such data must be deleted once it is no longer needed

for the purpose for which it was obtained. Through these measures, the EU AI Act ensures a balanced approach, promoting transparency and accountability in AI technologies while safeguarding the vital interests of AI providers and the health, safety and fundamental rights of EU citizens.

Challenges in Maintaining Confidentiality

Balancing regulatory compliance with the protection of confidential business information under the EU AI Act presents a complex challenge for AI providers. The Act mandates the sharing of technical documentation and incident reports to ensure AI systems' safety and compliance. However, this requirement raises concerns about safeguarding sensitive data, including trade secrets and intellectual property rights, which are crucial for maintaining competitive advantages.

The EU AI Act addresses these concerns by stipulating confidentiality obligations for the Commission, market surveillance authorities, notified bodies, and other entities involved in the application of the regulation. These entities are required to respect the confidentiality of information and data obtained in carrying out their tasks, protecting intellectual property rights, confidential business information, or trade secrets. Despite these safeguards, the potential for conflicts arises when the need for transparency and public safety intersects with the imperative to protect sensitive business information.

The Act mandates that authorities request only data strictly necessary for assessing the risk posed by AI systems and for exercising their powers, coupled with adequate cybersecurity measures to protect the security and confidentiality of the information obtained. While these provisions aim to mitigate risks, the practical implementation of these measures requires careful navigation to ensure that compliance efforts do not inadvertently compromise confidentiality.

Best Practices for AI Providers

For AI providers, navigating the EU AI Act's requirements while protecting confidential information necessitates a strategic approach. The Act mandates providers to maintain detailed technical documentation and make certain information available, ensuring compliance without compromising intellectual property or trade secrets. Here are best practices for AI providers:

1. Selective Documentation: Providers should carefully curate the information included in technical documentation and summaries. While being comprehensive in demonstrating compliance, it may be necessary to omit details that could reveal proprietary algorithms or data sets without compromising the meaningful explanation required. This aligns with the Act's allowance for protecting intellectual property and trade secrets.
2. Confidentiality Agreements: Before sharing sensitive information with regulatory bodies or third parties, AI providers should secure non-disclosure and confidentiality agreements. This measure is supported by the Act's emphasis on respecting the confidentiality of information obtained in regulatory processes.
3. Utilise Templates: The Act encourages the use of templates provided by the AI Office for documenting training data summaries. Providers should leverage these templates to ensure that the summaries are detailed yet do not disclose proprietary sensitive information.
4. Engage in Dialogue: Providers are encouraged to engage in structured dialogues with the AI Office or national competent authorities. This proactive communication can help clarify the extent of information required, including an understanding of what is meaningful transparency relevant for the relevant bodies involved, ensuring that providers meet compliance obligations both in the letter and in the spirit of the law without unnecessary exposure of confidential data.

By adhering to these practices, AI providers can effectively manage and protect their confidential information while fulfilling their obligations under the EU AI Act, ensuring a balance between regulatory compliance and the safeguarding of trade secrets and intellectual property.

Conclusion

The EU AI Act embeds confidentiality safeguards, underscoring their pivotal role in the AI regulatory framework. For providers of general-purpose AI models, these measures are instrumental in balancing the act of compliance with the imperative to protect sensitive business information, including trade secrets and intellectual property. The Act mandates the maintenance and disclosure of technical documentation, ensuring that AI systems are transparent, safe, and aligned with codes of practice (and in time with harmonised EU standards), while simultaneously safeguarding proprietary information.

Moreover, the confidentiality obligations outlined in Article 78 fortify this balance, ensuring that all parties involved in the application of the regulation respect the confidentiality of information obtained during their tasks, thereby protecting intellectual property rights and confidential business information. This framework not only fosters a trustworthy and secure AI ecosystem within the EU but also encourages innovation by ensuring that providers can share necessary information without risking their competitive edge.

The EU AI Act's confidentiality measures are crucial in promoting a responsible, transparent, and secure development and deployment of AI technologies, contributing to a sustainable and trustworthy AI ecosystem in the EU.

Glossary

Act or EU AI Act: European Union Artificial Intelligence Act

AI: Artificial Intelligence

Board: European Union Artificial Intelligence Board

EU: European Union

SME: Small and Medium-Sized Enterprise

How can we help?

AI & Partners
Amsterdam – London - Singapore

AI & Partners – 'AI That You Can Trust'

At AI & Partners, we're here to help you navigate the complexities of the EU AI Act, so you can focus on what matters—using AI to grow your business. We specialize in guiding companies through compliance with tailored solutions that fit your needs. Why us? Because we combine deep AI expertise with practical, actionable strategies to ensure you stay compliant and responsible, without losing sight of your goals. With our support, you get AI you can trust—safe, accountable, and aligned with the law.

To find out how we can help you, email contact@ai-and-partners.com or visit https://www.ai-and-partners.com.