Trustworthy AI

Implementing the EU AI Act as a value driver

pwc

Foreword

Artificial intelligence (AI) is a key technology for digital transformation in private and public organisations. By 2030, we expect that AI will be a direct or indirect component in all processes and products along the entire European value chain. From smart washing machines to (partially) autonomous vehicles, from automated application processing to intelligent chatbots, and from optimised maintenance processes to production robots – the rapid spread of AI can be seen all around.

In addition to the economic relevance of use cases, trust in the performance, security, reliability, and fairness of AI is an essential factor in deciding for or against the use of AI systems. Both dimensions are inextricably linked because, in practice, economic efficiency and trustworthiness are mutually dependent on the use of AI. Finally, the use of the technology requires certainty and trust on the part of users, customers, and decision-makers, which only sound governance based on best practices and generally accepted standards can provide.

For this reason, the institutions of the European Union have developed and launched harmonised regulations for AI systems. This comprehensive “EU AI Act” is directly applicable in the member states and affects both private and public organisations – regardless of whether they are providers or deployers of an AI system.

The regulation aims both to promote European AI value creation through uniform standards and to protect EU citizens. Specifically, the EU AI Act calls for holistic AI governance that promotes the development and use of high-quality AI systems and makes the risks of AI systems manageable and transparent throughout the entire lifecycle. Organisations must balance their legally compliant implementation of the requirements in and around AI systems case-by-case at an early stage. Otherwise, those affected expose themselves to many legal risks, these include not only potential fines under the EU AI Act, but also fines under the GDPR, and industry-specific regulations, or also liability claims if they use a defective AI system.

However, implementing diverse regulations not only poses challenges for organisations but also opens opportunities for them to improve their AI in terms of quality. This is all the more true because the increasing use of AI puts the responsible handling of data more strongly than ever before in the focus of Corporate Social Responsibility – companies‘ responsibility for the environment and society. With the right approach and the corresponding interdisciplinary competencies in AI governance and law, organisations can avoid efforts and risks, shorten (market) introduction times for their AI systems and take a pioneering role in digital transformation with AI.

1. The synergy potential of scaling and regulation

Scaling

AI systems are finding broader and more intensive applications and are thus increasingly becoming a core component of global value creation. One-third of larger German companies are already using AI systems, another 44% are testing AI in pilot projects[1], and three-quarters of these companies plan to put them into operation in the next two years.[2] This is also imperative because AI systems generate the greatest added value when they are applied on a large scale and quickly move beyond the pilot phase. Economies of scale take on a central role in AI projects because high fixed costs in development meet low marginal costs in operation.

However, there is often a lack of adequate AI governance systems to operationalise and subsequently scale AI-enabled solutions. Without them, operating the system becomes a risk to an organisation's reputation and profitability. In addition, the use of AI without adequate monitoring and governance systems can have a tangible negative impact on the quality of life of many people in a critical situation, as the child benefit scandal in the Netherlands showed, where migrants were discriminated against for years by an algorithm.[3]

Nevertheless, around 45% of companies in the private sector lack qualified staff to review AI adequately. Nearly 50% lack the human resources and expertise to implement AI.[4] Under these circumstances, there are immense challenges in making AI systems safe, robust, fair, and effective.

---

1 IBM Global AI Adoption Index 2022 (p.4)

2 Gartner Top 10 Data and Analytics Technology Trends for 2020

3 Politico "Dutch scandal serves as a warning for Europe over risks of using algorithms" 2022

4 PwC DE Responsible AI Survey 2022

Regulation

The EU AI Act imposes several far-reaching requirements on many AI systems. The regulation is expected to have a similar impact on the affected AI systems as, for example, the General Data Protection Regulation (GDPR) had on the processing of personal data.

In addition to the general regulation of AI by the EU AI Act, other requirements are relevant to AI use cases: Horizontal ones, such as the GDPR or the EU Data Act, and vertical or sectoral ones, such as the EU Medical Devices Regulation (MDR) or the German Regulation on the Approval and Operation of Motor Vehicles with Autonomous Driving Function in Specified Operating Areas (AFGBV).

Already today, 26 % of companies see regulatory requirements as a barrier to implementing AI systems[5] – a proportion that will increase significantly with the entry of the EU AI Act into force. To ensure that AI governance can address the multiple legal and technical aspects, aligning it with the regulatory framework early on is crucial. Failures are difficult to correct in retrospect and can lead to unintended liability risks, increased costs, or inefficient allocation of resources in AI development.

Building bridges between scaling and regulation

So are organisations only implementing AI governance for compliance reasons? In our experience, it is an opportunity to manage and improve AI transformation. A closer look at the EU AI Act reveals a great deal of overlap between the requirements set in regulation and the structures necessary for operationalising and scaling AI systems. Conformity assessments and specific documentation requirements are, of course, primarily relevant from a compliance perspective. However, flexible data, risk, and lifecycle management systems, especially, are indispensable to enable organisations to successfully transition their AI systems from the pilot phase to scaled operations.

Selection of relevant legal norms for the use of AI

[5] PwC DE Responsible AI Survey 2022

2. The EU AI Act and its implications

Quo vadis EU AI Act?

The EU AI Act is a regulation of the European Union to define harmonised regulations for systems with artificial intelligence. The term “AI system” is very broadly defined in the EU AI Act:

> “AI system” means a machine-based system designed to operate with varying levels of autonomy, 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.[6]

The EU Commission's original proposal for the regulation of AI systems was published in April 2021 and was revised and amended several times during the legislative process before finally being adopted in 2024. After the EU AI Act comes into force, transitional periods are provided for until it takes effect, within which organisations must implement the requirements. The deadlines are (with minor deviations) 6 months for prohibited AI systems, 12 months for general purpose AI (GPAI) systems, 24 months for high-risk AI systems in accordance with Annex III and finally up to 36 months for high-risk AI systems in accordance with Annex I. After this period, non-compliance could result in significant fines of up to 35 million euros or 7% of the previous year's global turnover for companies that fail to comply with the prohibitions and 15 million euros or 3% of the previous year's global turnover for companies that fail to comply with the rules for high-risk or GPAI systems.[6]

The impact of the EU AI Act on private and public organisations is complex and must be determined on a case-by-case basis. In the following sections, we therefore take a look at the three essential aspects of the regulation: the role of regulated organisations, the risk classification of AI systems and the requirements for high-risk AI systems and GPAI systems. We also look at the challenges involved in practical implementation and conclude by outlining a solution.

---

[6] Current version of the AI regulation

Addressees in the AI value chain

At the centre of the regulation are the deployers and providers of AI systems. They bear the main burden of the requirements. On the one hand, these are providers who place regulated AI systems on the market or put them into operation in the European Union, regardless of whether these providers are established in the EU or in a third country, as well as deployers of regulated AI systems who are established in the European Union. In addition, it also applies to providers and deployers of regulated AI systems established or resident in a third country if the result produced by the system is used in the European Union. The EU AI Act thus has an extraterritorial application and also indirectly regulates beyond the borders of the European Union and the European Economic Area.

Under certain circumstances, the EU AI Act also places other actors, such as importers, distributors or manufacturers of high-risk AI systems, on an equal footing with providers in terms of their obligations (see example 1).

At the same time, the deployers, representatives of the providers (so-called authorised representatives), importers and distributors of high-risk AI systems are required to ensure the extensive obligations of the providers – insofar as this is their responsibility (see example 2).

In order to ensure compliance with the requirements of the EU AI Act across the entire value chain, the obligations of deployers of high-risk AI systems apply alongside the requirements of other stakeholders. Deployers, i.e. natural or legal persons, including public authorities, institutions or other bodies under whose responsibility the system is used, are obliged, among other things, to ensure human oversight and must ensure that input data is subject to the specified purpose of the high-risk AI system.

In fact, the involvement of the various stakeholders means that every organisation dealing with AI systems must define and continuously review its own role in order to fulfil its obligations correctly. Furthermore, this involvement means that purchasers of AI systems in particular should optimise their legal position vis-à-vis manufacturers, providers and distributors. For example, drafting a contract under IT law offers the opportunity to ensure specific performance obligations, such as support and maintenance in particular, to ensure the suitability of the AI system for use within the scope of the EU AI Act under warranty for defects law and to enable recourse under liability law.

Example 1:

A company or public body changes the intended purpose of an AI system already placed on the market or put into operation by a third party, which does not pose a high risk, in such a way that the modified AI system becomes a high-risk AI system.

Example 2:

A distributor who considers or has reason to believe that a third-party high-risk AI system that it has made available on the market does not comply with the requirements of the EU AI Act shall take necessary measures to bring that system into compliance, withdraw or recall it, or ensure that the supplier, the importer or any other relevant actor, as appropriate, takes appropriate corrective action.

Classification of AI systems

The EU AI Act divides AI systems into several categories based on their risk characteristics, which are either permitted without restriction, permitted under certain conditions or prohibited.

This approach is based on an assessment of the expected risks that a particular AI system may pose to the health, safety or fundamental rights of EU citizens.

Prohibited AI systems

High-risk AI systems

- Risk management system
- Data governance
- Technical documentation
- Record-keeping
- Transparency obligations
- Human oversight
- Accuracy
- Robustness and cyber security

GPAI systems

All GPAI (General Purpose AI)

- Technical documentation, instructions for use, Compliance with the Copyright Directive and publication of a summary of the training content

GPAI with systemic risks

- Model evaluation, tracking, documenting and reporting about serious incidents, cyber security

Minimal-risk AI systems

AI systems with additional transparency requirements

AI systems with **unacceptable risks** are prohibited per se. These include, for example, systems for the subliminal manipulation of people or certain biometric real-time recognition systems in public spaces.

High-risk AI systems are at the centre of the regulation. These are permitted, but in order to develop and use them, they must comply with comprehensive documentation, monitoring and quality requirements. The group of high-risk AI systems comprises use cases defined by the Commission, some of which are sector-specific and presented on the next page.

High-risk AI systems according to Annex I and III

High-risk AI systems according to Annex I

Section A – High-risk AI systems for which flexibility in the implementation of compliance is granted to avoid double regulatory burden:

- Machines
- Toy safety
- Recreational craft and personal watercraft
- Lifts and safety components for lifts
- Equipment and protective systems in potentially explosive atmospheres
- Radio equipment
- Pressure equipment
- Ropeways
- Personal protective equipment
- Appliances for burning gaseous fuels
- Medical devices
- In-vitro diagnostics

Section B – High-risk AI systems that are regulated by sector but exempt from most requirements:

- Civil aviation security
- Two-, three- and four-wheeled vehicles
- Agricultural and forestry vehicles
- Motor vehicles and motor vehicle trailers as well as systems, components and separate technical units for these vehicles
- Marine equipment
- Interoperability of the railway system

High-risk AI systems according to Annex III

- Biometric systems
- Critical infrastructure
- Education
- Human Resources
- Essential private and public services, including finance and insurance
- Law enforcement
- Migration, asylum and border controls
- Administration of justice and democratic processes

AI systems with a general purpose of use (GPAI) were most recently included in the regulation. Providers of GPAI systems[7] must create technical documentation, provide instructions for use, draw up guidelines for compliance with EU copyright law and publish a summary of the content used for training. Providers of GPAI systems that pose a systemic risk should also carry out standardised model evaluations such as adversarial testing, mitigate systemic risks, track and report serious incidents and ensure cybersecurity protection.

Deployers and providers of certain AI systems with a direct or indirect impact on human end users are also subject to **increased transparency obligations**. This includes systems that are designed to interact with people or expose natural persons to biometric categorisation, for example. Certain systems that are used to generate or manipulate images, video, sound or text are also affected. Deployers and providers are obliged to ensure the transparency of these AI systems towards end users in future.

AI systems that do not fall into any of the aforementioned risk categories are considered to have a **minimal risk**. They are permitted without additional requirements. However, the EU Commission is authorised to adapt the list of regulated AI systems and add further use cases.

[7] Some providers of GPAI systems with a free and open license and without systemic risks are partially excluded.

Challenges of a legally compliant risk classification

The correct, i.e., legally compliant, allocation of AI systems to one of the risk categories of the EU AI Act is a decisive factor in avoiding the threat of massive fines due to missing or insufficient compliance. In addition to its role in AI value creation, it also plays a key role in defining the requirements that must be met by the organisation and the AI system in accordance with the EU AI Act.

However, one example illustrates that it can be challenging for organisations to classify use cases in a legally compliant way: In the face of advancing digitalisation, employers are increasingly opting to use systems that enable targeted advertising and efficient recruitment of qualified employees. Given the broad definition of AI systems in the EU AI Act as well as the definition for classifying use cases in the application and selection process as high-risk AI systems, the following question will arise in the future: Is the use of supporting software in these areas already subject to the comprehensive requirements of the EU AI Act?

The EU AI Act itself does not provide for a concretisation of this definition. Rather, it will be important – following the case law of the European Court of Justice (ECJ) – to determine the classification into risk categories per the purpose and recitals of the legislator, relevant case law, and recognised methods of interpretation to enable a selective and judicially robust classification.

Legally compliant risk classification is relevant not only in the context of the EU AI Act but also because of the proposal for the EU AI Liability Directive published in September 2022. This regulates the liability of deployers, distributors and manufacturers for damage caused by AI systems and takes up the risk levels of the EU AI Act. The classification thus also has consequences for the assessment of liability risks. Those who misclassify AI systems must consequently expect both fines and civil liability in the event of missing or unreliable AI governance and documentation.

AI Liability Directive

The proposal to harmonise liability rules for damage caused by the use of artificial intelligence builds on the definitions established in the EU AI Act. The Commission assumes that the proof of damage caused by the breach of a statutory duty is partly impossible or, at any rate, associated with considerable difficulties for injured parties of AI systems. To counteract this, the EU AI Liability Directive provides that causality should be presumed under certain conditions. In particular, the risk classification under the EU AI Act plays a role. A right of access to information and, thus, evidence in cases involving high-risk AI systems flanks the facilitation of proof of causality.

Requirements for high-risk AI systems

The EU AI Act requirements for high-risk AI systems are extensive and pose complex challenges for organisations. The following overview shows the key requirement areas for high-risk AI systems under the EU AI Act:

Challenges in implementing functional AI governance along the lines of the EU AI Act

Due to the heterogeneity of AI systems, the requirements in the EU AI Act have been deliberately formulated in a general and abstract manner by the European legislator. The major challenge in practice is to operationalise these requirements so that they can be seamlessly integrated into technical and organisational processes and, at the same time, provide legal certainty.

To implement the regulation in an innovation-friendly and efficient manner, the requirements should be integrated or merged as far as possible into existing processes and structures, such as IT compliance systems. If this is not done, redundant structures and processes will increase personnel and IT-related expenses. Moreover, there is a risk of inconsistencies in assessment and documentation, and acceptance of compliance with the legal requirements in IT and operational business units will decline.

An example of an overlap between existing compliance processes and the future regulations of the EU AI Act is the assessment of risks: While the EU AI Act requires an assessment of expected risks to health, safety, or fundamental rights, the GDPR requires a “data protection impact assessment” in the event of expected high risks to the rights and freedoms of data subjects.

From a corporate social governance perspective, a data ethics risk assessment may also need to be added to address whether such processing is warranted. All these legal risk assessments should be bundled into one process, embedded in a – possibly overarching – compliance management system and subject to a technical assessment with a consistent benchmark. This allows existing processes and documentation to be used and aligned with each other, thus avoiding risks of incongruence associated with compliance and creating efficiency.

There is also more work to be done by the EU institutions in the context of standardisation if they want to create regulation that promotes innovation as promised. At the level of a concrete AI use case, the reality of AI systems and regulatory ideas can quickly diverge. For human oversight, for example, the EU AI Act stipulates that it takes at least two people to verify the results of biometric identification with AI systems. However, whether humans are actually able to improve monitoring and decision-making in the relevant fields of application is questionable. Therefore, the specifications work against automation and efficiency enhancement by AI systems in such cases.

3.

Holistic AI governance for compliance and quality

Combining high-quality AI systems and AI compliance is the key to scaling and, thus, to the success of sustainable value creation with AI systems. Both aspects require recognised standards, best practices, and appropriate tools. The goal: fast, secure, and efficient development and operation of AI systems.

In the following, we want to outline an AI compliance management system that can be used step-by-step to build up organisation-specific AI governance that combines both aspects. To this end, we draw on concepts and principles from existing compliance management systems, which we blend with our experience and the requirements of the EU AI Act.

Cornerstones of compliance management systems

Essential preparation

Firstly, for any organisation that develops, operates, uses, imports, or distributes AI systems, it is important to prepare the setup of necessary compliance and governance components and develop a clear picture of the measures required for its use cases.

Example: An organisation can align its AI governance with existing IT governance processes and integrate it into these. Use case management can be derived from IT demand and portfolio management. In turn, requirements for high-risk AI systems can be integrated into the software development lifecycle. Of course, AI specifics must be taken into account here.

Qualification & value orientation

Fundamental to successful AI governance is not only the development and documentation of new processes and structures but also their actual implementation by qualified people.

As explained above, a lack of know-how is one of the biggest barriers to digital transformation. Training concepts that establish a common knowledge and value base concerning artificial intelligence among all stakeholders and achieve buy-in for the AI-induced transformation offer a remedy here.

In addition to the content imparted to participants, a continuous exchange between stakeholders is particularly important. Especially between departments that are far removed regarding responsibilities, such as technical AI development and the legal department, suitable workshops can help build bridges in terms of knowledge and generate targeted interactions for the permanent involvement of the respective persons. Based on cross-functional training, it is easier to find formats later in AI projects where different stakeholders jointly advance the development, operationalisation, and control of the AI systems.

Impact assessment

Identifying all AI projects in an organisation and their status (e.g., planning, development, and operation) forms the basis for all further decisions in establishing appropriate AI governance. If organisations have subsequently classified risks and assigned roles according to the requirements of the EU AI Act, an impact assessment for individual AI use cases can help better assess the impact of AI regulation on the organisation. In addition to the EU AI Act, other horizontal and sectoral regulations should also be considered.

Gap assessment

Carrying out a comparison of the existing structures and processes with the previously identified compliance requirements makes it possible to determine concrete work packages. A gap assessment also helps to identify overlaps from different regulations and plan a uniform implementation. Experience shows that many existing structures and processes can be expanded to include AI-specific measures, such as risk management, data management or cyber security.

Compliance strategy

The final step of the preparation is defining a compliance strategy to establish a holistic AI governance and compliance management system. This includes a list of concrete measures for establishing AI governance and closing previously identified gaps, the associated efforts, and a corresponding prioritisation. The prioritisation, on the one hand, results from the organisation-specific goals in terms of compliance, quality, and scaling the identified AI systems and, on the other hand, from a risk/impact analysis to define which regulations and business areas should be focused on first. This also marks the beginning of the implementation of the requirements of the EU AI Act.

Modules of holistic AI governance

Meeting all the requirements of the EU AI Act and other regulations is an undertaking that should not be underestimated. Therefore, it is worthwhile for all organisations to divide the implementation into smaller work packages and initially focus on governance components that bring immediate added value beyond the core compliance purpose. The implementation can follow an agile approach in which the organisation works on different components at different intensities simultaneously – always based on the organisational goals, the identified compliance risks, and time restrictions.

Example: A company can fall back on established structures in data management and, therefore, design effective processes relatively quickly. On the other hand, regarding documentation, it is better to start early with standardisation and collection – but only when it becomes necessary, for example, to push ahead with compilation and preparation for an external audit.

The modular approach has many benefits:

- Quickly improving governance of AI initiatives.
- Enhancing the quality of AI systems from the beginning.
- Successive and iterative work towards compliance with relevant requirements.

In the following sections, we present several attractive modules of holistic AI governance for compliant and high-quality AI systems and highlight the specific challenges in their implementation.

Roles and responsibility structure

The EU AI Act requires the establishment of role and responsibility concepts for high-risk AI systems. It remains unclear how these shall be concretely designed. It is, therefore, a good start to fall back on best practices from other fields for concrete implementation.

The necessary design varies between organisations, but the areas of responsibility and roles within AI systems are often similar in practice. The basis for structured processes: clear delineation of individual task areas and roles. Secondly, responsibilities for the individual task areas must be defined to avoid diffusion of responsibility and provide clear contact persons for internal and external stakeholders. In the future, this will also include auditors and regulators.

If organisations can link work steps and roles, they can also reduce cases where entire process chains collapse, e.g., due to the departure of essential people because it was not clearly defined which areas of responsibility needed to be re-staffed. Depending on a risk assessment, high-risk AI systems also lend themselves to concepts that place additional requirements on the scope of staffing and representation of critical roles to ensure greater security and reliability.

In role and responsibility concepts, restrictions are necessary to meet specific security requirements, depending on the area of application. For example, in data management, it is (mandatory) to establish processes for granting and removing access rights to protect sensitive data. In principle, the actual necessity is decisive for granting such access rights (“need-to-know” basis). However, this is not always easy to determine in practice and must, therefore, be clearly defined for decision-makers.

Risk management

With the use of AI, an organisation's risk management requirements must be adapted. The EU AI Act requires high-risk AI systems to establish, apply and document a risk management system (RMS). This RMS can be understood as a continuous and iterative process that is carried out throughout the lifecycle of a high-risk AI system and requires regular, systematic updates. These can occur in defined cycles and on specific occasions, such as new findings from monitoring or user feedback. Furthermore, developing sufficient measures to prevent, mitigate, and reduce the risks to an acceptable level is required.

The basis of adequate risk management is a systematic risk analysis. It serves to identify sources of risks, determine probabilities of occurrence, and quantify impacts. In the context of AI, this area of risk management probably poses the greatest challenges, as identifying scenarios and their effects requires a great deal of expertise. From degradation of performance metrics due to the changing characteristics of input data to discrimination due to underrepresentation, there are many ways in which the health, safety or fundamental rights of EU citizens or the environment can be put at risk. In some cases, a so-called "Fundamental Rights Impact Assessment" is also required by the applying organisation.

It remains to be seen how standardisation in AI risk management will develop in the EU. Through international regulatory exchange, especially with the US, there is at least a willingness to harmonise AI standards in many areas. This would allow the EU to build on existing international and national standards. For example, in January 2023, the NIST (US federal agency) published the "Artificial Intelligence Risk Management Framework (AI RMF 1.0)", the first national framework for AI risk management. The risk management guidelines from the joint ISO and IEC technical committee on artificial intelligence are also relevant (ISO/IEC 23894:2023).

For some use cases, standardisation has already reached a higher level of maturity and detail and can at least provide conceptual guidance for future frameworks. One example is test scenarios for autonomous vehicles (ISO 34502:2022), which can help with standardised testing and subsequent approval. However, by no means will all organisations benefit from sector-specific standards as in the automotive sector. They cannot avoid determining detailed risk scenarios and probabilities of occurrence for their use cases. Organisations will, therefore, continue to depend on the know-how of (domain) experts for a comprehensive risk analysis.

Data management

High-quality data is essential for the development of good AI systems. Data management in the EU AI Act is, therefore, primarily about taking appropriate steps to achieve high data quality and protect data from access by unauthorised persons.

An enormous challenge is the processing of data sets during development and operation. Thus, the EU AI Act calls for specific and comprehensive quality criteria and safeguards. This results in several requirements: Data sets must be assessed regularly and transparently. Organisations must identify and define permissible and, where appropriate, desired biases. They must check the availability, quantity, and suitability of the data sets and possible discrimination. Last but not least, they must identify potential data gaps or deficiencies.

However, the EU AI Act does not further specify these requirements. Regulated organisations must determine, document, and constantly check the legal minimum for the respective high-risk AI system in question. This is associated with legal uncertainty in individual cases and may lead to liability risks.

Any data sets used and processed must also comply with data protection requirements in case they contain personal data (IP addresses, names, IDs, etc.). These requirements must be verifiably met at an early stage, in some cases even before the development of the AI system begins. Regardless of whether the data is collected by the provider of the AI system, provided by a third party, or even made available (publicly) by the data subject, compliance must be proven. In addition, if copyright-protected content (e.g., texts, music) is used, it may be necessary to obtain corresponding rights of use in advance.

The requirements of the GDPR extend not only to the AI system development phase (for example, with training data) but to the entire lifecycle of an AI system – i.e., for the input and further use as well as deletion and transfer of personal data. Organisations must evaluate data protection risks on a case-by-case basis for each form of processing of personal data in AI systems. The challenge in this context is the potential of AI systems with the principle of combining data protection through data protection-friendly default settings (“data protection by design”). Here, the newly adopted standard ISO 31700, “Privacy by design for consumer goods and services”, can provide guidance.

Particularly delicate: Responsibilities for compliance with data protection law and AI law regulations can diverge. This has implications for different areas of an organisation. While the EU AI Act mainly addresses natural or legal persons, the GDPR links the responsibility to the body that determines the purposes and means of the data protection-related processing activity. Thus, it is conceivable that the use of a high-risk AI system in a group requires the parent company’s compliance with the requirements of the EU AI Act while, at the same time, determining the means and purposes of processing.

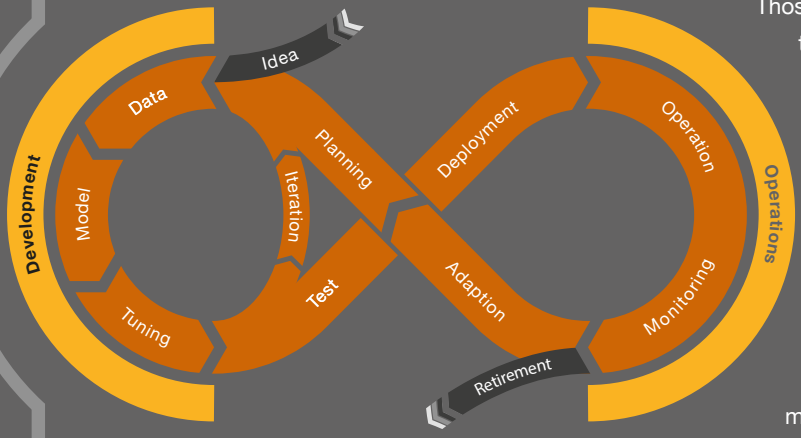
Personal employee data is primarily determined by the HR department of the subsidiary, whereby the subsidiary is qualified as the data protection controller. Such cases must be identified and documented to place the obligations of the EU AI Act in personnel terms.

Those who deal with the legal aspects early can better assess and, at best, reduce any liability risks. This includes, above all, drafting contracts with providers of AI systems or actors along the supply chain in an advantageous way.

Lifecycle management

The requirements of the EU AI Act are intended to ensure high-quality levels for all high-risk AI systems in the European market. Therefore, the regulation places particular emphasis on structured processes for the design, development, review and monitoring of AI systems. Furthermore, it requires that high-risk AI systems remain robust and accurate throughout their lifecycle. Combined with the already described requirements for responsibility management, data management and recording of logs and metadata, this results in the connection of abstract regulatory requirements to proven workflows and architectures of so-called Machine Learning Operations (MLOps).[9]

Trustworthiness and compliance, at their core, require the implementation of best practices throughout the lifecycle of an AI system. The fundamental purpose of MLOps frameworks and technologies is to build structured process flows that can manage AI systems from initiation to operation – in some cases, even fully automatically. Those who successfully combine methods from the fields of data management, software development and machine learning in MLOps enjoy far-reaching advantages for the quality and scaling of AI systems. Structured workflows make it significantly easier to develop and operate AI systems that meet the demands of deployers, regulators, and providers.



Organisations can implement lifecycle management with MLOps in a very practical way and thus create immediate added value for the scaling and quality assurance of AI systems. Prerequisite: To do this, they must be able to draw on a combination of competencies from the fields of AI development, software development, DevOps, and data technology/science.

The specific challenges for high-risk AI providers are primarily to implement proven MLOps architectures for all their systems, to align specific principles, components, roles, and processes with the requirements of the EU AI Act and then to document them fully. An example of a necessary extension is the control and test processes for individual case-specific quality assurance, which can be carried out at different points in the workflow.

The storage and monitoring of meta, input and output data is also extremely valuable, as it enables extended performance analyses and a higher degree of traceability. This pays off in risk management and AI liability and, together with the feedback loops in MLOps, represents an advantage for the continuous and rapid improvement of AI systems that should not be underestimated.

[9] Kreuzberger, Kühl und Hirschl (2022): Machine Learning Operations (MLOps): Overview, Definition, and Architecture

Enablement of (end) users

The EU AI Act explicitly requires users to be enabled to utilise the respective AI system. Furthermore, the EU AI Act obliges system providers to provide additional transparency to people who interact with the AI system and use its results. Analogous to the data protection notices and consent to the processing of cookies (“opt-in”) via so-called cookie banners, which many are now familiar with from every website, AI banners and declarations could also help to implement these requirements and thus become widespread in the future. This is because natural persons need to know that they are interacting with an AI system, who is providing it (for any queries), what it does, what its technical characteristics are (e.g., the level of performance), what is expected of users and how to interpret the system’s results. This, too, will ultimately have an impact on any liability issues. However, the EU AI Act does not make any concrete specifications in this regard.

Getting these aspects right poses several linguistic, technical, and economic challenges. To counteract fundamental uncertainties about AI systems among end users, the linguistic design must strike a balance between easy-to-understand language and concise texts and content that is technically and legally sufficient. The technical and economic challenges lie in explaining the results. It is not always easy to present the connection between an AI system’s input and output. There are solutions for so-called “explainable AI”, but here, too, the information must be adapted to the expected level of knowledge of the end user. On the other hand, the intellectual property of the AI provider must be protected accordingly, and no more than necessary must be revealed about the inner workings of an AI system.

Conformity management

To put high-risk systems into operation, onto the market or use their output in the EU market, regulated organisations must not only achieve compliance with the AI Act but be able to demonstrate it. This means that based on comprehensive documentation of the individual governance components (“technical documentation”), compliance must be assessed. Depending on the use case, an internal control system or an external body authorised to assess conformity can be used. Subsequently, the compliant high-risk AI system must be labelled accordingly and registered in an EU database. In addition, an AI provider has an obligation towards the relevant authorities to demonstrate a declaration of conformity and to provide access to data and documentation in case of justified requests. If a system is not compliant or there are incidents in operation, the relevant authorities must be informed and compliance restored.

The path to compliance with the EU AI Act is a team effort. Important contributions must be made along a wide range of functions and roles.

	1st Line of Defence	2nd Line of Defence	3rd Line of Defence
Units	Data Scientists Machine Learning Developers MLOps Engineers	Legal Department Compliance Department	Internal Audit External Audit
Tasks	Implementation at use case level Documentation	Definition of legal requirements Establishing controls Standards and specifications and Best Practices	Independent assessment of conformity with the EU AI Act

While compliance departments are often responsible for laying out the requirements for documenting systems, the teams providing technical support for the AI system must prepare and regularly update the documentation. The definition of control sets also usually falls within the scope of the so-called Second Line of Defence. However, technical and functional teams provide evidence of compliance with the controls. Internal departments or external notified bodies, which have an independent view of the evidence and the controls, then carry out the final evaluation.

One organisational challenge is to distribute responsibility optimally. The goal: a truthful and positive conformity assessment must be carried out, and obligations must be fulfilled. Although documentation tasks are often postponed, organisations should not avoid preparing the documents early on and in a disciplined manner because ex-post preparation is, in any case, more time-consuming and, in some cases, even impossible. So-called model maps, data sheets and other technical system information offer added value that goes beyond regulatory compliance. In our experience, this also simplifies the consolidation of best practices in the development process and the transition to scaled operations.

4. Conclusion and outlook

The EU AI Act will have a massive impact on the development, use and commercialisation of AI systems in the coming years. Organisations along the entire AI value chain must act now.

Early implementation of holistic AI governance and compliance management systems gives organisations not only a time advantage over the competition but also an immediate economic advantage through shortened time-to-market and high quality of their high-risk AI systems. The requirements are complex and present companies with new kinds of challenges that will take time to overcome.

Additionally, the expected shortage of AI (governance) experts in the transitional period will likely become a significant cost driver. Organisations with high-risk use cases that do not want to wait several years to benefit from using AI systems in regulated areas should, therefore, seize their opportunity now. If, in addition, it is possible to utilise as many synergies as possible with existing compliance structures and fall back on best practices for machine learning operations, organisations can build up the desired AI governance and compliance quickly and efficiently.

Companies and public organisations can set the course now to shape the digital transformation in Europe successfully with the use of AI. The ability to combine quality, compliance and scaling of AI systems will significantly determine the success of organisations in the European market – especially against the backdrop of technical developments by competitors in Asia and the USA. Linking these aspects requires interdisciplinary competencies in organisations along the entire AI value chain. This is how to create structures and processes for an AI governance that is technically, legally, and organisationally fit for the future.

Your contacts

EU AI Act Governance:
Hendrik Reese
Partner
Responsible AI

+49 151 70423201
hendrik.reese@pwc.com

EU AI Act Legal:
Dr. Jan-Peter Ohrtmann
Partner
Data Protection

+49 171 7614597
jan-peter.ohrtmann@pwc.com

Banking sector:
Konstantin Dagianis
Partner
Financial Services (Banks)

+49 171 9770067
konstantinos.dagianis@pwc.com

Insurance sector:
David Richter
Partner
Financial Services (Insurances)

+49 1511 0578093
david.richter@pwc.com

pwc

© March 2024 PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft. All rights reserved.

In this document, "PwC" refers to PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft, which is a member firm of PricewaterhouseCoopers International Limited (PwCIL). Each member firm of PwCIL is a separate and independent legal entity.

PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft adheres to the PwC-Ethikgrundsätze/PwC Code of Conduct (available in German at www.pwc.de/de/ethikcode) and to the Ten Principles of the UN Global Compact (available in German and English at www.globalcompact.de).