SOŁTYSIŃSKI KAWECKI SZLĘZAK

AI GOVERNANCE

A Framework for Responsible and Compliant Artificial Intelligence

September 2025

Agata Szeliga
Partner,
attorney-at-law

Anna Tujakowska
Senior Counsel,
attorney-at-law

Sylwia Macura-Targosz
Senior Associate,
attorney-at-law

SOŁTYSIŃSKI KAWECKI SZLĘZAK

Table of contents

SOŁTYSIŃSKI KAWECKI SZLĘZAK

Executive Summary

As artificial intelligence (AI) becomes increasingly integrated into business processes, organizations face growing pressure to leverage its benefits while managing its various risks, such as legal, ethical, operational and reputational. Effective AI governance is not the responsibility of a single team or function. It requires cross-functional collaboration that includes legal and compliance (to ensure regulatory alignment), data and IT personnel (to manage infrastructure and model lifecycle), business and operations leaders (to align AI use with strategic goals) and for more complex or high risk cases - ethics advisory (to embed responsible AI principles). Governance is not a barrier to innovation, it is rather a framework that enables AI to be used safely, fairly, and sustainably.

This white paper offers a comprehensive overview of how to responsibly govern AI systems, with particular emphasis on compliance with the **EU Artificial Intelligence Act (AI Act)**, the world's first comprehensive legal framework for AI.

It also outlines the evolving risk landscape that organizations must navigate as they scale their use of AI. These risks include:

- Ethical, social, and environmental risks – such as algorithmic bias, lack of transparency, insufficient human oversight, and the growing environmental footprint of generative AI systems.
- Operational risks – including unpredictable model behavior, hallucinations, data quality issues, and ineffective integration into business processes.
- Reputational risks – resulting from stakeholder distrust due to errors, discrimination, or mismanaged AI deployment.
- Security and privacy risks – encompassing cyber threats, data breaches, and unintended information disclosure.

To mitigate these risks and ensure AI is used responsibly, in this white paper we propose a set of governance recommendations, including:

- Ensuring transparency through clear communication about AI systems' purpose, capabilities, and limitations.
- Promoting AI literacy via targeted training and well-defined responsibilities across functions.
- Strengthening security and resilience by implementing monitoring processes, incident response protocols, and robust technical safeguards.
- Maintaining meaningful human oversight, particularly for high-impact decisions.
- Appointing an AI Champion to lead responsible deployment, oversee risk assessments, and foster a safe environment for experimentation.

Lastly, this white paper acknowledges the key implementation challenges facing organizations: overcoming internal resistance, balancing innovation with regulatory compliance, managing technical complexity (such as explainability and auditability), and navigating a rapidly evolving and often fragmented regulatory landscape.

We hope this white paper will serve as a practical guide for organizations seeking to build robust, forward-looking AI governance programs that not only meet legal requirements but also support responsible innovation, public trust, and long-term strategic value.

In order to help navigate these complex issues we also provide useful templates and recommendations which were made available by Microsoft for its generative AI tools, including Azure Open AI or M365 Copilot. We hope such practical guidelines will be a great support.

1. Introduction

Artificial Intelligence (AI) is reshaping the way organizations operate, offering unprecedented opportunities to automate, predict, and personalize at scale. However, with this transformative potential comes a corresponding need for structured oversight. AI governance refers to the system of policies, processes, and accountability mechanisms that ensure the ethical, lawful, and responsible development of AI systems.

Our intention is to provide a practical, high-level framework for organizations, irrespective of its size, seeking to build or improve their AI governance capabilities. It outlines key principles, risk areas, and implementation steps. It should be noted that in regulated sectors – such as healthcare, finance, or legal service – AI governance may be subject to additional obligations arising from sector-specific regulations and supervisory frameworks. Organizations operating in these areas should ensure compliance not only with general AI governance standards, but also with the specific regulatory requirements relevant to their sector.

It is intended both for organization deploying AI for the first time and those scaling up its AI use.

This white paper reflects the legal landscape as of September 8, 2025.

> *We are committed to supporting customers in Europe and around the world in implementing their own AI systems responsibly, including by developing responsible AI programs for our partner ecosystem. ... At Microsoft, all of our AI systems undergo a responsible AI process that includes reviews by a multidisciplinary team of experts to help us understand potential harms and mitigate these harms. Examples of mitigations include refining the dataset used to train models, deploying filters to limit the generation of harmful content, integrating techniques like query blocking on sensitive topics that helps to prevent misuse by bad actors, and applying technology that can return more helpful, representative, and diverse results.*
>
> **Anthony Cook** Corporate Vice President Microsoft Blueprint for Austria, AI for Austria

2. Legal framework

A comprehensive AI governance strategy should address the regulatory landscape, ensuring that legal compliance is integrated throughout the entire lifecycle of AI development and deployment rather than being treated in isolation.

2.1 AI Act

The key legal regulations which should be taken into account by the entities operating in Poland is AI Act[1] the world's first comprehensive legal framework for artificial intelligence. The AI Act as an EU regulation is directly applicable in all EU member states. The AI Act entered into force on 1 August 2024, but the application of its provisions has been phased in over several years.

When it becomes applicable?

6 months after entry into force
2 February 2025

- General provisions
- *AI literacy* regulations (Article 4)
- Provisions on so-called prohibited practices (Article 5)

12 months after entry into force
2 August 2025

- General-purpose AI regulations (Chapter V)
- Provisions on authorities overseeing compliance with the AI Act (Chapter III, Section 4)
- Management regulations (Chapter VII)
- Confidentiality provisions (Article 78)
- Provisions on penalties for violations of the AI Act (Chapter XII)

24 months after entry into force
2 August 2026

- **Almost full** application of AI Act, including high-risk AI systems: transparency obligations

36 months after entry into force
2 August 2027

- Certain **obligations for high-risk AI systems** that are safety components in products regulated by EU product safety regulations (e.g. civil aviation and medical devices)

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council on laying down harmonized rules on artificial intelligence and amending Regulations (EC) No. 300/2008, (EU) No. 167/2013, (EU) No. 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828.

SOŁTYSIŃSKI KAWECKI SZLĘZAK

Who must comply with the AI Act?

The AI Act applies to a broad range of entities across the AI value chain (Article 2 (1)), including: providers of AI systems, importers, distributors, authorized representatives, product manufactures, deployers. Most of the responsibilities lie with AI system providers.

PROVIDERS (DEVELOPERS)

any entities that develop an AI system/model or has it developed and place it on the market or put it into service under their own name or trademark, regardless of whether they are established in the EU (Article 3 (3)). It may be also the entity modifying AI model if the modification leads to a significant change in the model's generality, capabilities, or systemic risk. Such entities may be, for instance, a software company that develops a custom AI model for its clients or a research institute that develops an AI system and makes it available on the Internet.

DEPLOYERS (USERS)

any entities that use an AI system during a professional activity are subject to governance duties, including ensuring proper use and cooperating with national authorities (Article 3 (4)). Such entities may be, for instance, an e-commerce website owner using AI to personalize offers or entrepreneurs providing AI-based chatbots to clients to obtain information about the services.

It should be born in mind that deployer, who changes the purpose of an "ordinary" AI system so that it becomes a high-risk AI system, becomes a provider of that system under the AI Act.

IMPORTERS AND DISTRIBUTORS

any entities that import, distribute, or make AI systems available on the EU market must ensure that these systems meet applicable requirements and do not alter their compliance (Articles 3 (6)-(7)).

AUTHORIZED REPRESENTATIVES

any entities that perform and conduct on behalf of non-EU-based providers the obligations and procedures established by the AI Act authorities (Article 3 (5)).

PRODUCT MANUFACTURERS

any manufacturers of products subject to existing sectoral legislation (e.g., medical devices, machinery) if an AI system is integrated into such product.

Below, we focus mainly on the obligations of Deployers (Users), as most organizations fit in this role.

What is covered?

The AI Act applies to placing on the market, putting into service, and the use of **AI systems**, especially those deemed to pose a high risk to the functioning of society (e.g. AI systems applied in education or employment, or using biometrics) and **general-purpose AI models (GPAI) and systems** (such as ChatGPT). According to the European Commission's Guidelines, the AI model is a GPAI if its training compute is greater than 10^{23} FLOP and it can generate language. More details which models are GPAI and which are not and when modification is material may be found in Guidelines on the scope of the obligations for general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act).[2]

The AI Act does not apply, among other things, to areas outside of the scope of EU law, such as national security, and to systems used exclusively for military or defence purposes or for the sole purpose of scientific research and development. The AI Act does not apply also to private, non-commercial use of AI systems (e.g., hobby projects at home).

What is the AI System? AI system is **a machine-based system** that is designed to operate with **different 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** (Article 3 (1)).

AI systems will therefore be, for instance, recruitment systems that analyze CVs and/or support hiring decisions, chatbots that serve customers and provide advice or AI systems that analyze medical data and supports diagnosis.

To clarify AI system concept, the European Commission ("EC") has approved guidelines for the definition of an AI system. In particular, the guidelines indicate cases where a specific IT solution will not constitute an AI system (e.g., systems for improving mathematical optimization, simple prediction systems, systems based on classical heuristics). The EC guidelines are not mandatory and will be updated when needed, based on practical experience, new questions, and emerging use cases.

Thus, an AI system will not be a simple calculator or spreadsheet (e.g., Excel without AI functions), a contact form on a website, systems based solely on "if-then" rules without machine learning. Moreover, the AI Act obligations pertaining to GPAI other than those with systemic risks do not apply to open source GPAI.

Renata Zalewska
Senior Commercial Attorney at Microsoft

Microsoft is committed to building products and solutions that comply with the EU AI Act and helping our customers use AI compliantly. On Microsoft Blogs, we publish a lot of information on Microsoft efforts in this area. As Natasha Crampton, our Chief Responsible AI Officer wrote on her blog, Microsoft is "ready to help our customers do two things at once: innovate with AI and comply with the EU AI Act. We are building our products and services to comply with our obligations under the EU AI Act and working with our customers to help them deploy and use the technology compliantly.".

[2] Guidelines on the scope of obligations for providers of general-purpose AI models under the AI Act | Shaping Europe's digital future

Guidelines for the definition of an AI system

The concept of "AI system" should be also distinguished from the concept of an "AI model". The AI Act marks a "general-purpose AI model" ("**GPAI model**") which is an AI model capable of performing a wide range of different tasks (Article 3 (63)). As an example of GPAI models are indicated large generative AI models (ChatGPT, Claude or Gemini).

The AI Act's rules on GPAI models came into force on August 2, 2025. A Code of Practice for the GPAI Model (introduced in Article 56), that is a set of guidelines for compliance with the AI Act, was published on July 10, 2025. It is a crucial tool for ensuring compliance with the AI Act's obligations, especially in the interim period between when GPAI model provider's obligations come into effect (August 2025) and the adoption of standards (August 2027 or later).

A Code of Practice for the GPAI Model

The Code is divided into three chapters. The first, the Transparency chapter, provides all GPAI model providers with a framework to demonstrate compliance with their obligations under Article 53 (1)(a) and (b) of the AI Act. This includes requirements such as preparing and maintaining up-to-date model documentation and sharing relevant information with AI system providers who intend to integrate the GPAI model into their systems. To support this, the Code includes a Model Documentation Form.

The Copyright chapter outlines how GPAI model providers can demonstrate compliance with Article 53(1)(c) of the AI Act. It emphasizes the need for transparent and up-to-date copyright policies, lawful data collection practices, and respect for rights reservations expressed via machine-readable protocols. Providers are also encouraged to implement safeguards against copyright-infringing outputs and establish accessible channels for rightsholders to raise complaints.

The Chapter on Safety and Security is only relevant to the small number of providers of the most advanced models, those that are subject to the AI Act's obligations for providers of GPAI models with systemic risk under Article 55 AI Act.

Guidelines on the scope of obligations for providers of the GPAI models under the AI Act

On July 18, 2025, the EC published Guidelines on the scope of the obligation for the GPAI models established by AI Act. The guidelines clarify key concepts underlying the provisions in the AI Act on GPAI models, e.g., give detailed explanations on questions such as 'what is a GPAI model', 'which entities are providers', and 'which actions constitute a placing on the market'. The guidelines also lay out how the AI Office will provide support to facilitate compliance.

How does the AI Act divide AI systems?

The AI Act is built on a **risk-based approach which AI systems may pose to AI users**, which reflects a growing international consensus on how AI should be governed. Instead of applying uniform rules to all AI systems, this model recognizes that the risks vary depending on the system's context, purpose, and level of autonomy. It ensures that regulatory requirements are proportionate to the potential harm, safeguarding fundamental rights while enabling low-risk innovation. At the same time, it provides organizations with clarity and flexibility by aligning obligations with risk levels and allowing the framework to adapt as technologies evolve.

The AI Act categorizes the AI system into tiers such as prohibited, high-risk, limited-risk, and minimal-risk.

UNACCEPTABLE RISK

Prohibited AI Practices: Systems deemed to pose extreme risk are prohibited outright. Examples include AI that deploys subliminal techniques to manipulate behavior, exploits vulnerabilities of specific groups (like children or disabled persons), "social scoring" of individuals by governments, and certain uses of biometric surveillance. These practices are prohibited due to their threat to fundamental rights and safety.

The EC has approved detailed guidelines on prohibited AI practices. The guidelines are not binding.

Guidelines on prohibited AI practices

HIGH-RISK

Highly Regulated: AI systems with significant implications for safety or rights (e.g. in critical infrastructure, education, employment, credit scoring, law enforcement, etc.) are permitted only under strict conditions. High-risk examples (detailed in Annex III of the AI Act) include AI for hiring or worker management, algorithms for university admissions, creditworthiness assessments, medical devices with AI, and certain law enforcement or border control tools. These systems must comply with extensive requirements.

LIMITED-RISK

Transparency Obligations: Some AI systems aren't high risk but still require precautions. For instance, chatbots or deepfakes must be clearly identified as AI-generated so users know they are interacting with AI. Providers of such AI must ensure transparency (e.g. disclosure of AI involvement) but face lighter rules than high-risk AI.

MINIMAL OR LOW-RISK

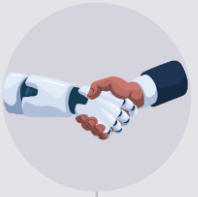
Minimal Regulation: The majority of AI applications (such as AI in spam filters or video games) fall into this category and are largely unregulated by the AI Act. While general laws (like consumer protection) still apply, either the AI Act imposes no new requirements on low-risk AI beyond existing legislation or it introduces limited transparency obligations.

SOŁTYSIŃSKI KAWECKI SZLĘZAK

Examples of High-Risk AI Systems

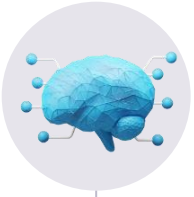
BIOMETRIC IDENTIFICATION AND CATEGORIZATION

- Real-time and post-remote biometric identification systems used in public spaces (e.g., facial recognition by law enforcement)
- Biometric categorization systems that infer sensitive attributes like race, religion, or political views



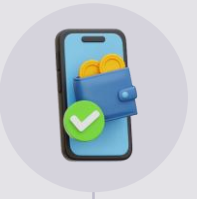
EMPLOYMENT, WORKERS MANAGEMENT, AND ACCESS TO SELF-EMPLOYMENT

- AI used in recruitment (e.g., CV screening, video interview analysis)
- Systems monitoring employee's behavior or performance (e.g., productivity tracking software)



EDUCATION AND VOCATIONAL TRAINING

- Tools that assess learning outcomes or student performance in high-stakes contexts



ACCESS TO ESSENTIAL PRIVATE SERVICES

- AI systems used to evaluate creditworthiness
- AI in insurance risk assessment or pricing

Not all AI systems are subject to the same level of regulation. The specific obligations depend on how the system is classified under the AI Act's risk-based approach:

RISK LEVEL	APPLIES TO NON-EU SYSTEMS?	KEY OBLIGATIONS
 PROHIBITED	❌ Not allowed in the EU	Cannot be placed on the EU market or used in the EU (e.g., social scoring)
 HIGH-RISK	✅ Yes	Full compliance required: risk management, data governance, transparency, human oversight, conformity assessment
 LIMITED-RISK	✅ Yes	Transparency obligations (e.g., inform users they are interacting with AI)
 MINIMAL-RISK	✅ Yes	No binding rules; encouraged to follow voluntary codes of conduct

Does the AI Act have an extraterritorial scope?

The AI Act applies not only to entities based in the EU, but also to companies and individuals located outside the EU, including those in the UK, U.S., or other third countries if they:

- place an AI system GPAI models on the EU market or grant them for use on the EU market, or
- the output of an AI system affects individuals located in the EU, even if the system operates entirely outside EU borders.

This means that, for instance, a UK-based provider must comply with the AI Act if their AI system is offered to EU users or delivers outputs (e.g., decisions, recommendations, scores) that impact people in the EU.

What are the key obligations under the AI Act?

The AI Act aims to ensure that AI used in the EU is safe, transparent, and respects fundamental rights. Most responsibilities are imposed on the AI system and GPAI models providers, especially for high-risk systems.

[Providers]

The AI Act provides for certain obligations that apply to all AI systems, regardless of the level of risk such as ensuring AI literacy (Article 4). This includes considering technical knowledge, experience, education, and training of the personnel, as well as the context of use.

Providers of the limited-risk AI system have certain transparency obligations, such as: informing users that they are interacting with an AI when that is not obvious and ensuring that when the AI system is used to generate graphics, audio, video and text material, the AI system must automatically label generated graphics, audio, video, and text as artificially created or manipulated (Article 50).

The catalog of obligations imposed under the AI Act on providers of high-risk AI systems is extensive and includes in particular:

- providing clear user instructions to deployers, explaining the AI's intended use and how to operate it in a compliant manner (Article 13);
- design of the AI system in such a way as to allow effective human oversight (Article 14),
- ensuring that high-risk AI meet standards of accuracy (minimizing errors), be robust against faults or manipulation, and be cybersecure against attacks (Article 15)
- ensuring that the high-risk AI system bears the provider's name, registered trade name or trademark, and contact detail (Article 16 (b));
- putting a quality management system (QMS) in place (Article 17). This QMS should cover policies and procedures to maintain compliance, like how ISO standards ensure product consistency;
- retention of documentation (including technical documentation, QMS documentation, documentation related to conformity assessment, including those related to approval of changes by a notified body, and the EU declaration of conformity) (Article 18);
- taking corrective action when non-compliance with the AI Act is found, informing other operators down the value chain (Article 20);
- registration of high-risk AI systems in an EU database before deployment for oversight purposes (Article 49).

Providers of general-purpose AI models have such responsibilities as: publishing summaries of their training data, providing instructions for use, and ensuring compliance with EU copyrights law (Article 53). If the GPAI model is deemed to pose "systemic risk" (high-impact, foundational models), providers will face additional requirements such as conducting risk evaluations, adversarial testing, ensuring cybersecurity, and reporting major incidents (Article 55).

[Deployers]

The AI Act also imposes specific obligations on **deployers**, regardless of their level of risk of AI systems they use. For instance, Article 4 establishes a responsibility to ensure AI literacy (Article 4).

In the case of the limited-risk AI systems, deployers must ensure transparency by informing users, in certain cases, that they are interacting with an AI or when the AI system generates or modifies content such as deepfake (Article 50).

Deployers are obliged to implement technical and organizational measures as per the user manual, assign human oversight to individuals who have the necessary competence, training and authority (Article 26) and conduct, in certain cases, a fundamental rights impact assessment of the AI system (Articles 27) – in case they deploy high-risk AI systems. Deployers must also report any serious incidents to providers (Article 73).

If a deployer changes the purpose of an "ordinary" AI system so that it becomes a high-risk AI system, it will also become a provider of that system under the AI Act, making the obligations of a provider of a high-risk AI system applicable to it (Article 25 (1) (3)). Moreover, the deployer may become the providers of GPAI models if they substantially modify it, i.e. if the modification leads to a significant change in the model's generality, capabilities or systemic risk. As further stated in the Commission's Guidelines on the scope of the obligations for general purpose AI models, an indicative criterion for when the modifier is regarded to be provider of the GPAI model is that the training compute used for modification is greater than a third of the training compute of the original model.

Polish legislation ensuring implementation of AI Act

Polish draft law on artificial intelligence systems aims to ensure AI implementation and supervision at the Polish level.

The draft provides for, among others, the establishment of the Commission for the Development and Security of Artificial Intelligence (a collegial body to oversee and monitor the AI systems market), the control of compliance with the provisions of the AI Act by the Commission (mainly remote control) and rules for imposing administrative penalties for violations of the AI Act.

2.2. Personal data protection regulations

General regulations on the processing of personal data, including the provisions of the GDPR[3] or sector regulations, apply to using AI tools[4]. The scope of the application of these regulations will depend on whether the entity engages in creating and training the AI model (or adjusting it – fine tuning) using information containing personal data or is a user of the AI system.

In the case of construction and training, the entity will have more responsibilities, including, first, properly identifying the legal basis for processing the personal data contained in the training data. A detailed explanation of how the GDPR is applied at this stage can be found in the Opinion of the European Data Protection Board on certain data protection aspects of processing personal data in the context of AI models[5].

In case the entity is a user of the AI system, personal data processing can potentially occur at three levels:

- **processing of the personal data contained in the input content (prompt)**

  If the entity wants to process personal data in the input content, it should, as the controller or a person authorized by the controller, verify, first and foremost, that it has an appropriate legal basis for such processing and what the purpose of the processing is. It should also verify compliance with other requirements for the protection of personal data, such as: whether data will be transferred outside the European Economic Area (EEA), whether it is necessary to enter into a data processing agreement, whether a Data Protection Impact Assessment (DPIA) is required, whether the realization of data subject's rights is ensured, etc.

- **processing of the personal data contained in the output content (output)**

  It should be considered that personal data might be inaccurate what could impact decisions based on it. Failure to do so with "limited confidence" in the personal data generated, may mean that the data processing does not comply with the principle of fairness. Uncritical reliance on such data may also be incompatible with the principle of data minimization as personal data, including applications, must be relevant and appropriate to the purpose.[6]

  The primary concern is whether AI-generated data can be associated with an identifiable individual. In such cases, the data can be considered personal data, even if it is inaccurate. EDPB's opinion confirms that this will be particularly true if AI is to provide conclusions (e.g., personal data) about individuals whose personal information was used for training.[7]

- **processing of the user's data when using the AI model**

  AI may also collect personal data of the user including data such as, name, surname, username, contact information; and in addition, input content, output content, and so-called user engagement data, e.g. pseudonymized identifiers randomly assigned to the user, usage time, activity history, encryption information, error information. An AI system provider usually processes user data as a processing entity, but may also be a data controller if, for example, it processes data to improve the AI system's performance, conducts research, or detect abuse.

[3] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

[4] For more information on the GDPR compliance for AI see, for example, here: Introducing Our New Whitepaper: GDPR & Generative AI – A Guide for Customers | Microsoft Community Hub.

[5] European Data Protection Board Opinion 28/2024 on Certain Data Protection Aspects of the Processing of Personal Data in the Context of Artificial Intelligence Models, 17 December 2024 Opinion available here: https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en.

[6] Information Commissioner's Office publication: How to use AI and personal data appropriately and lawfully, p. 9, https://ico.org.uk/media2/migrated/4022261/how-to-use-ai-and-personal-data.pdf.

[7] European Data Protection Board Opinion 28/2024 on Certain Data Protection Aspects of the Processing of Personal Data in the Context of Artificial Intelligence Models, December 17, 2024, paragraphs 29, 36-37; Opinion available here: https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en.

2.3. Other regulations

It should be borne in mind that when using AI systems, other regulations may also apply, such as:

- copyrights regulations (the Law on Copyright and Related Rights);
- general regulations on contractual relations (Civil Code in the case of contracts governed by Polish law);
- directive on liability for defective products[8];
- Act on the National Cyber Security System and other Polish regulations in the field of cyber security;
- DORA Regulation[9];
- CER Directive[10].

This whitepaper does not focus on these regulations. That topic has been addressed in, for instance, "AI in the work of an attorney-at-law" recommendations.[11]

3. Risks Associated with usage of AI-based tools

The decision to use AI systems in an organization may introduce a complex and evolving set of risks that organizations must proactively address.

This section provides an overview of the four non-legal categories of AI risk that should be considered as part of any governance framework, such as:

A. **Ethical, Societal and Environmental risks**, including bias, opacity, the lack of human oversight in automated decision-making and rising energy and water consumption of generative AI systems;

B. **Operational risks**, such as unpredictable model behavior, hallucinations, or poor reliability under real-world conditions;

C. **Reputational risks**, stemming from public backlash, loss of customer trust, or stakeholder scrutiny;

D. **Security and Privacy risks**, related to data misuse, adversarial attacks, or unintended information disclosure.

By understanding and addressing these categories holistically, organizations can lay the foundation for responsible AI use.

Renata Zalewska
Senior Commercial Attorney at Microsoft

The fine-tuning of AI models, especially with the use of personal data, is not yet very common in Poland, however, we may expect that it will become more and more popular. At this moment, the Clients are more interested in potential consequences of using AI models trained on materials that could have contained personal data. Microsoft provides Clients with many useful information in this area. In particular, they may review the Azure Open AI Foundational Privacy Impact Assessment or GDPR & Generative AI A Guide for Customers.

We also address this topic in the legal analysis for Azure Open AI and M365 Copilot which were prepared for Microsoft by the SK&S

[8] Directive of the European Parliament and of the Council (EU) 2024/2853 of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC. The Directive should be implemented in the Polish legal order by 9 December 2026.

[9] Regulation (EU) 2022/2554 of the European Parliament and of the Council 14 December 2022 on the operational digital resilience of the financial sector and amending Regulations (EC) No1060/2009, (EU) No. 648/2012, (EU) No. 600/2014, (EU) No. 909/2014 and (EU) 2016/1011. The regulation is effective as of January 17, 2025.

[10] Directive (EU) 2022/2557 of the European Parliament and of the Council of 14 December 2022 on the resilience of critical entities and repealing Council Directive 2008/114/EC. The CER Directive should be implemented in the Polish legal order by 17 October 2024.

[11] The Recommendations are available here: https://kirp.pl/wp-content/uploads/2025/05/rekomendacje-ENG-NET.pdf.

Renata Zalewska
Senior Commercial
Attorney at Microsoft

Working with many Clients, especially from regulated sectors who intend to implement M365 Copilot, Azure Open AI or GitHub Copilot, we often receive many questions on the risk associated with use of these AI solutions and the way these risks may be mitigated. Below, this White Paper presents some measures of risk mitigating. In addition, I would also like to refer to the documents available on Microsoft Service Trust Portal. There, you may find in particular "M365 Copilot and Copilot Chat Risk Assessment Quick Start Guide" which describes AI risks and their corresponding mitigations in the context of Copilot. Many of these mitigations are applicable beyond the Copilot service to AI in the broader context of the Microsoft cloud. It also includes a practical sample risk assessment which may be carried out by the Client:

Sample Risk Assessment: Questions & Answers

This section presents a sample set of questions and answers that can be used to assess the features, functionalities, and the surrounding security and compliance posture of Copilot, as well as the broader implications and impacts of using the service. The questions are derived from customer inquiries and the answers are based on information from various Microsoft internal teams and sources, including engineering, legal, compliance, privacy, trust, customer experience, the Office of the CTO, Azure Security, Azure OpenAI, and many more teams. Some responses also include direct attestation from OpenAI, the Microsoft strategic partner from which we source our Copilot foundation models.

The risk assessment questions are grouped by ID with the key as follows:

P = Privacy

S = Security

SR = Supplier Relationship

MD = Model Developer

From OpenAI = Official statement obtained directly from OpenAI

ID	Question	Microsoft Response
P1	Are the AI models trained on personal data? If not, then how does the AI solution provider ensure that the model is not being trained with personal data e.g. anonymization techniques?	**From OpenAI:** OpenAI maintains high standards for data anonymization prior to training models. OpenAI continues to improve these systems iteratively and maintain quality control through human data review to remove false negatives. OpenAI employs extensive PII scrubbing, PII filtering, and/or human review on all training data. Refer to OpenAI for additional information.

A. Ethical, Societal and Environmental risks

Ethical and societal risks arise when AI systems, particularly those used in decision-making processes, produce outcomes that negatively affect individuals, groups, or broader democratic values. These risks are often complex and difficult to detect, yet they have profound implications for fairness, transparency, and human dignity.

One significant ethical challenge in AI is the **risk of bias and discrimination**. AI systems frequently rely on historical data that may reflect existing inequalities or prejudices. When left unaddressed, these biases can lead to discriminatory outcomes in areas such as hiring, credit scoring, education, or law enforcement.

A second ethical concern is **the opacity of many AI systems**, particularly those based on complex or black-box models. The internal logic and the way the models reach their conclusions is often difficult to explain, even by their developers. This lack of transparency **undermines user trust and poses challenges for accountability**, especially when AI systems influence decisions that significantly impact individuals.

Another key societal risk is the absence of **human oversight**. When organizations rely too heavily on AI to make or influence critical decisions, without clear human supervision, they risk delegating authority to systems that lack ethical judgment or contextual understanding. This can lead to outcomes that are not only technically flawed but also socially or legally unacceptable.

Finaly, critical environmental risk is the rising energy and water consumption of generative AI systems. According to a recent UNESCO (2025). “Smarter, Smaller, Stronger: Resource-Efficient Generative AI & the Future of Digital Transformation” report, training a single large language model can consume up to 50 GWh, that is more than some developing countries use in a year.

Inference, or everyday AI usage, is now the largest energy driver; ChatGPT alone consumes around 124 GWh annually, equal to the usage of 1.3 million people in Ethiopia. By 2027, AI-related water consumption is expected to triple, potentially exceeding that of countries like Denmark. Mitigation strategies such as quantization, prompt optimization (providing shorter but more concrete prompts), and smaller task-specific models can reduce AI’s energy use by up to 90%, significantly lowering its environmental impact.

B. Operational risks

Operational risks refer to failures in how AI systems function in real-world settings. These risks are not limited to technical malfunctions; they include a wide range of challenges related to reliability, consistency, adaptability, and the alignment of AI behavior with intended outcomes. As organizations increasingly use AI for the purpose of business processes and decision-making, the consequences of operational failures can be significant. One of the most prominent operational risks is the **unpredictability of AI performance**. In machine learning models, especially those trained on dynamic or unbalanced datasets, this can result in degraded performance, inaccurate predictions, or unintended actions. Same applies to the phenomenon of **hallucinations**, which means that large language models may produce outputs that are factually incorrect, logically incoherent, or misaligned with user intent, all while appearing very convincing and accurate. Another example of such unpredictability is the phenomenon known as “**misbehavior**”, where AI chatbots based on language models, provide a final response that is at odds with the reasoning they appear to use to reach that conclusion. This stems from the fact that, even when the model presents a chain-of-thought explanation, the final answer may contradict or diverge from that reasoning. The output may seem well-structured and logical, creating a false sense of reliability, but the conclusion can still be flawed, incoherent, or disconnected from the steps that supposedly led to it. This highlights the challenge of interpreting and trusting model outputs, especially in high-stakes or sensitive contexts.

Additionally, operational risk encompasses issues related to **data quality** and lifecycle management. AI systems are only as reliable as the data on which they are trained and the systems in which they operate. Poor data quality, untracked model drift, and inadequate performance monitoring can result in outdated, biased, or misleading outcomes.

Operational risks also arise when **AI systems are not adequately integrated into broader organizational workflows or when human users lack clarity about how to interact with or supervise the system**. Misuse or overreliance on AI, particularly in environments without proper training can lead risk exposure.

C. Reputational risks

Reputational risks associated with AI arise when using the AI leads to a loss of trust among customers, employees, regulators, or the public. Unlike operational or legal risks, which may be contained within specific processes or transactions, reputational risks often spread quickly and unpredictably, in particular amplified by social media, public scrutiny, and the growing expectation that companies act responsibly and transparently when using advanced technologies.

AI-related reputational damage can result from a variety of causes. A single high-profile failure, such as a biased hiring algorithm, a hallucinating chatbot, or use of a hallucinated data, can undermine stakeholder confidence across an entire organization. In many cases, the reputational impact is not only linked to the technical failure itself but to the perception that the organization lacked the foresight, governance, or ethical commitment to prevent it.

D. Security and Privacy risks

The integration of AI into core business functions, and/or digital services exposes organizations to a new layer of vulnerabilities, both from malicious attacks and from unintended design flaws that compromise confidentiality, integrity, or availability of information.

From a privacy perspective, AI systems often process personal data in ways that can be difficult to trace or justify under traditional data protection frameworks. Large language models may infer sensitive information that was not explicitly collected, such as health status, political opinions, or behavioral profiles, potentially violating principles of purpose limitation and data minimization. In some cases, AI systems may unintentionally re-identify individuals from anonymized datasets, creating additional exposure under regulations such as the GDPR.

Security risks are equally pronounced. AI systems can become targets of sophisticated cyber threats, including model inversion, data poisoning, and adversarial attacks. These techniques can cause AI models to misbehave, leak confidential information, or become unreliable under targeted manipulation. Unlike traditional software systems, AI models often lack deterministic outputs, making it more difficult to detect whether they have been compromised.

Moreover, AI models can sometimes memorize and reproduce sensitive information from their training data, even if it was not meant to be retained. As organizations explore the use of such technologies, they must assess not only technical performance but also the risk of privacy violations, copyright infringement, or regulatory non-compliance.

4. Recommendations and best practices

AI governance should not be seen as a control mechanism imposed from above. Rather, it should become part of the organization's innovation culture, ensuring that AI systems are trusted, reliable, and accountable by design. A well-implemented governance framework empowers teams to innovate confidently. Knowing that risks are managed, users are protected, and the company is prepared for legal and public scrutiny.

While Section 3 outlined the key risks that may emerge when an organization adopts AI systems, Section 4 offers practical recommendations to avoid or mitigate those risks and presents best practices that can be embedded into organizational processes to support the responsible and effective use of AI.

4.1. Recommendations

An effective AI governance framework must be grounded in clearly defined principles that guide the ethical, legal, and operational management of AI systems throughout their lifecycle. The following core pillars represent basic elements of responsible AI use. Together, they help organizations proactively identify and mitigate the risks associated with AI, while promoting trust, transparency, and long-term value.

4.1.1. Transparency and engagement

Transparency is key to managing ethical, reputational, and legal risks related to AI. An organization should be clear, both internally and externally, about how and why AI systems are being used, especially when they influence decisions about people. External communication is in particular important as under Article 50 of the AI Act the deployers are obliged to design the AI in a way that users, in certain cases, are informed that they are interacting with an AI.

Recommended actions:

- Document and communicate the purpose of each AI tool, including its capabilities and limitations, in plain language
- Use explainable AI models where possible, or choose vendors that provide explainability features
- Engage stakeholders early, including employees and users, when introducing AI into sensitive processes
- Designate an AI Champion(s) to serve as a point of contact for internal questions and concerns related to AI systems
- Map your processes to establish if external users are interacting with AI systems

Risks addressed:

✅ A. Ethical, Societal and Environmental

✅ C. Reputational

4.1.2. AI Literacy

Many operational and societal risks arise not from the technology itself, but from how it's used. AI literacy across the organization is critical, especially as tools become more accessible to non-experts.

Recommended actions:

- Provide practical training for staff who use AI systems, covering not only how to use them to spot issues like bias, misbehavior or hallucinations, but also how to draft efficient prompts, which can help reduce unnecessary computation and lower energy consumption.
- Include AI awareness in onboarding for relevant roles (HR, marketing, compliance, etc.)
- Clarify user responsibilities and who is accountable for reviewing AI outputs before action is taken.
- Promote a culture of cautious experimentation, where employees are encouraged to ask questions and escalate concerns.

Risks addressed:

- ✅ A. Ethical, Societal and Environmental
- ✅ B. Operational
- ✅ C. Reputational

Microsoft offers structured AI literacy support aligned with Article 4 of the EU AI Act. Through its Trust Center, Microsoft makes available a "Getting Started Guide" and additional materials that help organizations ensure staff have the knowledge to use AI safely and responsibly. This includes understanding AI systems' risks and opportunities, recognizing their responsibilities under the AI Act, and staying up to date with evolving requirements. The resources are designed to be concise, practical, and ready for immediate use.

AI Literacy Starting Guide

EU AI Act: Compliance

EU AI Act: AI Literacy

4.1.3. Security and Robustness

AI systems introduce new vulnerabilities that any, even small organization must consider, especially regarding data handling and system reliability. While not all organizations may have large security teams, they can still take essential steps to reduce exposure.

Recommended actions:

- Ensure that AI tools (especially LLMs) don't process sensitive data unnecessarily, use anonymized or synthetic data when feasible.
- Choose vendors that comply with security best practices, including appropriate level of encryption, access controls, and regular audits.
- Regularly review AI system performance in real-world conditions, and monitor for drift (misbehavior), hallucinations, or incorrect outputs.
- Establish basic incident response protocols for AI-related failures, including reputational or data incidents.
- If your organization does not have a dedicated internal IT or cybersecurity team, seek guidance from trusted external consultants or specialists. They can help assess risks, guide system selection, and establish basic safeguards appropriate to your size and sector.

Risks addressed:

- ✅ B. Operational
- ✅ C. Reputational
- ✅ D. Security and Privacy

Microsoft provides detailed security recommendations for entities which implement AI tools. In particular, for M365 Copilot it issued guidelines how to build a strong foundation of security based on zero trust principle. This information is available at How do I apply Zero Trust principles to Microsoft 365 Copilot?

Apply principles of Zero Trust to Microsoft 365 Copilot

4.1.4. Human Oversight

Human oversight is a critical safeguard against ethical, operational, and reputational risks. While AI can automate tasks, decisions that impact people should always include a human check, irrespective of the size of the organization.

Recommended actions:

- Define which decisions require human validation, especially those with legal, financial, or ethical consequences
- Maintain human accountability by ensuring that someone is always ultimately responsible for reviewing and signing off on critical decisions

Risks addressed:

- ✅ A. Ethical, Societal and Environmental
- ✅ B. Operational
- ✅ C. Reputational

In its AI Code of Conduct Microsoft indicates that responsible use of AI includes also obligations on customers side. In particular they must ensure that all of their applications built with Microsoft AI Services, including applications that make decisions, or take actions, autonomously or with varying levels of human intervention:

- Implement technical and operational measures to detect fraudulent user behavior in account creation and during use.
- Implement strong technical controls on inputs and outputs, including decisions made and actions taken by their applications, to reduce the likelihood of misuse beyond the application's intended purpose.
- Disclose when the output, decisions, or actions are generated by AI, including the synthetic nature of generated voices, images, and/or videos, such that users are not likely to be deceived or duped – or able to deceive or dupe others – into believing they are interacting with a real person, that any generated content is authentic or, without their consent, attributable to a specific individual, or that any AI-generated decision or action is done by a human.
- Are tested thoroughly, continuously, and are subject to appropriate human oversight so customers can find and mitigate undesirable behaviors.
- Establish feedback channels that allow users to report abuse or issues, and ensure reasonable responses to feedback that is received.
- Implement additional scenario-specific mitigations, as appropriate, to ensure responsible use of the Microsoft AI Service, including meaningful human oversight.
- Provide all necessary notices and obtain all necessary consents as required by applicable law for both the customer and Microsoft to process data, including third-party data, as part of a customer's use of the Microsoft AI Service.
- Implement robust security and access control measures, including protecting the Microsoft AI Service resource permissions and having strong user authentication mechanisms.

4.2. Best practices

To translate those basic recommendations into daily practice, which is understandable for the end user, and ensure that recommended actions are taken, organizations need a step plan. The following actions are essential building blocks:

✅ Appoint an AI champion

Appoint an individual within the organization whose responsibility will be to monitor changes in AI-related laws, regulations, and industry standards. As AI technologies evolve rapidly, they help the organization stay informed, compliant, and adaptable in a fast-changing environment. An AI Champion promotes responsible and effective use of AI technologies. They help identify practical use cases, support teams in adopting AI tools, and ensure that AI initiatives align with the company's goals. Being AI champion does not have to be a full-time job or a stand-alone position. Depending on the work load, it is even recommended that the person who knows organization and fulfils other responsibilities is an AI champion.

✅ Review AI Use and Identify Potential Risks

When an AI system is being embedded into the organization, it is essential to assess where and how it will be used, and to identify points at which risks may. These may include legal (e.g. under Article 50 of the AI Act), ethical, operational, or data-related concerns. Running this analysis early helps ensure that AI is introduced safely, responsibly, and in line with the AI Act and organization's goals.

Given the rapid pace of AI development, this risk analysis should not be a one-time exercise. Organizations should re-evaluate AI systems regularly to account for changes in technology, regulation, or context of use. It is recommended that this process is led by the AI Champion, who ensures ongoing oversight, alignment with internal policies, and responsiveness to emerging risks.

✅ Train Stakeholders

AI governance is only effective when people across the organization understand their role in it. Offer tailored training, preferably delivered by the AI Champion, to ensure that all stakeholders are aware of:

- What AI is (and isn't) capable of in your business context.
- How to properly use the AI systems adopted within the organization.
- The systems' intended purposes, limitations, and acceptable use,
- Key legal, ethical, and operational risks, specific to your organization.
- How to recognize issues such as bias, hallucinations, or misuse and mitigate them.
- Their individual responsibilities in using AI safely and responsibly.

Training should be role-specific and updated regularly to reflect evolving technologies, regulations, and internal policies.

Microsoft offers a range of training modules on the Microsoft Learn platform that can help ensure not only AI Champions but also entire teams within your organization stay up to date with AI developments and understand how to use AI technologies safely and responsibly.

Training - Courses, Learning Paths, Modules

✅ **Draft a Clear and Actionable Internal AI Policy**

An internal AI policy is a foundational tool for effective AI governance. It sets shared expectations across the organization and provides practical guidance on how AI should, and should not be used. The policy should be short, simple, and relevant to your organization's actual use of AI. It does not need to be legalistic or overly technical.

Start by clearly stating the organization's commitment to using AI responsibly and in compliance with applicable laws, including the AI Act. Define key roles (e.g. AI Champion), outline basic principles (such as human oversight, transparency, and data protection), and set out high-level rules for adopting, deploying, and monitoring AI systems.

This document should be reviewed and updated as the organization's use of AI evolves. A one-pager, written in plain language, is often more effective than a detailed manual, its purpose is to guide decisions, not to sit unused.

4.3. Example of Good Practices: Microsoft

Microsoft has long been actively developing and implementing principles of responsible AI, becoming one of the global leaders in AI governance. The company not only complies with existing regulations such as the AI Act and GDPR but also creates its own ethical and technical standards that set the direction for the entire industry.

Microsoft's AI governance framework integrates ethical standards, organizational oversight, and practical tools to ensure that AI technologies are developed and used responsibly.

At the core is the **AETHER Committee** (AI and Ethics in Engineering and Research), which advises on ethical, legal, and societal risks. Microsoft enforces its **Responsible AI Standard** across all product teams, covering principles such as fairness, reliability, transparency, privacy, and accountability.

[12]

Responsible AI principle	Definition	Risk assessment question
AI Privacy and Security	AI workloads should respect privacy and be secure.	How might AI workloads handle sensitive data or become vulnerable to security breaches?
Reliability and Safety	AI workloads should perform safely and reliably.	In what situations could AI workloads fail to operate safely or produce unreliable outcomes?
Fairness	AI workloads should treat people equitably.	How could AI workloads lead to unequal treatment or unintended bias in decision-making?
Inclusiveness	AI workloads should be inclusive and empowering.	How might certain groups be excluded or disadvantaged in the design or deployment of AI workloads?
Transparency	AI workloads should be understandable.	What aspects of AI decision-making could be difficult for users to understand or explain?
Accountability	People should be accountable for AI workloads.	Where could accountability be unclear or difficult to establish in the development or use of AI?

[12] Govern AI - Cloud Adoption Framework | Microsoft Learn

Microsoft's high-risk systems undergo mandatory Responsible AI Impact Assessments (RAIAs), and built-in controls, such as content labeling, human oversight mechanisms, and audit tools to help operationalize compliance.

Microsoft also provides resources to support entities implementing AI in their organisations. Service Trust Portal platform offers documents, reports, and resources for organizations, including those related to AI.

Service Trust Portal

Applying Microsoft-Inspired AI Governance in Smaller Organizations

Even small and mid-sized organizations can adopt proven governance principles by tailoring them to their size and resources. The following approaches, inspired by best practices implemented by Microsoft, provide a practical governance baseline:

- Start with a short, internal AI policy (1-2 pages) that outlines your commitment to responsible AI use. You can use Microsoft's Responsible Ai Standards as an inspiration.
- Use risk assessment checklists before adopting AI tools, especially those affecting decisions about people, customers, or compliance-sensitive areas.
- Favor vendors that support governance, for example by offering AI audit logs, usage monitoring, human oversight features, or model transparency (such as Microsoft 365 Copilot or Azure AI).
- Embed responsible AI into procurement and onboarding processes by asking vendors for documentation aligned with AI Act obligations.
- Promote continuous awareness via role-specific training, internal briefings, or curated resources (e.g. Microsoft Learn – Responsible AI).

5. Challenges and Limitations

While the need for AI governance is increasingly clear, implementing it in practice presents a range of challenges. These obstacles can arise at organizational, technical, and regulatory levels. Ignoring them can lead to stalled initiatives, ineffective policies, or ungoverned AI usage.

5.1. Organizational Resistance and Lack of Awareness

Many organizations are still in the early stages of AI maturity, and awareness of governance requirements may be limited. Leaders and employees may view AI governance as a bureaucratic burden, an impediment to speed and innovation, threat or temporary hype. In some cases, fear of redundancy, especially among so-called white collars & knowledge workers, can lead to silent or unconscious resistance. Employees may worry that automation will replace their roles, making them hesitant to engage with AI initiatives.

These concerns are intensified by the rapid speed of AI advancements, which can leave employees, and also management, feeling disoriented, overwhelmed, or left behind. The pace of change creates uncertainty about what skills will be needed, which roles will evolve, and how to stay relevant.

In this context, management plays a pivotal role. Leaders must actively prepare their organizations for AI transformation, and lead by example.

What can be done to mitigate those risks?

- **Communicate early, often and in understandable way:** Share the purpose of AI adoption, clarify what it will and will not do, and emphasize that AI supports and augment, not replaces, human decision-making. Use the language and communication methods understandable for particular group of end-user (a different approach may be needed with respect to the lawyers and IT developers).
- **Appoint an AI Champion:** Identify a trusted internal person who can demystify AI, serve as a liaison between teams, and model responsible use.
- **Offer upskilling opportunities:** Provide accessible training on how AI works, how to use it safely, and how it can make everyday tasks more efficient—not obsolete.
- **Make AI a team effort:** Involve employees in testing and shaping AI use cases. Encourage feedback and create safe spaces for raising concerns.

5.2. Balancing Innovation with Control

AI works best in environments that encourage testing and quick changes. But these environments often clash with the strict oversight needed for governance. Finding the right balance between flexibility and accountability is tough: too much control can limit innovation, while too little can lead to risks.

Governance models should be flexible and based on risk, providing clear guidelines without blocking progress. This is especially important when AI is used in fast-paced business units or areas that interact directly with customers.

What can be done to mitigate those risks?

- **Adopt a risk-based governance model:** Apply stricter oversight only to high-impact or sensitive AI use cases, allowing greater freedom for low-risk experimentation.
- **Define "guardrails" rather than restrictions:** Set clear, simple rules (e.g., “No sensitive data in public AI tools”) instead of detailed technical policies, such as Clear and Actionable Internal AI Policy.
- **Encourage pilot projects:** Start with small-scale tests.

Microsoft supports a risk-based approach to AI governance, consistent with the EU AI Act, which applies the most stringent requirements to high-risk AI systems while allowing flexibility for low-risk use cases. Through its Trust Center EU AI Act Compliance | Microsoft Trust Center, Microsoft provides practical guidance on how organizations can assess risk, apply proportionate oversight, and implement safeguards without stifling innovation.

5.3. Technical Complexity and Lack of Tools

AI systems, particularly those based on machine learning or generative models, are inherently complex, often functioning as "black boxes" with outputs that are difficult to explain or audit. Many organizations lack the technical infrastructure or tooling needed for explainability, bias detection, or auditability.

Even when tools exist, they may not be integrated into workflows, or teams may lack the expertise to use them effectively. Bridging this gap requires investment in technical capacity, cross-functional collaboration, and, where possible, the use of standardized templates or open-source governance toolkits.

What can be done to mitigate those risks?

- **Use explainable models where feasible:** Choose simpler or better-known models when transparency is critical.
- **Select vendors with built-in governance features:** Prioritize AI tools that come with dashboards for usage, logs, audit trails, or model explainability. Ask about input data, model use, and decision accountability.
- **Invest in foundational skills:** Train or hire at least one person who understands the basics of AI system operations.

Microsoft provides organizations with practical tools to address the challenges of explainability, bias detection, and auditability in AI systems. For example, the Responsible AI Dashboard Use the Responsible AI dashboard in Azure Machine Learning studio - Azure Machine Learning | Microsoft Learn in Azure Machine Learning offers capabilities for model explainability, fairness assessment, error analysis, and data exploration, enabling teams to identify and mitigate risks throughout the AI lifecycle. In addition, the Azure AI Content Safety Azure AI Content Safety – AI Content Moderation | Microsoft Azure service helps detect harmful or biased outputs, while built-in monitoring and logging features in Azure AI services support transparency and accountability. These resources are designed to integrate into development workflows and assist organizations in meeting governance requirements under frameworks such as the EU AI Act.

5.4. Evolving and Fragmented Legal Environment

The legal landscape surrounding AI is in flux. The AI Act is set to become a global benchmark, but national and sector-specific regulations are still developing, and they often diverge. Moreover, the legal framework implementing AI Act is still being developed. This regulatory fragmentation makes it difficult for multinational organizations to create a unified governance approach.

Moreover, many legal obligations related to AI are indirect, flowing from data protection, consumer protection, anti-discrimination, and sectoral rules. As a result, organizations must navigate a patchwork of obligations, often without clear precedent or enforcement history. Staying ahead requires continuous legal monitoring, alignment between legal and technical teams, and a proactive, rather than reactive, compliance strategy.

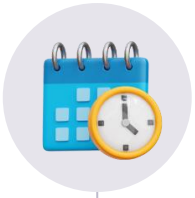
What can be done to mitigate those risks?

- **Monitor key areas, not every update:** Focus on data protection, non-discrimination, consumer rights, and AI risk categories most relevant to your sector.
- **Assign legal responsibility to a specific individual:** Even in small organizations, ensure someone is tasked with tracking legal and compliance risks, for instance AI Champion.
- **Encourage collaboration between legal and tech:** Regular check-ins between whoever manages compliance and whoever operates AI tools is essential.
- **Use sectoral guides and associations:** Rely on industry groups, local chambers, or bar associations to stay informed on emerging rules.

General AI Act Compliance Tips

COMPLIANCE TIP 1

At the outset, **classify your AI systems according to the AI Act's risk tiers**: prohibited, high-risk, limited-risk, and minimal-risk. This determines the scope and scale of your compliance obligations and ensures you are building the right controls from day one.



COMPLIANCE TIP 2

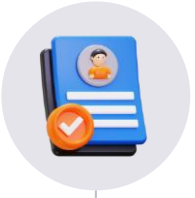
Do not wait for the final deadline. **Start aligning now with the AI Act's stages timeline**, especially the February and August 2025 milestones for general rules and GPAI models. Initial action reduces the risk of last-minute fire drills.

COMPLIANCE TIP 3

Even if your system runs outside the EU, you must comply if its outputs affect EU individuals. Document on how and where your system is used, including indirect or API-based access.

COMPLIANCE TIP 4

Maintain an internal inventory of all AI systems used or developed, with details like risk classification, owner, deployment status, and regulatory documentation. This registry is your first line of defense in any audit or inquiry.



COMPLIANCE TIP 5

Ensure all AI systems are designed and implemented with **appropriate human oversight** mechanism. Document who has the authority to review, override, or halt decisions — and when that must happen.

COMPLIANCE TIP 6

Everyone involved in high-risk AI systems needs to understand their role in keeping the system compliant and safe. Therefore, legal, IT, product, and HR teams should all receive AI Act awareness training.

General AI Act Compliance Tips

COMPLIANCE TIP 7

Even if your system is not high-risk AI system, you may still face **transparency obligations**. For example, (1) disclosing that users are interacting with AI; (2) labelling AI-generated content, such as text, images, audio, or video, that could be mistaken for human-created material or (3) informing individuals when emotion recognition or biometric categorization technologies are being used. Always check whether any duty applies before deployment, taking into account the current state of legal obligations under the AI Act. Review guidance with practical examples of compliant disclosures, templates or minimum standards, and best practices provided by the EU bodies and regulators.

COMPLIANCE TIP 8

Ensure your organization conducts regular **bias and fairness assessments** across all AI systems, especially those used in decision-making. Involve diverse stakeholders during development and design user-facing explanations that clarify how decisions are made and how users can contest them.

COMPLIANCE TIP 9

Implement a **robust testing and monitoring protocol** for all AI systems, before and after deployment. Include performance validation under real-world conditions, model drift monitoring, and clear incident escalation procedures to ensure systems behave reliably over time.

COMPLIANCE TIP 10

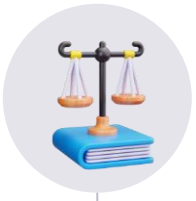
Establish ahead a clear **AI transparency and communication strategy**. Proactively inform users, clients, and partners about how AI is used and what safeguards are in place. In case of error or failure, respond with speed, openness, and accountability to preserve public trust.

COMPLIANCE TIP 11

Integrate **privacy-by-design and security-by-design** principles into the AI development process. Apply encryption, access controls, and adversarial robustness testing. Regularly audit both models and data pipelines for compliance with GDPR and other applicable regulations.

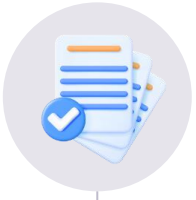
SOŁTYSIŃSKI KAWECKI SZLĘZAK

AI Act Compliance Tips for Providers



PROVIDER'S COMPLIANCE TIP 1

Embed compliance into each phase of AI development: ideation, training, testing, deployment, and post-market monitoring. Ensure that AI Champion are in control over these processes and control traceability.

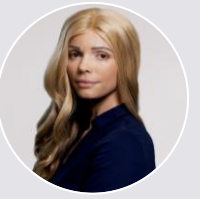


PROVIDER'S COMPLIANCE TIP 2

High-risk AI systems require extensive documentation that must be created, maintained, and kept up to date. Establish a system of checks to ensure that all required documentation is produced and easily retrievable at every stage of development. This helps ensure that documentation is not treated as an afterthought, but as an integral part of responsible AI development.

Agata Szeliga
Partner, attorney-at-law
✆ +48 698 660 648
✉ agata.szeliga@skslegal.pl

Anna Tujakowska
Senior Counsel, attorney-at-law
✆ +48 698 694 645
✉ anna.tujakowska@skslegal.pl



Sylwia Macura-Targosz
Senior Associate, attorney-at-law
✆ +48 694 415 447
✉ sylwia.macura-targosz@skslegal.pl