Workstream AI Governance

White Paper

Authors

Furkan Eser
Birgit Bartz
Sebastian Döll
Eric Joachim Liese

In cooperation with other CBA Lab member companies and HTWG Konstanz – University of Applied Sciences

Contents

1. Introduction

1.1 Putting the hype in perspective

With the popularity of Stable Diffusion, MidJourney, and ChatGPT, the excitement and buzz surrounding artificial intelligence (AI) has reached a high point. It's not just software manufacturers like Microsoft that are adding AI capabilities to their products (for example with GitHub Copilot or M365 Copilot, Salesforce with Einstein AI), or the many public cloud providers with expanded AI services, as companies in the mechanical engineering sector and other manufacturing industries are also equipping their products and processes with AI capabilities. They do this in the hope that such an approach will give them a competitive edge and increase their efficiency. However, using AI without thinking hard about what you actually want to do with it[1] poses substantial risks that are often overlooked.

Such risks include potential damage to a company's reputation, negative effects on corporate culture and communication, liability risks, and risks relating to high investment costs[2].

In order to ensure successful long-term use of AI, it is important to determine the actual added value AI will create as well as the actual need for AI in the specific context of the company in question. A company thus needs to analyze its processes and requirements in detail and deploy AI in those areas where it can lead to real improvements that will benefit both the company and its customers. Instead of simply buying into the hype, companies should select a strategic and well-considered approach for implementing AI.

For example, it is absolutely crucial that a comprehensive AI governance system be introduced in order to successfully integrate AI into a company and avoid the high costs that are incurred when planning is ill-considered, not strategic, or not in line with regulatory requirements. Such a governance system creates a clear framework for the development, utilization, and monitoring of AI systems and ensures that these systems are aligned with ethical standards, regulatory requirements, and the company's strategic goals. The governance system defines areas of responsibility, identifies risks in a timely manner, and develops risk control measures for the company.

[1] Fear of missing out (FOMO) phenomenon.
[2] https://www.blackstone.com/insights/article/the-convergence-of-data-centers-and-power-a-generational-investment-opportunity-the-connection/

1.2 The need for AI governance

Along with defining the general objectives for an AI governance system, the participants also worked together to formulate the following specific goals:

- **Effective assessment and management of use cases:** A central goal of any AI governance approach is to ensure that AI use cases can be effectively assessed and implemented and then continuously managed. This requires a detailed analysis of the return on investment (ROI) and timely and comprehensive risk assessment as early as during the demand management process. The dimension of assessment must be continuously reviewed throughout the entire lifecycle to ensure that the use case in question will contribute to the success of the company over the long term.
- **Risk management regarding special aspects of AI:** AI governance must ensure that specific risks that arise through the use of AI are quickly recognized, effectively reduced, and reliably monitored. This includes an individual risk assessment as early as the demand management stage. The risks are to be discussed at the beginning of the process in order to be able to make sound decisions regarding the actual investment, or develop a risk minimization strategy. This initial risk analysis is documented and monitored throughout the entire lifecycle of the given use case.
- **Making use cases available to the entire company:** A further goal is to define a clear and effective way to make AI use cases available to the entire company. This means that AI use cases and their parameters must be documented and published as best practices in order to ensure that all stakeholders have access to the relevant information. In addition, automated processes – for monitoring, for example – should be established in order to ensure smooth and efficient use of AI systems in daily business operations.
- **Use and adjustment of existing governance processes:** Finally, an AI governance system should be designed in a manner that allows it to make use of existing governance processes at a company, and also make adjustments where necessary. Functioning and established processes, such as those for data management or innovation, can serve as a basis for the implementation of AI systems. This ensures that the integration of AI can be seamlessly embedded into the existing corporate structure and that necessary adjustments can be made with regard to matters such as strategy, compliance, and monitoring. (See Section 4 as well)

1.3 Opportunities that artificial intelligence offers companies

The implementation of artificial intelligence (AI) offers companies many opportunities, such as:

- **Efficiency gains through automation:** AI can automate repetitive and time-consuming processes, which in turn results in a significant increase in efficiency.
- **Personalized customer interaction:** The analysis of customer data can be used to create customized products and services that increase customer loyalty and satisfaction.
- **Supply chain optimization:** AI analyzes complex data in realtime and recognizes patterns that can help optimize supply chain processes, inventory management, and processes for predicting fluctuations in demand.
- **Support for innovation:** AI creates new possibilities for developing products and services and makes it possible to enter new business sectors.
- **Improved decision making:** AI-supported data analysis enables sound and precise decisions to be made.
- **Risk management and fraud detection:** Predictive maintenance helps minimize machine and facility downtime and detect fraud at an early stage.
- **Increase in employee productivity:** AI can help employees increase their productivity and focus on activities that create value.

It cannot be assumed at the moment that processes will be managed completely autonomously with AI solutions, or that AI represents a stand-alone product. Instead, it can be assumed that processes and products will be enriched by AI.

The potential AI offers can only be fully exploited if the use of AI is accompanied by the implementation of a comprehensive and well-thought-out governance strategy and system. This will in fact significantly increase the available opportunities by taking into account not only the direct technological benefits that result but also the long-term strategic advantages that AI will lead to.

- **Risk minimization and ensuring compliance:** A clear governance structure helps minimize legal risks and ensure that regulatory provisions are consistently complied with.
- **Establishing trust among stakeholders:** Successful AI governance establishes trust among customers, investors, and the public by ensuring that AI is used in a responsible and ethical manner.
- **Competitive advantages with ethics and responsibility:** Companies can position themselves as pioneers for responsible AI and thus gain a competitive edge.
- **Promoting innovation and growth:** Stable governance is the foundation for the safe and efficient use of AI, which in turn promotes innovation and growth over the long term.

- **Efficient use of resources:** Clear guidelines and processes help ensure human and financial resources are utilized more efficiently, which improves a company's overall performance.
- **Long-term sustainability and resilience:** Continuous adaptation in line with new technologies and regulatory requirements ensures sustained success and the avoidance of liability.

Companies that use AI without a clear governance system risk running into a variety of problems:

- **Economic dimension:** A lack of AI governance can lead to inefficient processes, financial losses, and failed projects. There is also a risk of damage to a company's reputation, as well as high liability payments that could threaten the economic stability of the company.
- **Ethical and social dimension:** Uncontrolled AI systems can reinforce discriminating decisions, violate privacy, and lead to a loss of trust among customers and employees. Without clear ethical guidelines, there is a high risk that AI systems will lead to unethical results, or even to results that would be damaging to society.
- **Strategic dimension:** A lack of AI governance can put a company in a strategically unfavorable position because the company will not be able to efficiently use the AI technologies to strengthen its competitiveness. Clear governance creates internal synergies, increases efficiency, and enables a company to react flexibly to market changes and new regulatory requirements, which in turn will allow it to expand its innovative capability in a targeted manner.
- **Technological dimension:** Without a structured governance system, there is a risk that AI systems will operate on the basis of poor-quality data and will become unreliable and also vulnerable to attacks. In addition, an inability to integrate AI systems into existing systems, as well as scaling and maintenance problems, will threaten long-term technological performance capability.
- **Regulatory dimension:** Companies also run the risk of violating legal provisions, which can lead to problems with certifications, approvals, and compliance. This in turn can have legal consequences in the form of sanctions, as well as negative effects in terms of market viability and competitiveness.

1.4 Types of AI

Artificial intelligence refers to the ability of machines and software to achieve a level of intelligence performance that is similar to that of human beings. This includes the ability to learn, to solve complex multilevel problems, and to interact with the surrounding environment. However, AI is more than just ChatGPT, Microsoft 365 Copilot, or GitHub Copilot. These solutions are examples of generative AI. There are also other solutions, however, such as AlphaFold3, with the help of which

[3] https://alphafold.ebi.ac.uk/

its researchers won the 2024 Nobel Prize in Chemistry and which itself is an example of a predictive AI system. These types of AI fall within the category of machine learning (ML).

The different types of AI also require different approaches and technologies for using them.

Predictive AI

This type of AI is used to analyze non-linear connections and relationships in data and then make predictions regarding future events or types of behavior on the basis of the analyses. A typical example involves predicting the need for maintenance work or predicting attack patterns in web traffic. However, this type of AI is also used in connection with financial analyses, customer behavior analyses, and risk management approaches in order to improve products or business processes. It is designed to support decisions and provide strategic insight.

Generative AI

This type of AI specializes in creating new content such as texts, images, or music. It uses technologies like neural networks in order to produce creative results for a very wide range of applications. Generative AI is now used in many areas – for example in programming (creating program text) and in business processes (summarizing e-mails or creating presentations). A distinction is often made between multimodal and unimodal models for AI. A multimodal model is an ML model that can process information from a variety of modes, including images, video, and text. Unimodal models, on the other hand, usually can only process information from a text source. In other words, AI is not always the same thing in all cases, so to speak – there are different types that can be used for different applications. The hype surrounding AI that emerged after OpenAI's introduction of GPT-3 focused especially on the fact that machines were now able to perform even the most complex tasks, including those previously unfamiliar to them. However, even though this generative AI has become more and more effective, it is still not able to act completely autonomously.

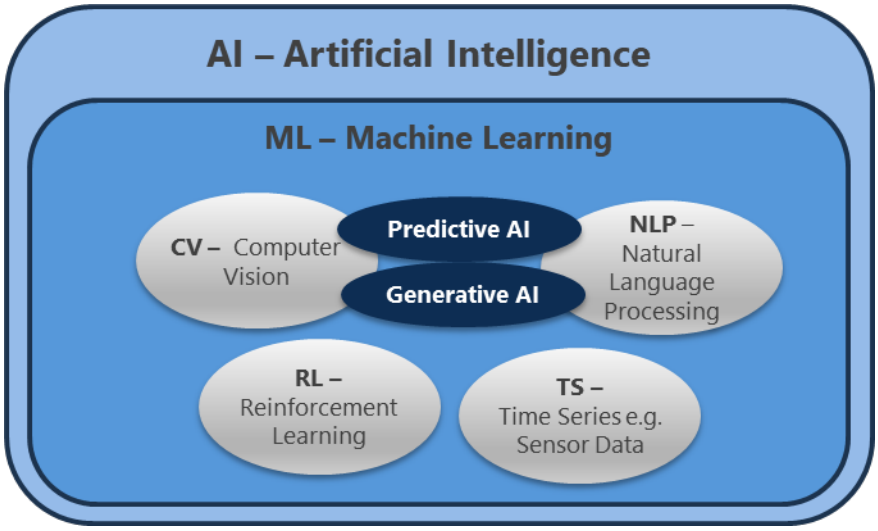

Figure1: Artificial intelligence (AI) and machine learning (ML)

2. Methodology for creating an AI governance system

The creation of an AI governance system forms the foundation for responsible and efficient use of artificial intelligence at a company. The various requirements and conditions at a company must first be examined and relevant aspects must then be identified on the basis of the analysis. The resulting governance ecosystem specific to the company in question will then establish the framework for the specification of an AI governance approach and its characteristics. In the workstream, the individual dimensions of the ecosystem and the core elements and attributes of AI governance and its dependencies were brought together in an AI governance model and depicted methodologically.

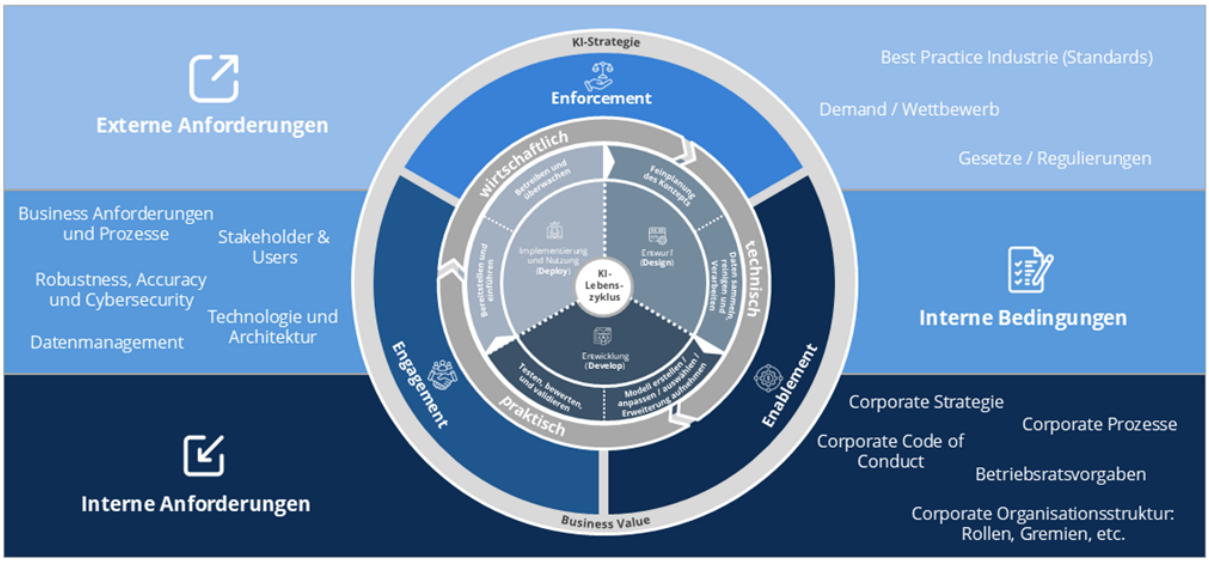

Figure2: AI governance ecosystem

2.1 The AI governance ecosystem

The AI governance system forms the foundation for responsible and efficient use of artificial intelligence at a company. This ecosystem consists of various requirements that must be seamlessly integrated into a company's strategy, structure, and processes.

The requirements in turn are divided into three central dimensions: **external requirements**, **internal requirements**, and **internal conditions**.

External requirements in the AI governance ecosystem relate to the external framework conditions that companies must take into account in order to be able to comply with legal provisions and manage competition and market requirements. The framework conditions here include best practices in the given industry, market demand and competition, and laws and regulations.

- **Best practices in the industry:** Industry standards and tried-and-tested processes and procedures offer guidance for developing safe, ethical, and effective AI systems.
- **Market demand and competition:** The pressure to develop innovative and marketable AI solutions forces companies to design AI systems in a manner that makes them trustworthy, efficient, and user-friendly.
- **Laws and regulations:** National and international laws and regulations ensure that AI systems meet certain minimum requirements.

Internal conditions are the technical and organizational conditions companies need to establish in order to be able to function on the market. These include:

- **Business requirements, and business processes:** These define the business goals and targets that AI systems are to help achieve, and how processes are to be adapted in line with this.
- **Stakeholders and users:** The needs and expectations of stakeholders and end users need to be taken into account in order to be able to provide user-centered AI solutions that will be accepted.
- **Robustness, precision, and cybersecurity:** AI systems must be resistant to attacks, deliver reliable results, and meet strict security standards.
- **Technology and architecture:** A solid technology infrastructure and a clear system architecture are needed in order to be able to implement AI solutions in an efficient and scalable manner.
- **Data management:** Effective data management is crucial for ensuring that high-quality, secure, and relevant data can be made available for AI operations.

Internal requirements relate to the organizational, strategic, and ethical framework that companies have defined internally and which must guide the AI governance approach. These include:

1. **Corporate strategy:** The overall strategy for the company determines the focus of the AI strategy and the AI governance strategy.
2. **Corporate code of conduct:** Ethical guidelines / common values define how AI systems can be developed and deployed in a responsible manner and in line with the company's values.
3. **Corporate processes:** On the one hand, corporate and business processes must be designed in such a manner that they can support AI and react in an efficient way to the integration of new technologies. At the same time, it must be possible for AI governance processes to use established corporate and business processes and build upon them.

4. **Works council stipulations:** Works council stipulations ensure that the use of AI also takes into account labor law and codetermination aspects.
5. **Corporate organization structure:** Clear roles and responsibilities and corporate bodies ensure effective management and control of AI deployment at a company.

The internal requirements and external requirements define how companies need to plan, implement, and monitor and control the use of AI in their organization. These requirements must be clearly incorporated into the strategy and the development of capabilities, as well as into roles and responsibilities, compliance structures, tools, technologies and processes, policies, documents, and guidelines. The important thing here is that companies must adjust and further develop these requirements individually and in line with the specific situation in their organization.

With regard to incorporating the requirements, it is important that this is not done in an isolated manner or only in certain departments or units. Instead, experts from the various key corporate functions (e.g. Legal, Technology, Compliance, HR, and Strategy) should work closely together. This type of interdisciplinary cooperation ensures that the requirements are understood and implemented in a holistic manner. Close cooperation also ensures that the requirements are smoothly and efficiently implemented in normal operations, without any risks to the work of the development teams caused by unnecessary delays due to complex bureaucratic processes. In this way, strategic and operational aspects can be aligned in an optimal manner and pave the way for innovations to be developed, tested, and scaled with as few obstacles as possible.

2.2Attributes of AI governance

The **AI lifecycle** shown in the center of the graphical depiction of the AI governance model symbolizes the continuous development and iteration of AI systems. Around this core are grouped the three central attributes that characterize the internal conditions at a company as the key principles and drivers:

- **Economic**: This attribute relates to the cost-benefit analysis and the assessment of the financial return that an AI system offers. It is crucial that investments generate income over the long term and lead to efficiency gains. Important governance elements can be derived from the “economic” attribute. Ensuring that an AI system will generate a financial return requires careful planning with a long-term perspective. It is important here to focus not only on direct benefits (e.g. greater efficiency) but also on potential costs for implementation and maintenance, for example, as well as possible risks that could arise in connection with legal provisions and ethical considerations.
- **Technical:** Within the framework of this attribute, the feasibility, performance capability, and integration of AI systems into existing infrastructure is examined. A successful system must be reliable, scalable, and adaptable.

Technical feasibility plays a key role in the assessment and development of AI systems and is therefore an important attribute of AI governance, as it mainly involves the management of access to governance. Issues such as infrastructure, scalability, and adaptability are incorporated here.

- **Practical**: This attribute is focused on user friendliness and the acceptance of AI systems by stakeholders. A practical system must meet the needs of users if it is to be successfully implemented and utilized.

  This also involves an analysis of the various stakeholders and users. Systems that are technically mature but also complicated in terms of their operation may meet with resistance, with the result that they will not be used as much as they should be. As specified in the AI Act, AI management approaches must also ensure that systems are not only efficient, but also easy to understand and adjust in line with different user groups.

2.3The core functions in the governance model

Three core governance functions make it possible for us to ensure sustainable and responsible AI implementation in line with the internal conditions and with due consideration of external requirements and internal requirements:

- **Engagement**: It is crucial that an active community be created that promotes the exchange of knowledge and experiences. The AI Office plays a key role here by coordinating initiatives and organizing training courses in order to increase awareness of AI technologies and promote acceptance of the same.

  In this connection, active engagement and knowledge management at a company are particularly important, as there is often a lack of central knowledge at companies regarding the interface between technology, compliance, and legal affairs departments and other units that interact with these. It's therefore very important to have a proactive community in place that manages the following derivative actions: Support of AI development as regards compliance with various regulations and their complex requirements, establishment of a change strategy and awareness-raising and training measures in order to allay any fears or anxiety among the general workforce, development of AI capabilities and measures to ensure compliance with regulatory provisions when all the different AI instruments are used.

- **Enablement**: This involves creating a solid infrastructure that supports the use of AI systems. Key aspects here include technology resources, data management, and training programs that expand employee capabilities and expertise.

  A solid technological foundation ensures that AI systems run smoothly and are scalable, while high-quality data management results in AI applications that work with reliable and representative data. Training and further education for employees makes it possible for staff to not only understand the technology but also actively use it to create value and develop

and implement innovations. With regard to governance, enablement also relates to other aspects such as prioritizing flexible and adaptable technologies – as technological progress is moving so fast that organizations have problems keeping up to date at all times. Training programs represent another derivative action for ensuring that employees not only obtain theoretical knowledge but also the practical skills that enable them to develop and use AI systems responsibly.

- **Enforcement**: Enforcement ensures that all AI applications comply with ethical standards and legal provisions. The implementation of control mechanisms and audits makes stakeholders more trusting of AI technologies and promotes the responsible use of these.

  The principle of enforcement is essential for ensuring that AI systems are not only effective but can also stand up to quality and security audits and are in compliance with legal provisions. Although the implementation of audits and control mechanisms provides a solid foundation, this alone is not sufficient, as audits are often conducted only periodically, which means that risks occurring between audits might be overlooked. Derivative actions here are mainly carried out within the framework of compliance checks and measures to strengthen accountability and responsibility.

Around all of this, the outer realm encloses the **business value**, which makes clear the overall utility of AI and its strategic importance to the company.

2.4Preventing shadow AI, and adaptability of AI governance

Effective governance requires the clear assignment of domains of responsibility that not only focus on products or business processes but are also arranged in line with specific areas. This model involves the division of governance domains on the basis of **conformity, security, and operation**. In the case of conventional internally provided AI models, the focus is on compliance with regulatory provisions and on ensuring data quality and the robustness of the systems. At the organizational level, on the other hand, there are additional requirements, such as preventing so-called shadow AI and ensuring that the content created meets ethical and legal standards.

Shadow AI refers to the use of AI systems outside the official governance structures at a company – i.e. without adequate monitoring of security and compliance. This would be the case, for example, if a department were to introduce an unapproved AI application for automating work processes without checking the application's data protection aspects or searching for potential security risks. This can actually occur relatively often in these times of low code, no-code, and API and browser applications. Such behavior must absolutely be prevented using a variety of measures, as the consequences that shadow AI can have for a company can be severe.

An awareness program for employees is essential for ensuring that only tested and approved AI applications will be used at a company. It must be made clear that only tested systems may be deployed, and that self-developed applications must be tested before they can be used. At the same

time, testing processes must be designed in a lean manner that also conserves resources in order to promote a culture of experimentation and be able to maintain trust in employees' own sense of responsibility.

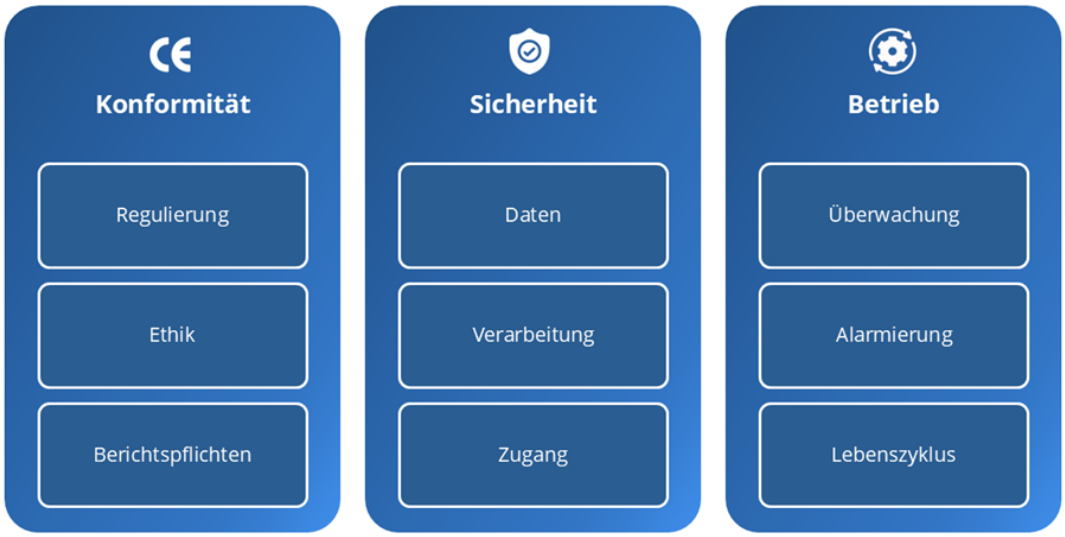

Figure3: Regulation, security, operation

2.5 Types of AI systems the AI governance model focuses on

1. Self-developed models or adapted external models that have been integrated into a company's products and services.
2. Distributed AI systems that remain unchanged and are made directly available to the market.
3. AI systems or service components that a company only uses internally in order to meet certain operational or business requirements.

All three system types are subject to the same governance mechanisms in order to ensure security, compliance, and ethical use.

3. AI governance artifacts

AI governance artifacts are central components (essential documents, tools, and methods) that are absolutely required if an AI governance system is to be established at a company. These artifacts

are the foundation for managing and monitoring AI solutions. They are targeted at the unique challenges that arise at companies that deploy artificial intelligence.

AI governance requires a broader-based approach than traditional IT governance. It takes into account both technical complexity and the ethical implications of using AI, as well as the continuous adaptability of the systems (even after launch) and the need to establish new forms of risk management.

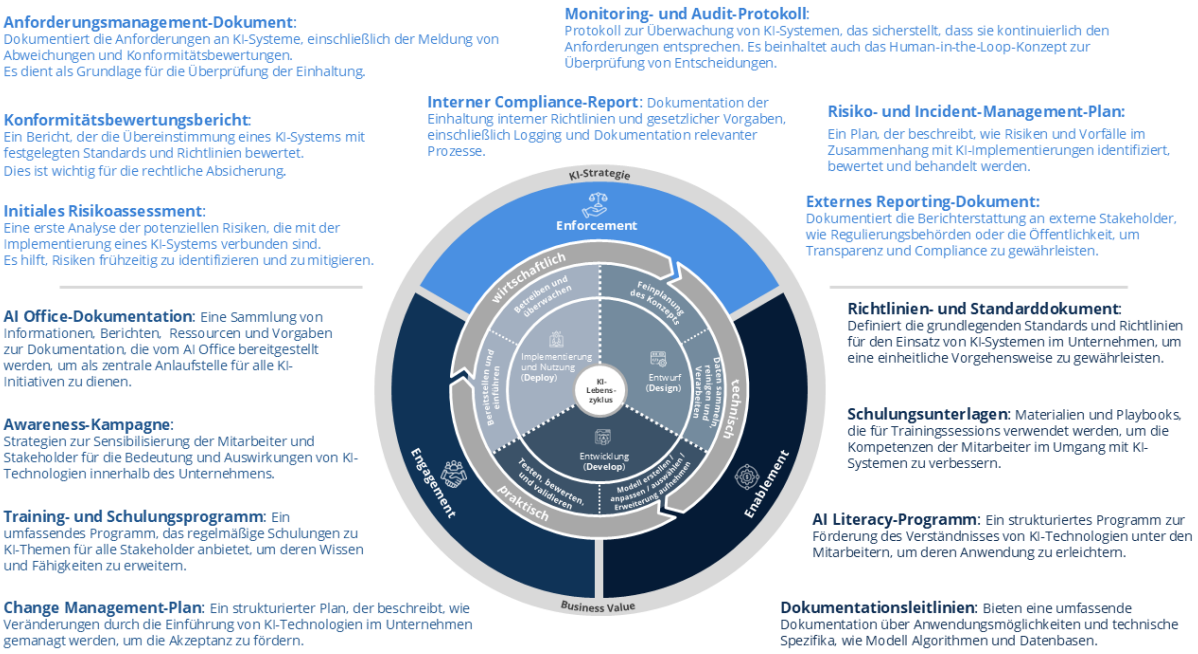

Figure4: Governance artifacts

In accordance with the principle of holistic and iterative processes for AI governance, which requires periodic checks and updates, these artifacts also need to be continuously maintained and updated.

The process has three areas where AI governance has an effect on a company: Enforcement, enablement, and engagement. All AI governance activities can be assigned to one of these “effect factors.” Either certain provisions need to be made and complied with (enforcement), decision-making assistance needs to be provided (enablement), or decisions regarding the use of AI solutions need to be supported.

3.1 Artifacts for enforcement of AI governance

Requirements management document: This document lists all requirements that an AI system must meet, including a notification obligation when deviations occur, and a conformity assessment

as well. The document serves as a reference for checking compliance with internal and external provisions and forms the basis for structured implementation of the requirements.

Conformity assessment report: This report documents the conformance of an AI system with defined standards and rules. It thus also serves as the basis for reports to external stakeholders such as government agencies, for example.

Initial risk assessment: An initial analysis of the potential risks associated with the implementation of an AI system. This early-stage risk report helps a company identify potential threats and then take suitable measures to minimize these risks, which in turn ensures that development teams can address potential risks at an early stage and mitigate risks on a technical basis in their fundamental decision making. It also helps make the use of resources for compliance and security checks more effective, and might also make it possible to benefit from best practices at the company.

Monitoring and audit log: This log documents the monitoring of AI systems to ensure that they continuously meet the defined requirements when they operate. It also involves a "human-in-the-loop" approach that ensures important decisions are reviewed by people, whereby this approach must be adequately aligned with the risk and sensibility of the given use case.

Internal compliance report: The compliance report ensures compliance with internal guidelines and legal provisions. It encompasses logging and documentation of all relevant processes and activities and thus ensures seamless traceability of activities and decisions. These reports need to be archived in line with the given legal stipulations.

Risk and incident management plan: This plan describes how risks and incidents in connection with AI implementation and use arise, and how they are identified, assessed, and managed. Depending on the sensitivity of the application, this plan must be evaluated and adjusted, and conditions for the rapid analysis of such plans must be established as well.

External reporting document: The external reporting document ensures effective reporting to external stakeholders such as government agencies or the public. It establishes transparency and helps ensure requirements relating to compliance are met.

3.2Artifacts for engagement in AI governance

AI Office documentation: A central collection of information, reports, resources, and stipulations that is made available by the AI Officeso that the AI Office can serve as a central point of contact for all AI initiatives at the company. This brings together not only the artifacts from the area of enforcement but also supportive and helpful information such as best practices, transparency obligations, GTC modules, etc.

Awareness campaigns: A strategy for sensitizing employees and stakeholders with regard to the significance and effects of AI technologies at a company. Awareness campaigns ensure better understanding and acceptance of AI systems at a company, work to reduce fear and anxiety, and help establish company-wide legal certainty in order to make it easier for development teams to do their jobs.

Training and further education programs: A comprehensive training program that offers courses on various AI topics and issues for all stakeholders on a regular basis. The knowledge and capabilities of employees should be continuously expanded, and it must be ensured that everyone is familiar with the latest developments, the applications that are used, and all best practices. This will also allay fears and anxiety at the company and make it possible for all employees to learn how to automate (partial) aspects of their work in a meaningful way on the horizontal level as well.

Change management plan: A structured plan that describes how changes brought about by the introduction of AI technology to a company can be managed. The goal here is to promote the acceptance and successful implementation of AI systems and to be able to scale this as well.

3.3Artifacts for enablement of AI governance

Guidelines and standards document: Defines the basic standards and guidelines regarding the use of AI systems at a company in order to ensure a uniform approach at the company-wide level or in certain defined sectors and application areas.

Training documents: Materials and playbooks that are used for training modules in order to increase employee expertise as regards their use of AI systems.

AI literacy program: A structured program to promote understanding of AI technologies among employees in order to make it easier for employees to use such technologies. Depending on the use case in question, it may be necessary to create a dedicated AI literacy program for users of particularly sensitive applications.

Documentation guidelines: Offer comprehensive documentation of application possibilities and specific technical aspects regarding how models, algorithms, and databases are to be used in AI projects.

4. AI governance and existing processes at a company

Companies generally have numerous established processes that have proven themselves over a period of many years and provide for a solid foundation. When AI technologies are to be implemented, it is important that these processes be carefully reviewed, and in some cases redesigned as well. The aim here is to avoid the duplication of work and unnecessary bureaucratic obstacles that can arise as a result of uncoordinated parallel structures and workarounds. AI innovations are horizontal issues that affect all departments and business units, which is why

existing efficient processes need to be adjusted in line with the new requirements associated with AI development and implementation.

It is important here not to view AI as an isolated system. Instead, the interfaces between the various departments, units, etc. at a company need to be redefined and must work together in a more closely coordinated manner. Existing structures represent a valuable foundation for implementing the new requirements. It is crucial that the existing structures and their processes be optimized in order to be able to exploit the full value of AI throughout the company and ensure that all units work together smoothly and efficiently.

The central company processes presented below must be optimized and adjusted in line with the new challenges:

4.1 Strategy and planning

The strategy and the associated principles define the framework within which AI systems are developed and used at a company. They determine the long-term focus and prioritization of AI initiatives and ensure that these further the achievement of the overriding goals at the company. The principles serve as guidelines that all operational decisions relating to AI must take into account.

Why do strategies and principles have to be changed?

The introduction of AI technologies at a company leads to fundamental changes with regard to strategic planning requirements. AI initiatives are often complex and interdisciplinary and affect different departments and units, and frequently those that have had little experience with digitalization to date. A clear strategy is important in order to be able to select use cases in a targeted manner and maximize their utility for the benefit of the entire company. Without such a strategic focus, there is a danger that AI projects will become isolated or be conducted without a clear goal, which can lead to inefficient use of resources and missed opportunities.

The implementation of AI systems also necessitates the introduction of new decision-making principles that take into account specific challenges such as the selection of technical infrastructure (“cloud first” versus “build before buy”) and the definition of architecture principles. These principles must be flexible and adaptable.

- **Alignment of the AI strategy with the corporate strategy:** The selection and prioritization of the use cases must be aligned with the overriding AI strategy and the company’s strategy.
- **Reuse and/or adjustment of existing principles:** Already established strategic principles should be reviewed and used again wherever possible. It must be ensured that the

company's principles – e.g. agility or sustainability – are also complied with when AI is introduced.

- **Development of new decision-making principles:** There is a very great need for clear architecture principles in the AI context because the mostly small-scale initiatives undertaken at a company can benefit from standardized best practices and principles. Decisions such as "cloud first" or "build before buy" must be made at an early stage in order to define a clear strategic direction and simplify technical implementation.

4.2Demand process

The demand process refers to the structural path that is used to identify and assess new ideas or innovation projects at a company and then transfer them to the implementation phase.

Why is it necessary to make changes to the demand process?

The introduction of AI technologies must take into account additional factors that previously did not play a key role in traditional innovation processes.

Above all, the risk involved must be recognized and thoroughly assessed in an early phase of the process. AI projects harbor potentially higher risks in terms of ethical questions, data security, and liability, among other things. These risks should not first be identified and analyzed in late project phases but instead from the very beginning if sound decisions are to be made.

It's also crucial that compliance costs be integrated into the process at an early stage. The demand process must ensure that all regulatory requirements relating to AI are clearly defined and that compliance with these requirements is taken into account as part of the project costs. These compliance costs flow directly into the calculation of the return on investment (ROI) and are a key factor in the process for assessing the economic feasibility of an AI project. Not taking these costs into account at an early stage poses a risk in that projects might be underestimated with regard to their financing requirement, and ROI calculations may turn out to be unrealistically optimistic.

Adjusting the demand process:

- **Transparency regarding compliance costs at an early stage:** In order to be able to make sound decisions, compliance costs must be calculated in detail and assessed as early as the design phase.
- **Implementation in accordance with value and resources:** The demand process prioritizes projects on the basis of their expected value contribution and the available resources.

- **Risk and role classification:** A structured risk analysis, especially in the case of high-risk models, is used to introduce a dual control principle that helps clearly define areas of responsibility and minimize risks.
- **Iterative review:** Projects and their risks, as well as costs, need to be reviewed and reassessed on a regular basis in order to be able to react flexibly to changes.

4.3 Compliance process

The compliance process includes all steps taken to ensure that a company complies with legal provisions, regulatory requirements, and internal guidelines.

Why is it necessary to make changes to the compliance process?

The introduction of AI technologies leads to new and complex regulatory requirements that can vary greatly depending on the risk class of the AI system in question. AI systems with high risks in particular need to meet strict requirements, and compliance with these must be monitored and assessed on a regular basis. An AI system that is not in compliance can lead to substantial legal and financial difficulties. The early incorporation of compliance requirements can also help development teams at least partially minimize the risk level of the AI model in question. For example, dividing a task into individual steps with models separated from one another can lead to a situation in which only models that affect significant decisions are considered to be high risk, which in turn can substantially reduce compliance costs for the entire application.

Adjusting the compliance process:

- **Early communication about requirements:** All compliance requirements must be clearly communicated to the relevant departments in order to ensure consistent implementation or develop a minimization strategy.
- **Transparency regarding compliance requirements:** All relevant compliance requirements must be clearly defined and integrated into the process at the very beginning of the AI lifecycle. This makes it possible to assess risks at an early stage, and it also reduces any delays that might occur further on down the line.
- **Obligatory conformity assessment and audit processes:** The compliance process requires continuous monitoring of the systems, supported by audits and periodic conformity assessments, in order to ensure compliance with provisions – after the AI introduction as well.

4.4Solution design

Solution design relates to the process in which a technical solution – in this case an AI system – is designed, assessed, and prepared for implementation.

Why is it necessary to make changes to the solution design?

The introduction of AI systems leads to unique challenges that extend beyond traditional software development processes (e.g. in relation to bias, fairness, transparency, ...). An insufficient design can lead to unanticipated risks such as distorted results or serious security problems.

The solution design must also ensure that cybersecurity and the scalability of the AI system is adequately taken into account. In addition, the design must be structured in such a manner as to enable the system to also function reliably in different environments and under conditions that vary.

This means that the system needs to be continuously monitored and it must be ensured that the defined criteria for performance, security, and transparency are met. Clear lines and areas of responsibility, as well as a functioning incident management system, also need to be in place in order to be able to respond rapidly and efficiently to problems or security incidents.

- **Governance tools:** Tools such as Collibra or LeanIX should be used to continuously monitor the health, service health, and security of the AI system.
- **Analyze risks at an early stage:** Consideration of aspects relating to bias and fairness at an early stage
- **Performance assessment and scalability:** Clear metrics for performance assessments must be defined and these metrics must be aligned with the specific use case. This also includes an assessment of scalability and consideration of environmental constraints.
- **Cybersecurity:** In order to minimize potential threats, suitable security mechanisms aligned with the specific attack vectors associated with the system must be incorporated into the system. Security measures must be a part of the basic design and should not be added later on down the line.
- **Roles, responsibilities, and incident management:** Clear areas of responsibility must be defined in order to safeguard implementation of both the security and performance requirements and a structured and effective incident management system.

4.5 Monitoring / Operation management

Monitoring and operation management relates to the continuous monitoring and maintenance of AI systems during ongoing operations.

Why is it necessary to make changes to monitoring / operation management?

AI models as we define them continue to develop on an ongoing basis even after they are implemented. A lack of effective monitoring can lead to a deterioration of robustness or the quality of results, and there is also a risk that the models might behave in an unanticipated or undesired

manner. This is especially difficult to recognize if such distortions creep in slowly. In addition, a monitoring strategy for high-risk models is an absolute must according to the AI Act.

All key stakeholders – such as members of operation teams, compliance departments, and management – must always have access to the latest information about the condition and performance of AI systems. This in turn requires custom metrics tailored to the specific requirements of the company and the use case in question.

- **Continuous monitoring:** The results of AI models must be continuously monitored with regard to quality and behavior. This also involves ensuring the robustness and accuracy of the models and their compliance with defined performance metrics.
- **Custom metrics and audit:** Tailored metrics need to be developed and used to monitor the performance, security, and conformity of the models. Audit trails must also be set up in order to document all essential decisions and model changes in an understandable manner. These metrics, as well as the extensity and intensity of the monitoring and the audits, should always be designed and implemented with the use case adequately taken into account.
- **Incorporation into the existing company infrastructure:** For example into company processes such as ticketing systems, health scores, and reporting – dashboards
- **Death Switch:** A Death Switch should be installed so that a model can automatically be deactivated in the event of a severe error. The process to activate this must be explained and made available to all stakeholders in a clear manner.
- **Reporting to the AI Office:** Reporting to the AI Office ensures that all relevant data and incidents are documented, which in turn enables decisions to be made on the basis of reliable data and ensures that any required adjustments can be implemented in a timely and strategic manner.

5. New requirements and capabilities that are needed to ensure successful implementation of AI

When AI technologies are introduced at a company, new requirements arise and new capabilities are created that previously did not exist in traditional processes, or else weren't needed in the form that they will be needed in connection with AI.

At the superordinate level, the following capabilities must be developed at a company in order to ensure effective management of AI applications and the associated regulations:

- Expertise in addressing regulatory requirements (AI Act)
- AI-specific risk assessment expertise
- Ability to cooperate in an interdisciplinary manner, in particular in terms of effective cooperation between compliance/legal departments and the development teams. The way the compliance monitoring processes and tools are designed is crucial here because bureaucratic obstacles, complicated processes, and tools that cannot be understood or are difficult to use can slow down or obstruct the innovation process.

A detailed description of all these requirements and the capabilities that are needed was compiled in the context of the EU AI Act analysis and can be viewed in the AI Act Deep Dive.

6. Ethics

Companies are expected to conduct themselves in an ethical manner because this strengthens trust among the public and promotes long-term customer loyalty[4]. Societal expectations and regulatory requirements such as the EU AI Act ensure compliance with ethical standards. Companies that conduct themselves in an ethical manner are less vulnerable to legal and financial risks and benefit economically over the long term[5]. Moreover, consumers and talented potential employees are now paying more attention to the way companies behave, and they favor those that act responsibly[6]. Ethical conduct not only makes a decisive contribution to improving a company's reputation and making it more stable; it is also a key factor when it comes to obtaining and retaining highly

[4] Bhattacharya, C. B., & Sen, S. (2020). Sustainability: How stakeholder engagement leads to competitive advantage. *Journal of Business Ethics*, 161(2), 317-329.
[5] Eccles, R. G., Ioannou, I., & Serafeim, G. (2019). Corporate sustainability: A strategy? *Management Science*, 65(12), 5661-5680.
[6] Jones, D. A., Willness, C. R., & Heller, K. W. (2019). Corporate social responsibility attributions and employee engagement: The role of perceived external prestige and internal respect. *Journal of Business and Psychology*, 34(2), 239-252.

qualified specialists who increasingly expect to be able to work in an environment marked by compliance with ethical standards[7].

In view of such increasing expectations regarding ethical conduct, which is demanded not only by society at large but also by the regulatory framework (e.g. the AI Act), companies need to have clear ethical guidelines and frameworks in order to safeguard trust and stability and thus avoid liability issues and a loss of reputation. Ethical guidelines play a key role in this context because they govern the use and management of AI systems in a manner that satisfies societal and regulatory expectations.

Ethical principles relating to the use and management of artificial intelligence can vary greatly depending on the region and cultural context in question. Whereas in Europe the focus is on the protection of individual rights (e.g. privacy and data sovereignty), China is pursuing an approach that emphasizes social stability and control over technological developments, for example. These differences are reflected in the respective ethical frameworks that define how AI is to be used.

The European Union has formulated specific guidelines for trustworthy AI that are to serve as the basis for the responsible use of AI technologies. The guidelines address key aspects such as human agency and oversight, technical robustness and safety, privacy and data governance, and ensuring transparency. They also emphasize the need for diversity, fairness, and social responsibility.

This framework offers a solid foundation for companies to ensure that their AI systems not only function properly but are also ethically sound. The European guidelines are supplemented by additional national and global ethical frameworks that can serve as a guide for companies.

Incorporating ethical principles into governance

Incorporating ethical principles into specific AI applications presents a special challenge, as it is not sufficient to define ethical principles on a general level; instead, they must be translated into specific use cases.

In order to effectively manage ethical risks in a corporate governance system, a specific special process needs to be introduced that focuses on AI applications that display a high reputational risk potential. This special process is meant to ensure that particularly sensitive projects are extensively analyzed in order to identify and address potential problems at an early stage.

A special governance process can also be established for particularly sensitive AI projects that potentially pose high ethical risks. The process begins with a **check** in the planning phase that uses a risk assessment or targeted questions to identify critical projects in terms of their ethical risks. After that, an **ethics assessment** is made by having a specialized ethics team or board conduct a detailed analysis that evaluates the ethical implications of a project. The **results** of this assessment

[7] Gond, J. P., El Akremi, A., Swaen, V., & Babu, N. (2017). The psychological microfoundations of corporate social responsibility: A person-centric systematic review. *Journal of Organizational Behavior*, 38(2), 225-246.

are then documented and integrated into project planning in the form of recommendations or binding provisions. This process ensures that ethical considerations are systematically taken into account and that risks are addressed in a timely manner.

Important frameworks for identifying and assessing the risk associated with artificial intelligence

- **ISO/IEC JTC 1/SC 42:** International standards for the implementation of AI systems, with a focus on ethics and fairness.
- **AI RMF (AI Risk Management Framework):** Developed by NIST, this framework provides guidelines on security, fairness, and transparency.
- **EU Ethics Guidelines for Trustworthy Artificial Intelligence:** Focus on transparency, responsibility, and data protection.
- **ASAM:** Assesses the risks associated with algorithmic systems and the effects they have on society.
- **AI Incident Database (AIID):** Documents AI system malfunctions and risks.
- **FAT/ML (Fairness, Accountability, and Transparency in Machine Learning):** Tools for ensuring fairness and transparency.
- **IEEE Ethically Aligned Design:** Develops ethical principles for AI development and implementation.

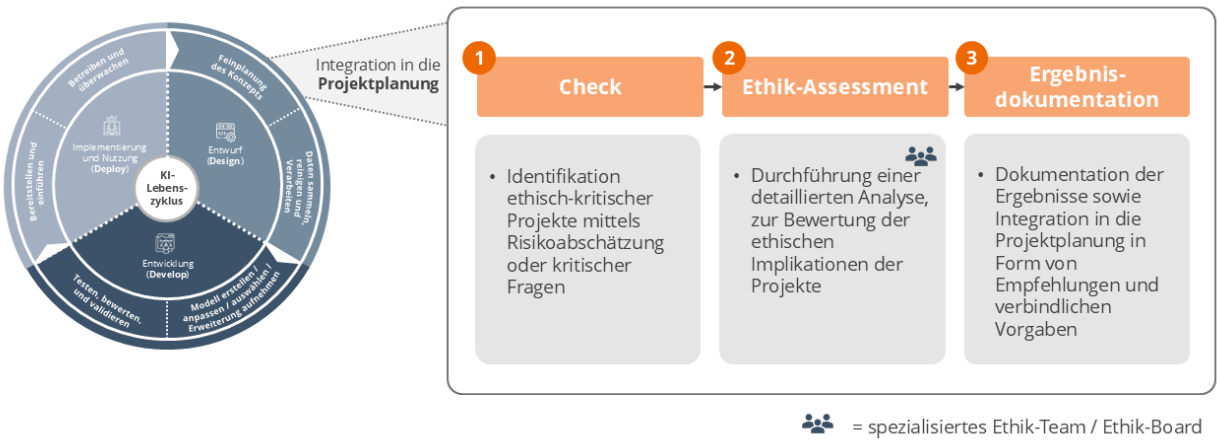

Figure5: Special governance process

7. Regulatory aspects – EU AI Act Deep Dive

Effective AI governance must be based on relevant regulations that influence the use of AI systems. Below we take a closer look at the AI Act, although it is only one of many regulations that apply to AI

systems in one way or another. Along with data protection and data security provisions, sector-relevant or locally relevant stipulations also need to be taken into account, whereby these can vary depending on the use case in question.

The analysis below relates to the *Artificial Intelligence Act, Official Journal version of 13 June 2024*[8] and all amendments made up until 1 August 2024.

7.1 The EU AI Act

The AI Act, which went into effect on 1 August 2024, provides the first comprehensive legal framework for artificial intelligence in the European Union. The goal of the act is to ensure secure, transparent, and ethically responsible use of AI systems. For companies that develop, deploy, or use AI systems on the European market, the act requires that their AI applications be designed and used in line with strict regulations, especially if such applications have been assessed as being high-risk.

The AI Act is being implemented incrementally: Beginning in February 2025, all AI systems that are considered to pose unacceptable risks will be prohibited. Starting in August 2025, the rules for general models (General Purpose AI Systems) will go into effect and the authorities that will be responsible for these rules will be announced. The European Commission will present monitoring plans for high-risk AI systems by February 2026, and additional obligations regarding such systems in sensitive areas will become mandatory in August 2026. While this staggered implementation of the AI Act does give companies time to make adjustments, it also necessitates quick integration of all the regulations into a company's strategy. It is essential that companies which operate internationally view the AI Act in conjunction with other local and sector-specific regulations in order to ensure continuous across-the-board compliance.

The AI Act results in numerous mandatory processes, tools, and specific roles and responsibilities that need to be integrated into a company's governance framework.

The following aspects must be taken into account for every AI application at a company:

- All AI systems that fall under the definition of AI pursuant to the AI Act must be carefully examined in terms of several key factors. First, it must be determined whether the system in question does in fact correspond to the definition of AI within the meaning of the AI Act. After that, the territorial area must be considered, as the AI Act applies to both companies that provide or deploy AI systems within the EU and companies that are headquartered in a third country but whose AI results will be utilized in the EU. It is also crucial to be clear about the role a company plays in connection with the AI system – i.e. provider, developer, or user. In addition, it must be determined whether the system in question contains a General Purpose AI (GPAI) model. Finally, the application in question needs to be examined in order to determine a risk classification.

[8] EU AI Act: https://artificialintelligenceact.eu/

- Local and sector-specific regulations can have a massive influence on the development of AI models along with the AI Act and must therefore be carefully examined as well. Regulations on the EU level include the General Data Protection Regulation, the Machinery Directive, the Product Liability Directive, and the General Product Safety Directive. Depending on the sector in question, other relevant industry-specific and environmental directives and guidelines may apply.

7.2The role of a company within the meaning of the EU AI Act

In general, three different types of AI models currently exist at companies, and each of these are associated with different governance requirements in accordance with the AI Act. These three types can be identified on the basis of the associated degree of responsibility in the supply chain. A company's responsibility varies according to the role (AI Act) it plays in terms of how it uses and manages the AI model in question. These roles are what determine access to governance, in particular with regard to the registration and risk categorization of the AI systems.

1. **Provider**: A company that develops the AI model itself, makes significant changes to an existing model (e.g. fine tuning), or distributes the model under its own name or brand name.

   The AI Act defines a provider as any person or entity that develops an AI system, or has one developed, and makes it available with the aim of putting it on the market or into operation. A company in this role bears full responsibility for compliance with all regulatory provisions and system conformity.

2. **Distributor**: A company that resells or hosts an AI model without making any significant changes to it becomes a distributor.

   According to the AI Act, a distributor makes an AI system available without making any fundamental changes to it. Above all, this role encompasses ensuring compliance, but there is less technical responsibility involved than is the case with a provider.

3. **Deployer**: A company that uses an AI model in a business context, for example to automate processes or for interaction with customers, becomes a deployer.

   The AI Act defines a deployer as any person or entity that utilizes an AI system, particularly within the framework of a commercial activity. This role requires the company in question to monitor use of the system and ensure that it is operated in line with the intended purpose and in a secure manner.
   The challenge for companies is to make it possible to integrate an AI system into existing systems and tools used in work processes – for example online collaboration tools. It is often the case that end users rather than the purchasing or procurement department is informed about this. Because companies must perform a risk assessment for all AI applications in their value chains, processes for risk assessment of all AI systems must be developed and a sufficient level of awareness must be established among all employees with regard to

utilizing these processes. (See “New requirements and capabilities that are needed to ensure successful implementation of AI”)

Figure6: Roles

All scenarios require clear access to governance, which also includes registration and risk categorization in accordance with the AI Act, as well as documentation of compliance-related measures.

The AI Act makes a distinction between AI models and AI systems, whereby the two concepts lead to different governance requirements. An AI model consists of the algorithms and technologies that are developed for specific tasks but which are generally not provided directly to end users. Instead, the model is integrated into an AI system that also includes additional components such as user interfaces and a technical infrastructure in order to enable the model to be used in normal business operations or in daily life.

An example of this distinction is offered by systems such as ChatGPT or Microsoft Copilot, which are based on the GPT-4 model from OpenAI. Whereas the GPT-4 model takes care of the actual calculation work, AI systems like ChatGPT create the connection to end users by providing a user-friendly interface.

Governance systems need to be flexibly designed in order to effectively regulate both AI models and AI systems. Governance systems must be able to ensure control over the development and use of AI models, particularly in environments in which both purchased and internally developed models are deployed.

7.3Determining requirements on the basis of the AI Act

What is regulated:

According to the AI ACT, AI is a "system based on machine learning or other technologies that is designed to perform, with varying levels of autonomy, tasks that would normally require human intelligence to complete." Systems that meet the criteria of this definition therefore need to be regulated.

Which different types of AI systems are regulated?

Along with the definition of AI shown above, there is also a definition for the concept known as GPAI:

GPAI models – or General Purpose AI Models – are defined as AI models that:

- Are trained on a major scale with large amounts of data and with the use of self-monitoring learning techniques
- Are extremely general in nature, meaning they can competently execute a large number of different tasks
- Can be integrated into a large number of downstream systems or applications

Timeline for implementation:

The AI Act is being implemented in several phases in order to give companies sufficient time to make adjustments:

- **August 2024:** The AI Act officially goes into effect.
- **February 2025:** Prohibition of all AI systems that are considered to pose an "unacceptable risk." This means that companies need to take an inventory of all AI systems so that they can remove from the market those AI systems that are subject to the prohibition.
- **May 2025:** Completion of a Code of Practice for General Purpose AI.
- **August 2025:** Application of the rules for General Purpose AI and announcement regarding the authorities that will be responsible for these rules. This means companies will need to ensure that their GPAI models have a CE marking and can be used for the intended application.
- **February 2026:** Presentation of the specifications of the guidelines for the practical implementation of the AI Act, including monitoring plans for high-risk AI systems.
- **August 2026:** Application of the obligations for high-risk AI systems in specific areas. This means companies must have implemented a complete AI governance system in line with EU standards by this time.
- **August 2027:** High-risk AI systems that are deployed in critical areas with regard to safety/security must implement additional safety/security components. This means companies will need to expand their AI governance systems

- **End of 2030:** Obligations relating to the large-scale deployment of certain AI systems in IT infrastructures.

How is regulation carried out:

First it must be determined whether an AI system falls within the **territorial scope** of the AI Act:

Provision and placement on the market in the EU:

- The AI Act applies to AI systems that are offered on the European market or put into operation there, regardless of where they were developed.

Use and impact within the EU:

- AI systems that are developed or made available outside the EU are also subject to the AI Act if the results produced by these systems are used in the EU or if the system has an influence on people or companies in the EU.

In the next step, an analysis must be conducted to determine whether an AI system has anything to do with the **prohibited practices** described in Article 5, or whether the system is an exception.

Three other aspects must also be analyzed:

1. Whether the system was developed on the basis of one or several GPAI Models (Article 3 (63)).
2. The role the company in question plays (Article 3 (3/4)).
3. The risk classification the application should be assigned to (Article 6, Annex I & II, Article 50).

This analysis can be used to derive the requirements that should be applied to the use case in question.

7.4 Resulting requirements for companies in order to comply with regulatory provisions

The following capabilities must be developed at a company in order to ensure effective use and management of AI applications and compliance with the associated regulations:

Definition of the (regulatory) requirements

The risk classification of an AI system determines the scope of compliance management activities:

- The deployment of AI systems, as well as the requirements laid out in the relevant regulations (e.g the AI Act), make it necessary to gain a clear understanding of compliance stipulations as early as at the beginning of the AI lifecycle so as to ensure that effort, expense, and resources can be reliably estimated at the start, or that the risk can be minimized by means of various (technical/architectural) decisions.
- If an external AI model is used, it also must be ensured that this model is in line with the defined compliance requirements. License conditions, and in the future the CE marking for models as well, need to be checked thoroughly to ensure that the system is legally sound and complies with all relevant regulations with regard to its intended use.

Requirements and capabilities:

- Standardization of the risk assessment: First a possibility must be created for a standardized evaluation of the extent of the risk or risks and the documentation of the results.
- Examination and validation of external AI models: This includes gaining an understanding of the license conditions and the legal framework, evaluating the CE marking, and ensuring that the model is in line with legal stipulations and the compliance requirements.
- Teams must be able to analyze and combine technical and legal requirements and develop strategic plans and decisions for minimizing risks, which must then be incorporated into project planning. After that, the most realistic estimation possible of resource requirements must be generated on this basis.

Internal development of General Purpose AI Models versus external service

If use cases are based on a GPAI Model, specific model requirements need to be met along with the requirements relating to risk class and the definition of roles. In the future, these additional requirements will either need to be met by the company itself, or else compliance will have to be documented via the external model provider through license verification procedures. This is especially important because providers will be able to exclude their models for certain use cases or risk classes in the future if the associated requirements are not met.

Companies must take comprehensive measures to ensure compliance with all license requirements in connection with AI models from external providers. The models deployed also need to be examined with regard to their suitability for the intended use case and the given risk class. This requires a detailed evaluation of the technical specifications and legal requirements in connection with the model in question.

Requirements and capabilities:

- If companies choose to deploy their own internally developed systems, they need to develop the capabilities, processes, and tools that will enable them to design their models in a manner that ensures compliance with regulatory and technical requirements.

- Companies that make use of an external service must develop the ability to maintain partnerships with model providers and develop the processes needed for this. This involves aspects such as license verification procedures, contractual negotiations, an analysis as to whether the model can continue to meet the company's requirements over the long term, and possibly periodic audits as well.

Mandating

Mandating refers to the official authorization and obligation and/or permission for a company or certain persons to develop an AI application.

AI initiatives are complex and interdisciplinary, and also pose risks such as violations of data protection provisions, algorithmic distortions, and security problems. Calculating ROI for an AI application is also a complicated process. Mandating establishes a central authority or central process that authorizes development teams to develop an application and make decisions concerning it. It also makes it possible to determine whether the risks to the company can be adequately managed or mitigated and ensures that AI systems meet the defined requirements.

Risk assessment and ROI analysis: Mandating requires the integration of the risk and role assessment into the strategic planning process and the financial analysis. This means that the risk assessment must be conducted together with the analysis of the expected return on investment (ROI) and serve as the foundation for mandating decisions. Indeed, the AI ACT stipulates that AI systems are to be assigned to risk classes, and a higher risk classification means stricter compliance requirements. During the mandating process, the extent of the risk, the anticipated associated costs, and a precise definition of roles for the responsibilities that result are presented and documented.

Requirements and capabilities:

- A standardized calculation of ROI must be created that makes it possible to analyze the financial evaluation of an AI project in connection with the risk assessment.
- A central authority or process must be established that is authorized to approve AI projects or reject them. This authority is responsible for deciding whether an AI system is to be developed and deployed. It ensures that all compliance requirements, strategic goals, and risk assessments are taken fully into account and that the investment is a sound and viable one for the company.
- Precise roles and areas of responsibility must be defined for each instance of mandating. This means that the people who are to be responsible for risk assessment, compliance

monitoring, and the development of an AI system must be clearly defined. This will ensure clear lines of responsibility and authority that cover the entire process.

Conformity assessment

As a result of the AI Act and other regulatory provisions, most AI systems will be subject to specific stipulations in the future, whereby these will vary depending on the system's risk classification. A white listing will only be provided for systems with a minimal risk, while all other systems will be subject to more extensive compliance checks and related processes. Compliance with the stipulations will have to be checked before a launch in the future, and this will require the development of internal processes, although external audits might also be necessary.

Requirements and capabilities:

- Companies need to develop appropriate processes to ensure compliance across all risk classes.
- With regard to AI systems classified as being high risk, the AI Act stipulates that these need to be audited by an external entity that has been designated a notified body in order to ensure that all regulatory requirements have been met.
- Companies must ensure such compliance not only prior to the launch of an AI system but also continuously throughout the system's entire lifecycle. This requires continuous monitoring and possibly repeated conformity assessments, especially if significant changes are made to the system.
- The stipulations regarding the creation and administration of technical documentation need to be met, and the documentation system must be audit-proof in this regard.
- Capabilities and processes for cooperation with external notified bodies that are responsible for certification of high-risk AI systems: Cooperation here involves the provision of all required documents and the performance of audits in cooperation with the external auditors.
- The ability to identify significant changes to AI systems and initiate a new conformity assessment in good time if the changes made could lead to changes in the regulatory requirements.

Monitoring and examining

Many AI models, especially those that pose a high risk (e.g. GPAI Models with a systemic risk), are subject to strict stipulations, which means they need to be checked on an ongoing basis to ensure they remain compliant. A one-time examination before deployment is often not sufficient because AI systems are dynamic and must be continuously reassessed due to changes in data or the surrounding environment. Companies therefore need to establish central units that monitor risks, compliance, and possible incidents, although monitoring is also important for internal assessments.

In addition, the obligation to notify the responsible EU authorities if incidents occur also necessitates the establishment of a fast and structured incident management system.

Requirements and capabilities:

- Companies should establish a central unit or processes to monitor risks and compliance activities in connection with all AI systems. The risks should be monitored proactively in order to be able to identify potential problems at an early stage. This requires the use of specialized tools for risk monitoring that continuously check all relevant KPIs and compliance with the applicable requirements.
- Companies must implement a structured incident management system that makes it possible to immediately detect, document, and report to the responsible EU authorities all incidents that arise in connection with an AI system.

Training, awareness, and knowledge

In view of the increasing use of AI systems at companies, the AI Act stipulates that both end users and employees need to be sufficiently AI-literate in order to ensure that the technologies are used and managed properly. Effective training programs and continuous knowledge sharing also make a key contribution to promoting the use of AI in a company as part of a successful change management approach.

If employees and managers are extensively trained and understand the benefits and risks of AI technologies, as well as the application possibilities, it increases not only the acceptance of the new systems but also the trust people have with regard to their use. This in turn helps minimize opposition to the transformation by reducing feelings of insecurity and presenting specific application examples. A high level of knowledge and practically focused training ensure that the introduction of AI is not viewed as a burden but instead as an optimization opportunity.

Requirements and capabilities:

- Development of practically focused training programs aligned with employees' specific tasks and responsibilities.
- Employees and end users who come into direct contact with AI systems must be provided with knowledge about the systems' functions, as well as system restrictions and limitations. To this end, mechanisms and processes need to be developed to enable open and transparent communication about the deployment of AI technologies. This requires the ability to prepare technical information in a manner that ensures it will be understood by non-experts and lead to the right expectations.

8. The AI lifecycle and governance

The lifecycle of an AI solution encompasses all phases of an AI application – from the concept design to provision of the system and ongoing operation. The process here is continuous and iterative, which means that AI solutions are to be monitored and improved on a regular basis in order to be able to respond to potential changed demands, requirements, and data.

Overview

The design of an AI solution represents the conceptual foundation of the system. In this phase, the AI solution is planned and specified in accordance with the given requirements profile. This comprises the precise detailed planning of the solution concept and the compilation of the data needed for the application. In some cases, this data will need to be cleaned and processed beforehand in order to create an optimal foundation for development.

The individual phases are described in a systematized manner below:

- **Focus** relates to the special aspects of this phase in the lifecycle and also the areas where special attention needs to be paid.
- **Internal and external requirements** refer to AI-specific demands in the sense of these supplementing general requirements for IT systems, products, and services.
- **Affected processes** mean those processes at a company that will be influenced by the introduction of an AI solution.
- **Artifacts** are the tools, methods, documents, and instruments addressed in this White Paper. These must be developed as part of the AI governance system to be established (especially for AI applications) and should have attributes specific to a company's structure and situation.

8.1 Holistic assessment and classification of business demands

A precise and systematic demand analysis is one of the key foundations for the successful deployment and management of AI at a company. Such an analysis makes it possible to precisely assess the necessity and feasibility of AI deployments and avoid bad strategic investments. An in-depth analysis here enables both technical and organizational risks to be identified at an early stage, which ensures successful integration of AI solutions into existing business processes and leads to long-term competitive advantages. It should also be ensured here that a problem or a need will be recognized and that a realistic assessment can then be made as to whether an AI deployment actually makes sense at all. This type of early analysis ensures that AI will only be implemented if its use offers clear benefits and added value for the company. In addition, the

demand analysis helps make it possible to make realistic calculations regarding all relevant dimensions, such as costs, technical requirements, organizational impacts, and ethical considerations.

8.2 Demand management

It is difficult at the moment for companies to measure or calculate the value of an AI system. This is due to the following issues:

- The scope of influence of an AI system is almost impossible to delineate and isolate – for example it is difficult to determine whether an increase in revenue has been brought about by an improved usage experience or an AI system[9]. Many advantages such as customer satisfaction, risk management, and brand loyalty are difficult to measure monetarily, but nevertheless contribute to long-term growth[10],[11]. On the other hand, incorrect use of AI can cause damage in exactly these areas, whereby the actual influence here often isn't measurable until months or even years later.
- Implementation of AI often requires extensive investment in data infrastructure, skilled specialists, and tools, which makes the cost calculation more difficult. To this can be added high costs for maintenance and further development that can further delay ROI.[12]
- The establishment of AI governance, cybersecurity, and a suitable data management system can also initially lead to high costs.
- In addition, a special challenge here involves the fact that ROI can display a sharp deviation from the estimate if the analysis of the risks and costs that arise due to the various implications of an application turns out to have been insufficient. This would particularly affect compliance-related costs. If an application is deemed high risk, compliance-related costs can be expected to be high. Transparency at an early stage – i.e. as early as the demand process – is critical in order to either plan the structure of the application differently or incorporate the costs into the ROI estimate in as realistic a manner as possible.

All of this knowledge can be used as a foundation to create a framework on whose basis the ROI of an application can be analyzed:

Business case

In the first step, an analysis is conducted within the framework of the business case to determine whether AI should be deployed and whether a product or process can be enriched by AI. Here, steps are taken to determine the level of added value the AI solution will generate for the company and how the solution can be embedded into the overriding business process. An important aspect

[9] PwC. Defining and Measuring Return on Investment for AI. LINK
[10] AiExponent. AI Return on Investment: How to Measure the Business Value of AI. LINK
[11] Slalom. ROI in AI: Measure Value to Deliver Value. LINK
[12] Wallaroo.AI. Why 90% of AI Projects Fail to Hit ROI Targets (And What to Do About It). LINK

here involves the question as to whether the use of AI will be worthwhile and whether the return on investment (ROI) can be maximized through the use of AI.

Central questions:

- Which specific problems or challenges is the AI solution supposed to solve?
- Which quantifiable benefits does the use of AI offer as compared to existing solutions?
- How can the success of an AI solution be measured in the context of the business case?

Compliance

Compliance is an essential component of the requirements profile. This is the point where an analysis is conducted to determine which regulatory provisions (territorial or sectoral) need to be taken into account when developing and implementing an AI solution. The goal here is to ensure that the solution is in conformance with all relevant laws and regulations. Liability risks that might arise from a system failure or improper system behavior also need to be analyzed here.

Before an AI solution that directly affects employees is deployed, the company works council, as the body that represents employee interests, must be incorporated into the process. This is required in order to ensure that the solution implementation is in conformance with labor laws and employees' rights. Getting the works council involved at an early stage not only promotes trust among employees, which is important in these times marked by a shortage of skilled professionals; it also ensures that the AI application will not adversely impact work conditions or the work environment.

The type of legal requirements that an AI solution needs to meet depends heavily on the business area and location in question, as well as the options selected for making the solution available. Different territorial and sector-specific stipulations need to be analyzed in order to make sure that the AI solution will be implemented in a legally compliant manner. This aspect needs to be extensively considered before the application is developed in order to avoid what in some cases might be high liability payments in the event of problems associated with non-compliance.

As was described above and in the section on the AI Act, a risk classification on the basis of the AI Act should be performed as part of the demand process, as this will identify the requirements that the application will need to meet.

Central questions:

- Which specific regulations and laws need to be complied with in connection with the implementation of the AI solution (e.g. GDPR, AI Act)?
- If the solution is to be deployed globally, are international regulations also relevant?
- Which continuous costs for the solution must be taken into account with regard to compliance?

Quality management

Quality management ensures that the AI solution is aligned with the company's business requirements and that the expected quality will be maintained throughout the entire lifecycle. Data quality and forecasting quality are key factors here. A high level of data quality establishes the foundation that enables the AI model to produce precise and reliable results. In order to guarantee consistent quality standards, comprehensive standardization needs to be implemented so as to enable uniform assessment and monitoring of the solution across different systems and processes:

Some sectors already have specific standardization requirements that can be applied to AI systems. It is important for companies to incorporate these requirements at an early stage. This not only establishes trust in the technology and is also expected by partners; it also improves the scalability and adaptability of the solution.

A further focus of quality management is the implementation of testing, training, and operating procedures that ensure the AI solution will remain stable and reliable during ongoing operations as well. Periodic tests and training minimize the risk of bias or model drift. Automated systems for monitoring data quality help guarantee correct and consistent forecasts.

For companies that plan to deploy AI solutions in sensitive areas, it is particularly important that the quality management system can prevent potential improper system behavior or damage to a company's reputation. This not only increases efficiency; it also protects against legal and ethical risks.

Central questions:

- Can we ensure the consistency and reliability of data quality for the specific application? Is the data complete and correct?
- How often and with which processes will the performance of the AI solution be analyzed and validated? Are these processes already established?
- Which quality standards are to be applied for the implementation of the AI solution (e.g. ISO standards, internal guidelines)?

Business requirements

The business requirements refer to the added value the AI solution will deliver for the company or a product. Here it is determined which types of expertise exist at a company or need to be developed in order to be able to effectively implement the AI solution. The risks associated with the introduction of the AI application also need to be assessed here, especially with regard to potential operational challenges. These risks should always be viewed in conjunction with the compliance risks.

The value of an AI solution is measured on the basis of its contribution to the company's business results. This means that the solution must provide clear and measurable added value that is in line with the overall strategy for the company and the existing product portfolio.

Along with the benefits, the potential risks associated with the introduction of an AI solution must be taken into account. These risks relate to business, technical, regulatory, and ethical aspects, which should result from the other dimensions and be derived from them. They are then calculated in this dimension to assign them a monetary value.

Central questions:

- Which operational changes are necessary in order to be able to efficiently utilize an AI solution at a company?
- Is it necessary to use AI to enrich the business process or product in question in order to meet the business requirements?
- How does the introduction of an AI solution influence existing business processes or structures? How "ready" is the company for the introduction?

Requirements for the solution

The requirements for the solution relate to addressing the question as to how the technical solution can be designed in a way that optimally aligns it with the existing requirements.

When AI solutions are introduced, the potential of the technology must be weighed against the potential risks. AI solutions can be integrated into companies in different forms, as they are used in products, specific features, or as SaaS (software as a service) solutions. Another key point for decision making is the question as to whether an AI solution should be purchased, an existing solution should be expanded, or a new solution should be developed within the company. This decision depends on existing resources, the urgency in terms of time, and the company's long-term goals. The successful introduction and operation of an AI solution necessitates the availability of specific types of expertise at a company, including everything from technical capabilities for implementing the solution to knowledge regarding how the required infrastructure can be provided and maintained. Companies must therefore ensure that internal expertise is either available or can be developed within the organization.

Central questions:

- Which form of AI solutions (product, feature, or SaaS) best matches the company's current business strategy and the use case in question?
- Which resources (employees, budgetary, time) are needed to implement and operate the AI solution?
- Should the AI solution be purchased, an existing solution be expanded, or a completely new solution be developed within the company in order to solve the current problem?

- How scalable is the solution?

Options for making the solution available

The analysis and consideration of the business demands and the requirements of the business case can be used to define the options for making the AI solution available.

Such options for an AI solution are based on the solution's purpose. Is it a solution that is meant to make forecasts (for machine maintenance intervals, for example), or is it a solution that should generate artifacts such as images or texts (e.g. video subtitles)? Another important factor when selecting the options for making the solution available involves weighing the costs and benefits. The long-term usefulness of an AI solution is also influenced by the options for making it available. The development of internal generative models involves a long-term decision that is associated with substantial investment in further development and utilization. Here, it must be decided whether the flexibility and control of an internal development is preferred, or if a quickly available standard solution that requires less adaptation and adjustment should be chosen.

Central questions:

- Which models are best suited to making the AI application available? Beginning in August 2025: Is the model approved in the EU (CE marking) and can it be used for the specific application in question?
- How flexible is the solution in terms of future adjustments or expansions? How flexible are our processes in terms of adapting and expanding the models?
- What types of long-term maintenance and support requirements arise in connection with the selected option for making the application available?

8.3 The design phase

The design phase involves developing concepts for generative or predictive AI solutions for products or business processes in cooperation with the relevant/affected specialist departments. This process, which forms the foundation for the successful deployment of AI at a company, concentrates on the assessment of feasibility and the added value contribution of the solution.

Focus

The focus in this phase is the risk and opportunity analysis for the deployment of AI solutions in the specific business case. Here, it is determined whether an AI solution is technically feasible and economically viable. Aspects such as feasibility, cost-effectiveness, and the actual effectiveness of the solution are examined in detail. Attention should be paid here to the fact that not every requirement necessitates an AI solution. The results of the design process therefore do not always

have to lead to an actual implementation – in some cases it may make more sense to forgo an AI solution. This would especially be the case with regard to applications in highly regulated areas, or if the data situation is not secure or sufficient.

In many cases, the goal of the design process is to create a high-level solution design that can serve as the basis for further steps.

Internal and external requirements

It is essential in this phase to realistically evaluate the technical and economic conditions for a given AI solution. Instead of succumbing to all the AI hype, a sober analysis of the risks and opportunities must be carried out. Along with technical feasibility, the cost-benefit aspects also need to be considered. A comprehensive risk assessment here ensures that potential problems can be identified and managed at an early stage. Legal and operational framework conditions and cost factors must also be considered in order to ensure that the solutions not only make sense from an internal point of view but also meet external requirements.

Affected processes

- Demand process
- Security process
- Software development lifecycle process
- Enterprise architecture management
- Solution scouting

Artifacts

- Initial risk assessment
- Guidelines and standards document
- Change management plan
- High-level solution design

8.4 The development phase

During the development phase, the AI solution is planned and prepared in detail and then tested while taking into account all legal, ethical, and economic requirements. This step in the lifecycle of an AI solution ensures that the solution will meet all business requirements and will remain viable after being put into operation.

Focus

This phase involves determining whether the AI solution can meet the requirements defined in the business case. Development focuses on various aspects, including time to market, time to value, and the performance of the solution. It also needs to be considered here how well the solution recognizes anomalies and deviations, and whether it is scalable and explainable. If necessary, adjustments to the solution design can be made here in order to achieve the desired performance.

The solution is extensively tested to ensure that it corresponds to the defined criteria. A particular focus here involves conformity with specific documentation specifications to ensure that all regulatory and technical requirements are transparent and clearly understood, and will thus be complied with.

Internal and external requirements

- **Documentation:** All processes and technical steps must be perfectly documented in order to ensure transparency. The AI Act stipulates that high-risk applications in particular must meet strict documentation requirements, including a retention period of 10 years (Article 18).
- **Business case:** The solutions need to be aligned with the defined requirements of a business case that has been previously defined, and must also support the achievement of the objectives or targets of the business case.
- **Policies and guidelines:** The solutions should be implemented in line with internal and external guidelines in order to ensure compliance with legal and ethical standards. The special challenge here involves translating guidelines into practically focused requirements for specific actions.

Affected processes

The affected processes are to be listed in an exemplary manner and should be designated using common terms as is customary at many companies. With regard to the individual instances of implementation, the processes need to be adjusted in line with the overall strategy for the company.

- MLOps process

- DevOps process
- Testing process
- Security process
- Information security process

Article 15 of the AI Act stipulates that high-risk AI systems should be designed and developed in such a way that they achieve an adequate level of accuracy, robustness, and cybersecurity. Such systems must also perform consistently in these respects throughout their lifecycle. Technical and organizational measures must be taken to guarantee security.

Attention must be paid here to ensuring that the GDPR is taken into account throughout the entire lifecycle.

Artifacts

- Internal compliance report
- External reporting document
- Solution documentation
- Change management document
- Documentation guidelines
- Conformity assessment report

High-risk systems must undergo a dedicated conformity assessment process before they can be launched on the market (Article 43)

8.5 Implementation and use

Focus

This phase focuses on the introduction of an AI solution to the market or at a company. This process is to be accompanied by documentation procedures and training as a means of ensuring that everyone involved knows how the solution can be deployed and monitored. Users often have reservations regarding these new technologies, particularly in the case of semi-autonomous or fully autonomous systems that can make decisions on their own or perform tasks without human intervention. It is therefore important to precisely define and explain how the AI solution is to

function and which limits will be set on its operation. The technical further development of the solution is also essential for ensuring that it remains state of the art in terms of technological standards.

Internal and external requirements

- **Solution documentation:** Steps must be taken to ensure that solutions are documented in a way that enables them to be operated, serviced and, if necessary, further developed – and also employed internally for other use cases in a modular manner. Documentation for high-risk systems is subject to stringent and standardized requirements.
- **Accessibility and clarity:** Documentation must be accessible to all parties involved and must also be clear so that they can understand it. This especially applies in the case of people who do not possess in-depth technical knowledge. The documentation makes clear which stakeholders will be using the application in the future and how they are to be trained in such use. Pursuant to the transparency obligations (Article 50), it must be ensured that the affected persons have been provided with sufficient information before they interact with an AI system.
- **Monitoring:** The approach used to monitor an AI solution must be in line with legal requirements and ensure that potential risks are identified and managed at an early stage. Article 72 defines the following requirements for high-risk systems: The creation of a strategic monitoring plan, the establishment of a monitoring system, and active and systematic collection of data that shows how the system operates and acts.

Affected processes

- Information security management
- Knowledge management processes
- Security incident and IT service continuity management
- Enterprise architecture management

Incident management: Definition of processes for dealing with disruptions and errors while the AI solution is operating. Here as well, high-risk systems are subject to specific legal requirements pursuant to Article 73.

Risk management: Monitoring of potential risks in connection with the introduction and use of an AI solution. Particular attention is paid here to the transfer of these risk monitoring operations to the company's overall risk management process. The liability risk associated with AI applications after the implementation of the two most important regulatory frameworks – the AI Act and the GDPR –

amounts to 7% and 4%, respectively, of global revenue, which means it is essential that the risks of AI applications be viewed at the overall company level.

Appendix

A. Method, procedure, and result

The purpose of CBA Lab's "AI Governance" workstream was to develop a comprehensive AI governance model that helps companies minimize the risks associated with AI deployments while also making it possible to exploit the full potential of the technology. In addition, this development process was to cover all the technical, ethical, regulatory, economic, and organizational aspects of the entire AI lifecycle.

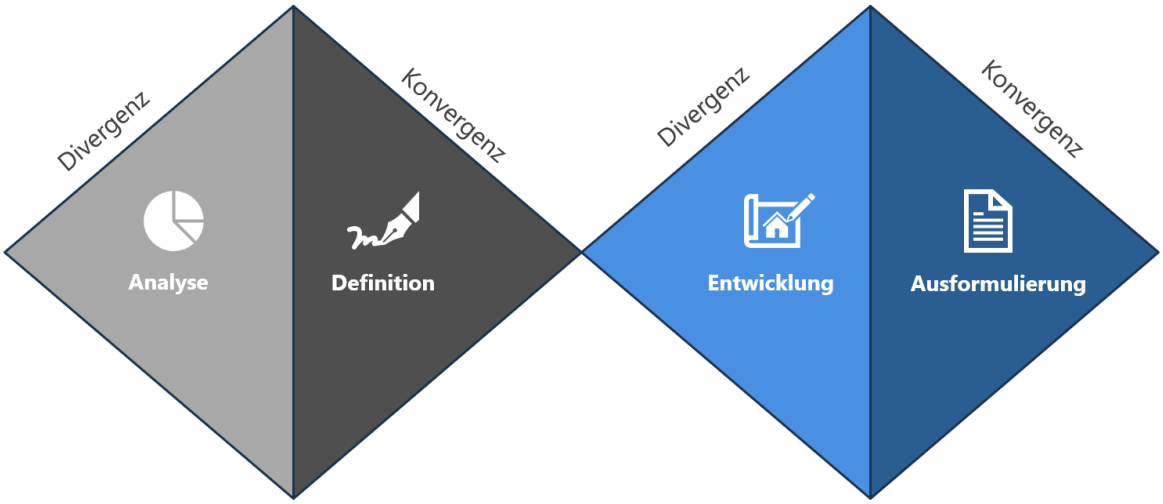

Figure7: Design thinking – Double Diamond

The development process used the Double Diamond process from the design thinking method as a guide, whereby the Double Diamond process is divided into four phases: Discover, Define, Develop, and Deliver.

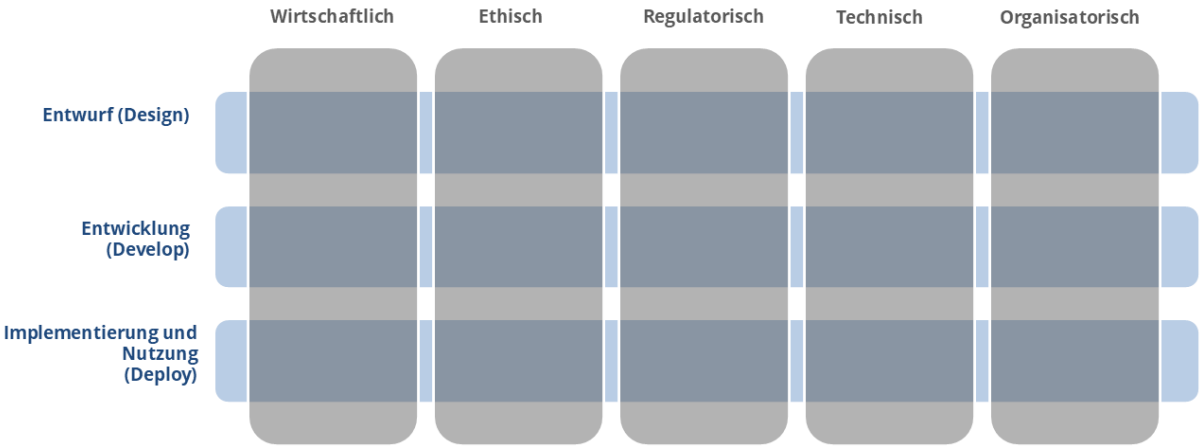

Figure8: Matrix teams

During the Discover phase, the progress being made at the moment in terms of thematic areas at companies was analyzed, problem areas and requirements were identified, and participants were divided into horizontal and vertical teams. The horizontal teams worked on the various phases of the AI lifecycle (design, development, implementation, and use), while the vertical teams were organized along the lines of the five dimensions of economic, ethical, regulatory, technical, organizational. The horizontal teams worked in a cross-functional manner in order to avoid a silo mentality, while the vertical teams promoted knowledge sharing within their areas of expertise.

The Define phase made it possible to prioritize important topics and problem areas and assign these to topic owners, who were then responsible for ensuring they were continuously taken into account. During the Develop phase, the teams created a process-focused AI governance model that covered all of the identified challenges and requirements. After that, in the Deliver phase, the results were brought together, visually depicted, and integrated into a white paper.

The workstream consisted of nine workshop units, beginning with a three-hour kickoff meeting that was followed by eight weekly check-ins that ran for two hours each. The kickoff meeting featured presentations, status reports from the participating companies, and a brainstorming session that led to a consolidation of the topics and the creation of the horizontal and vertical teams. During the weekly check-ins, participants discussed the progress that had been made, clarified definitions, and presented model proposals.

The result of the workstream is the White Paper, which presents the comprehensive and practically focused AI governance model that was developed. The contents of the White Paper consist of the results of all the teams, which were iterated and challenged several times throughout the course of the process. The model offers companies a comprehensive approach for implementing an AI

governance system that takes into account the economic, ethical, regulatory, technical, and organizational aspects throughout the entire AI lifecycle.

B. Glossary

AI Office

An AI Office at a company is a specialized unit that manages, controls, and supports the deployment of artificial intelligence (AI) systems at the company. It coordinates initiatives for implementing AI systems, ensures that the necessary training and awareness-raising measures for employees are taken, and assists with the safe, secure, and effective integration of AI technologies. The AI Office also serves as a central point of contact for knowledge sharing and the identification of best practices. It promotes the safe, secure, and ethical use of AI in order to increase acceptance of AI systems and make them more understandable to everyone throughout the entire company. The AI Office may be structured in different ways depending on the size of the organization in question. For example, it may be designed more like a community, with one or several part-time employees, or it can be set up as a larger center of expertise.

AI literacy

(From the AI Act): “AI literacy means skills, knowledge, and understanding that allow providers, deployers, and affected persons, taking into account their respective rights and obligations in the context of this Regulation, to make an informed deployment of AI systems, as well as to gain awareness about the opportunities and risks of AI and possible harm it can cause.”

C. Charts, images, and illustrations

Legal Notice

Publisher
Cross-Business-Architecture Lab e. V.
Hinter Hoben 149
53129 Bonn
Tel.: +49 228 55 51 131
e-mail: info@cba-lab.de
www.cba-lab.de
https://twitter.com/cba_lab
https://de.linkedin.com/company/cba-lab

Board members authorized to represent CBA Lab
Joachim Schmider, Chairman
Dr. Arun Anandasivam
Dr. Johannes Helbig
Dr. Karsten Schweichhart (responsible within the meaning of German press and media law)