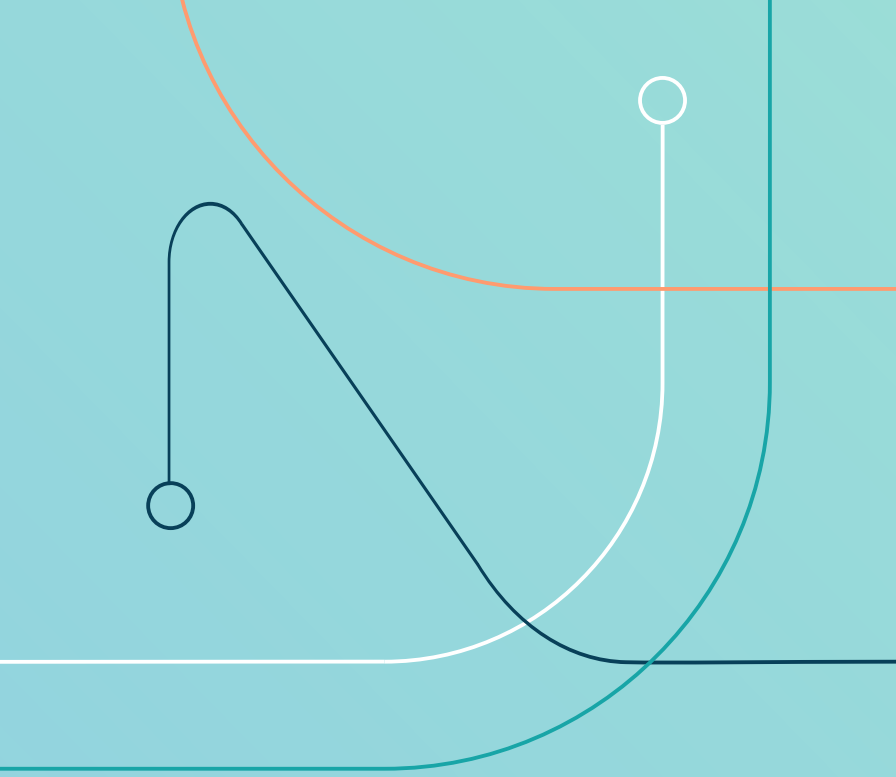




DEEP DIVE

Content break-down for High Risk Requirements on the Use Case level

An abstract graphic in the top left corner consisting of several overlapping lines in teal, orange, and white, with small circles at their ends, set against a light teal background.

The EU Act is going to be regulating the development, offering, and usage of AI Systems and General Purpose AI Systems in the EU, meaning operators of all kind, including provider and deployer, need to demonstrate compliance with applicable requirements before placing such systems on the market or putting them into service.

The AI Act adopts a risk-based approach, that is, the higher the risk of a given AI System (aka. Use Case) the more stringent are the requirements. Amongst the allowed AI Systems, so called High-Risk AI Systems face the most comprehensive set of requirements, which are predominantly on the shoulders of the provider, because they specify the intended purpose and make crucial design choices.

While such requirements ought to prevent unreasonable damage to health, safety and human rights, they also demand attention and expertise, as they are challenging from a technical and legal point of view. And since the AI Act is brand new, there is close to zero experience in companies and authorities on how to operationalize such requirements, i.e. the learning curve is not only steep, but also uncertain and potentially iterative.

appliedAI Institute for Europe intends to support individuals across the AI Ecosystem to reach compliance with minimal invest in time and resources, so that attention and expertise can be “spend” on driving trustworthy AI Innovation. This new content resource is a step into this direction.

Who is it for? Providers of High Risk AI Systems

Providers of High Risk AI Systems are required to demonstrate compliance with certain requirements. Some of those apply on the Use Case level, while others are on the Organizational level. Here, we focus on the Use Case level.

What is the problem? The technical implementation is difficult and uncertain

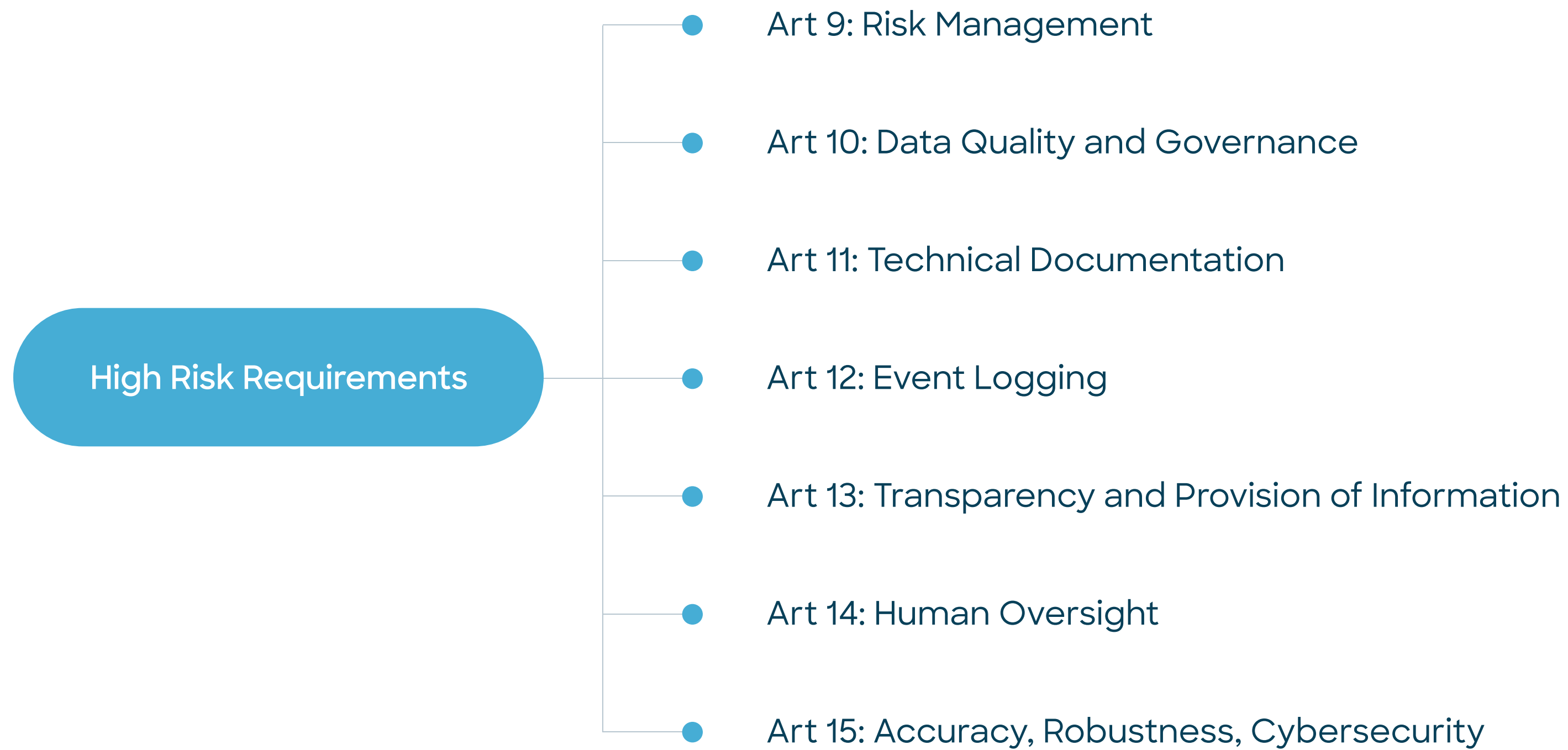
Operationalising the requirements on the Use Case level (e.g. Data Governance, Transparency, Accuracy) is challenging from a technical perspective as they refer to fundamental design choices made during the ML Lifecycle. Plus, the requirements need to be interpreted in the context of a given Use Case to ensure an adequate level of safety.

What is our contribution? content break-down for High Risk Requirements on the Use Case level

The first step is to foster a solid understanding of the requirements as such, link them to design choices, and to consider how they might apply to a Use Case at hand. With this in mind, we focus on the Articles 10-15 and provide for each requirement:

- A breakdown of each requirement in the form of a mind-map for easy reading
- An indication how the requirement might link to design choices
- Prompts for implementation for a given Use Case

Title III Chapter 2



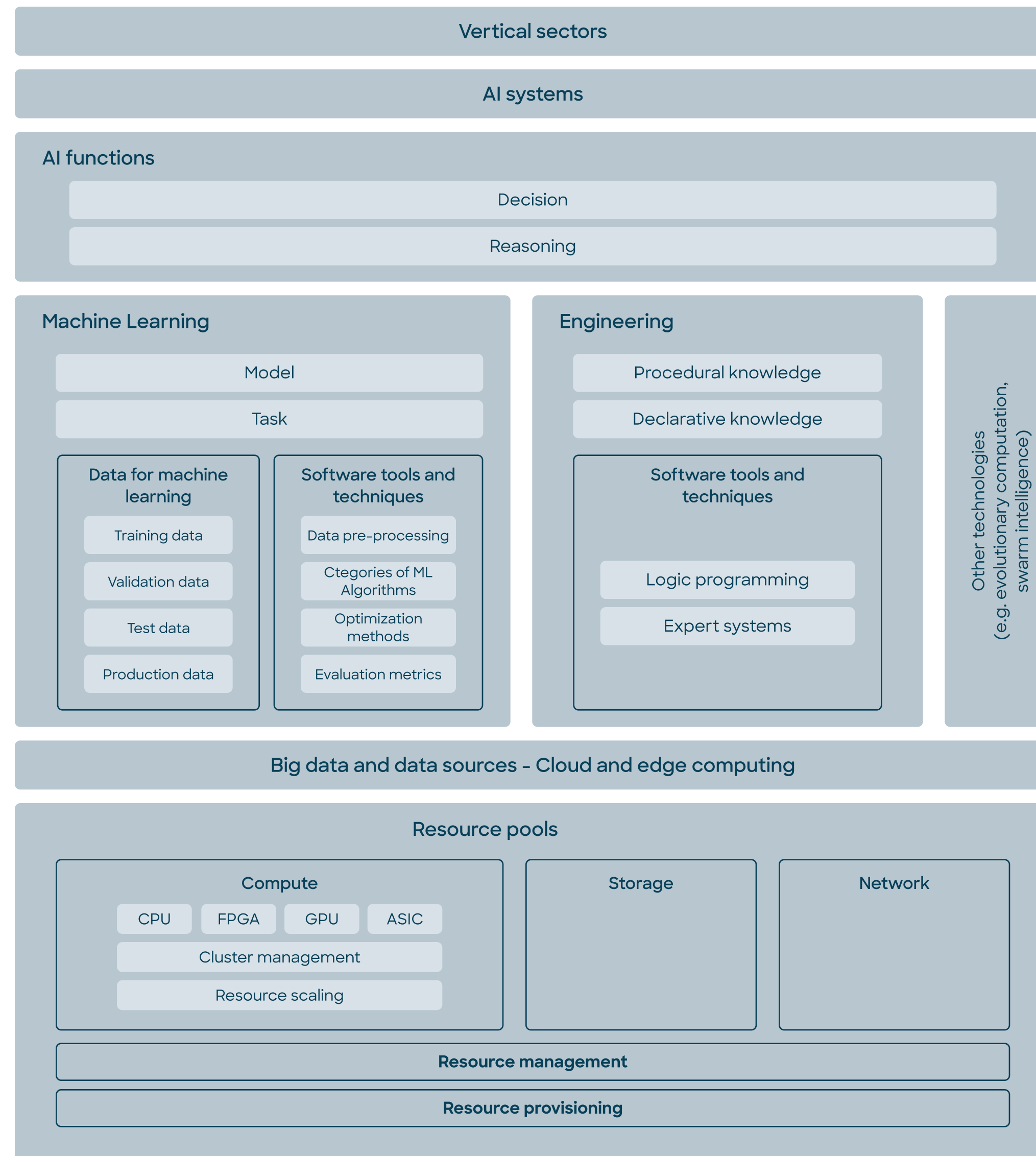
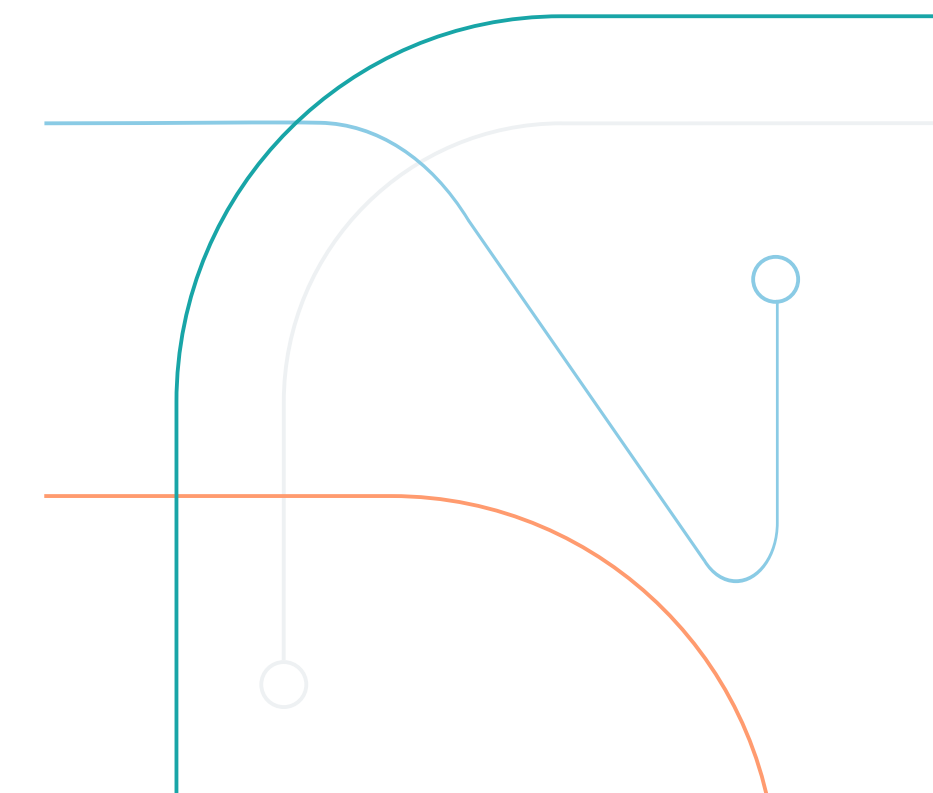


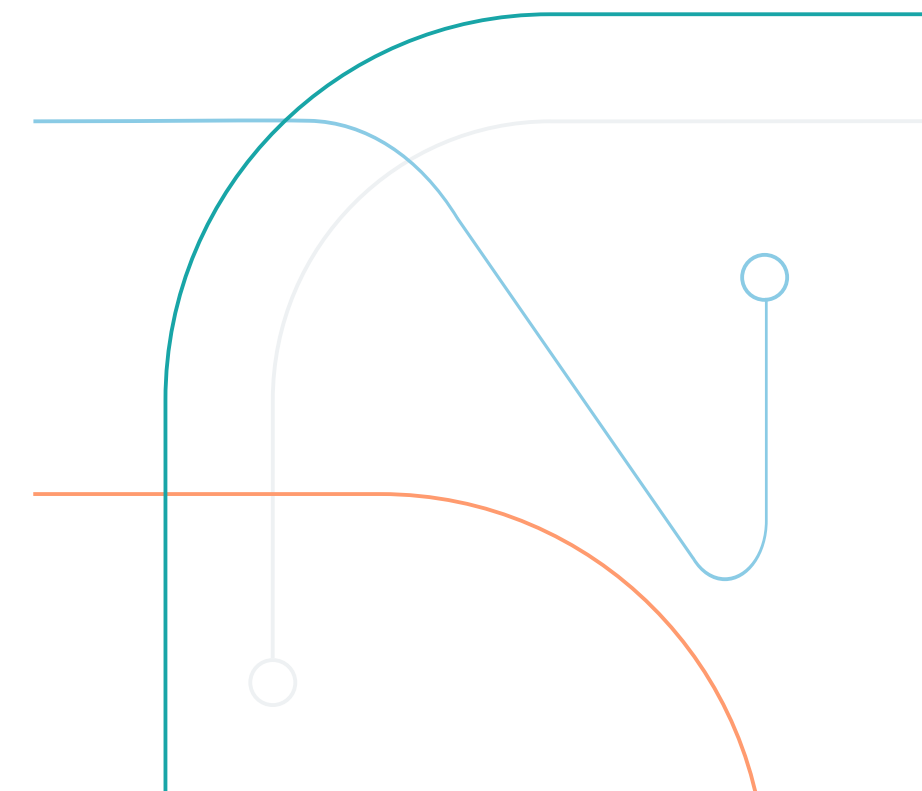
Figure 6 - AI ecosystem

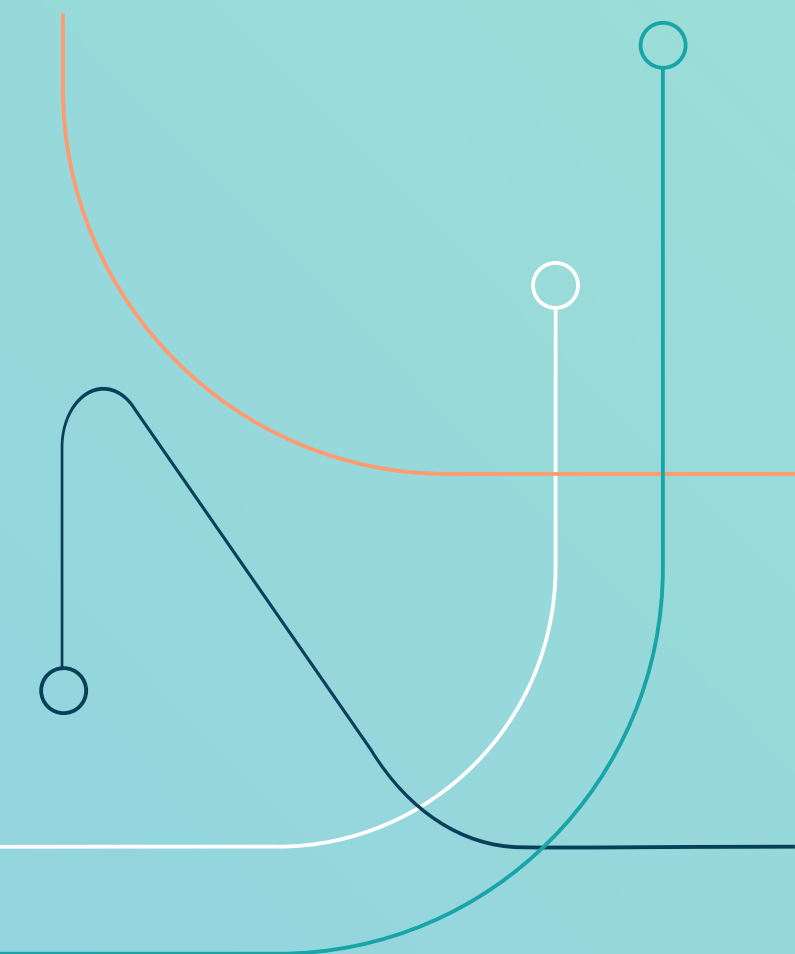
- **Vertical Sectors:** For example, healthcare, finance, manufacturing etc.
- **AI system:** Engineered system that operates with varying levels of autonomy and generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives
- **AI Functions:** The output of a system - recommendations, predictions, etc.
- **Machine Learning:** Computational techniques to learn from data
 - **Data for machine learning:** The various data sets involved in ML
 - **Software tools and techniques:** The pipelines, analytics and other software engineering techniques applied to data-sets
- **Engineering:** Software approaches that support or complement an AI use case
- **Cloud and edge computing:** Centralised or distributed computing
- **Resource Pools:** Compute, networking, storage etc.



Title III Chapter 2

- **Intended use:** Classifying individuals into loan eligibility buckets based on prior financial activity
- **Application domain:** credit institutions
- **Type of AI system:** classification model
- **Data source:** Merged data from third-party vendor and in-house data from financial transactions
- **Intended deployer:** European banks
- **Human operators:** Bank IT staff & ordinary bank employees
- **Resource pool:** Third party cloud compute service provider



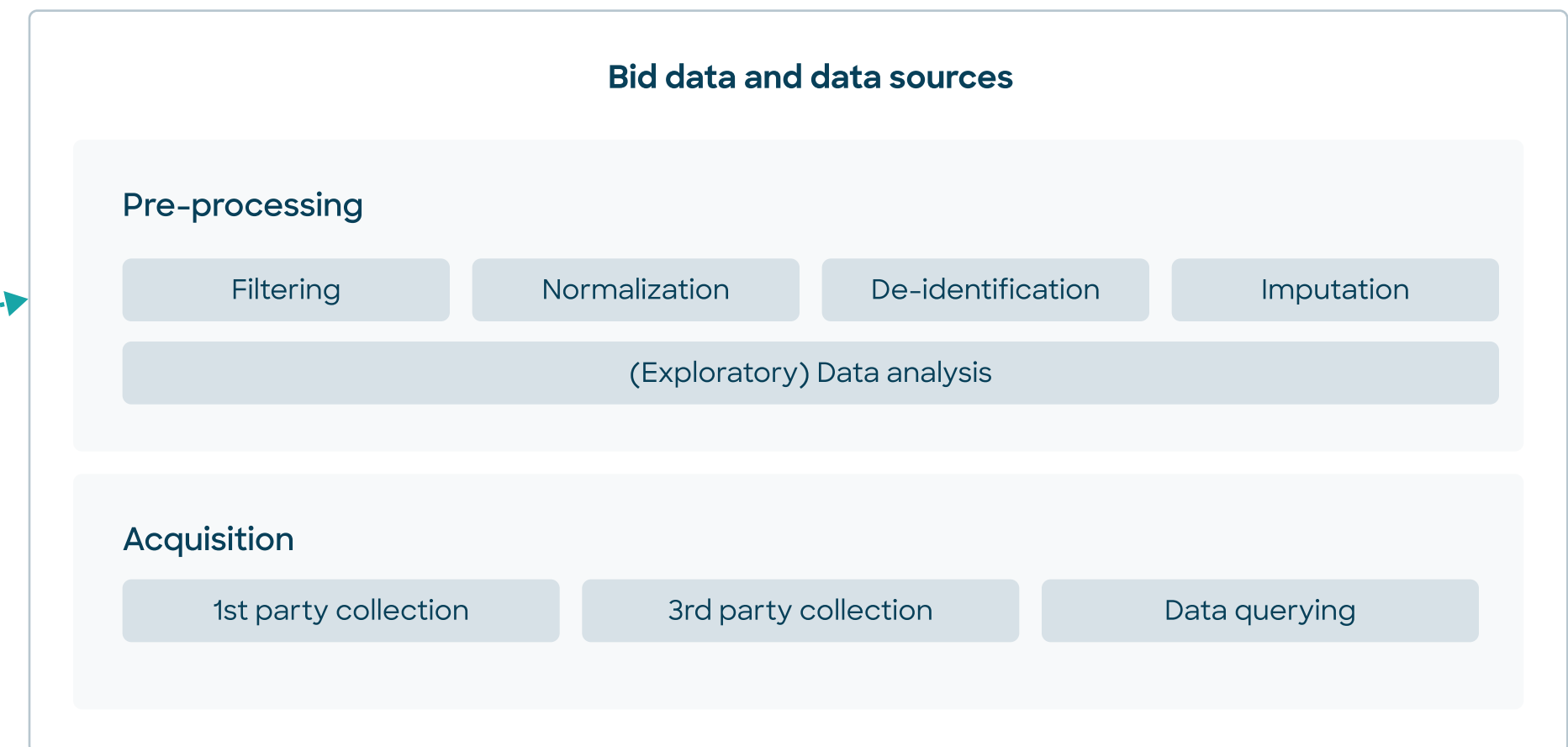
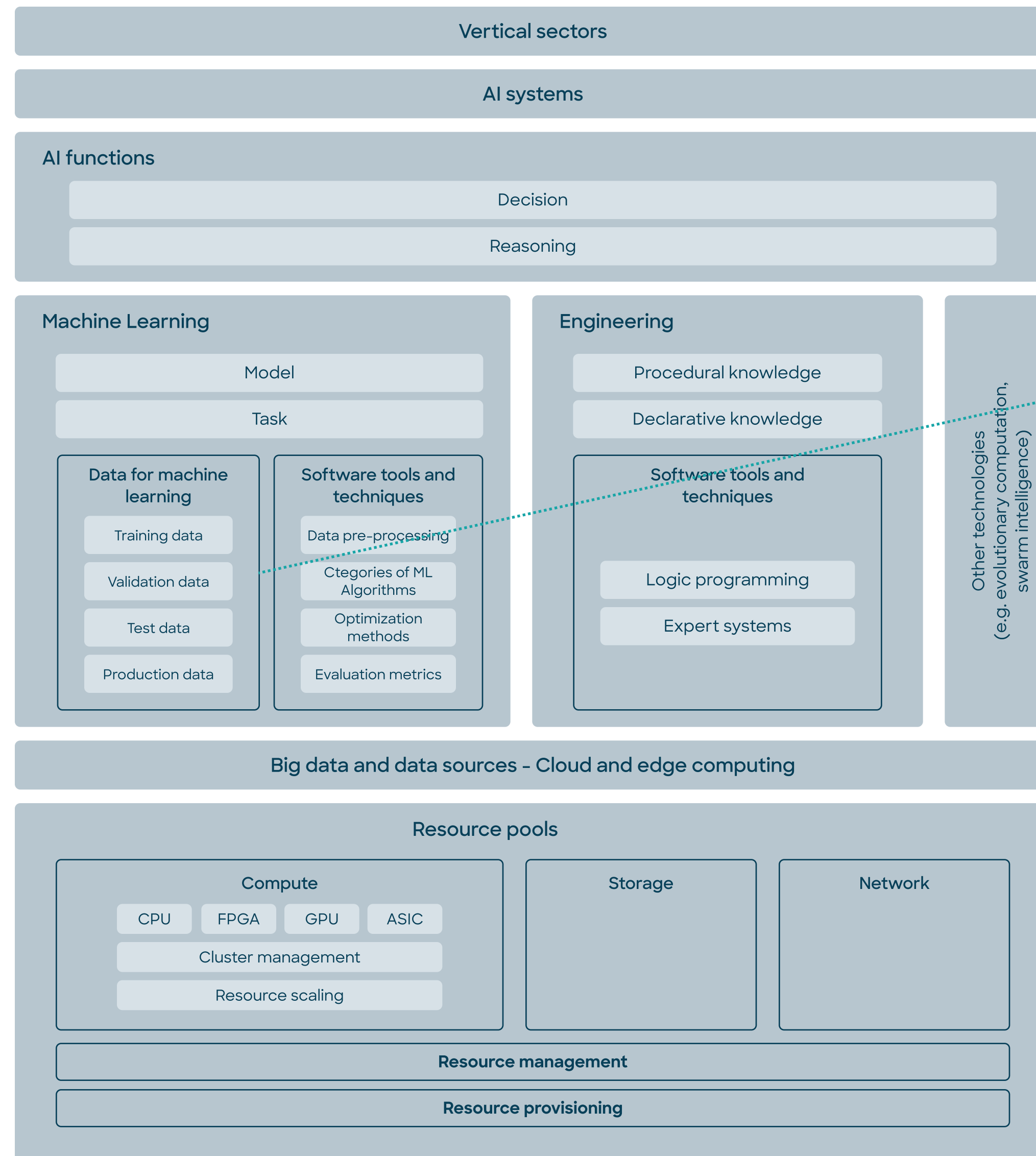


ARTICLE 10

Data Governance

Article 10

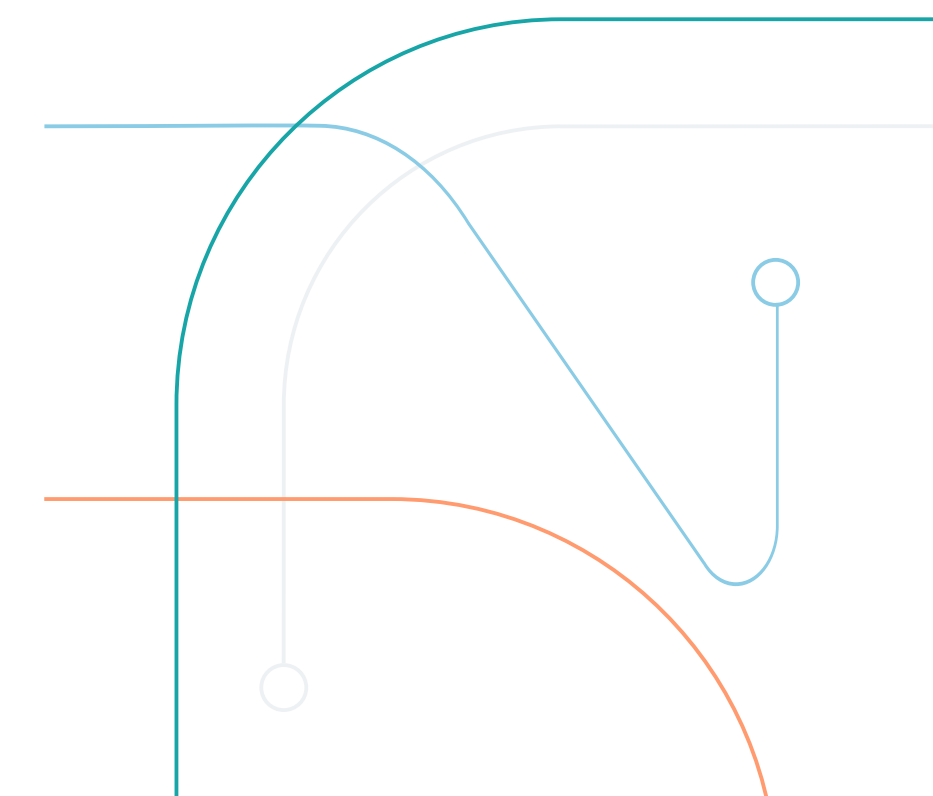


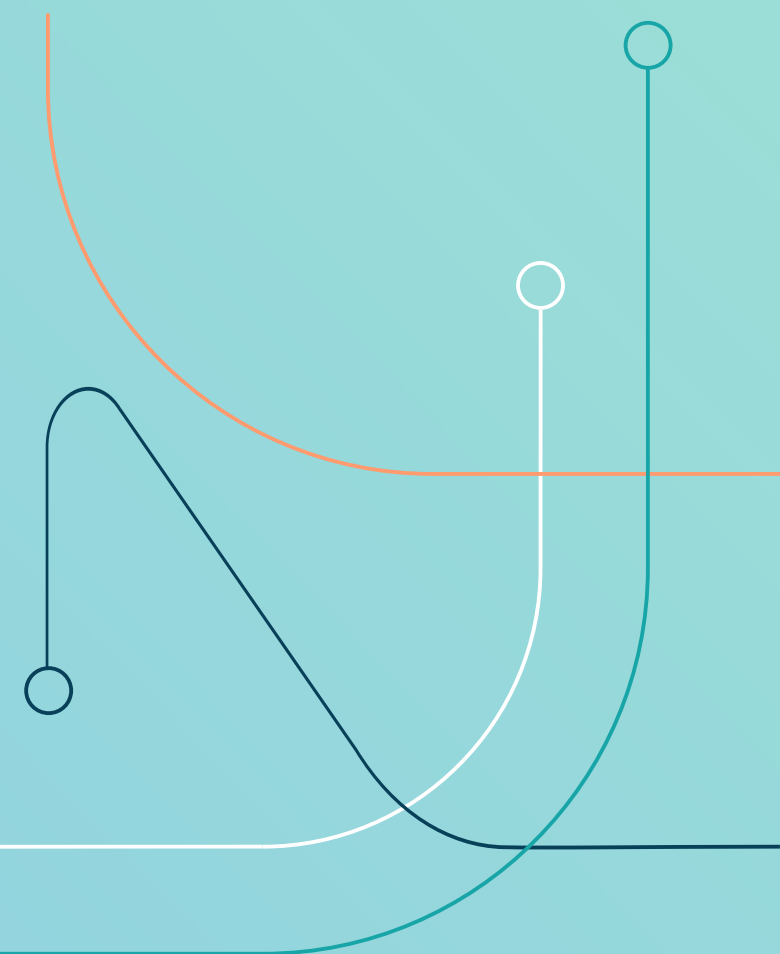


- **Where** did my data come from?
- **What** are the properties and quality of my data?
- **How** was the data processed or modified?
- **Why** was the data processed or modified?

Figure 6 - AI ecosystem

List at-least two data governance and data quality concerns with your use case





ARTICLE 12

Event logging & Record-keeping

Article 12

Event Logging

In ISO 27001, a log file is defined as a "record of events having significance for the management of the information system"

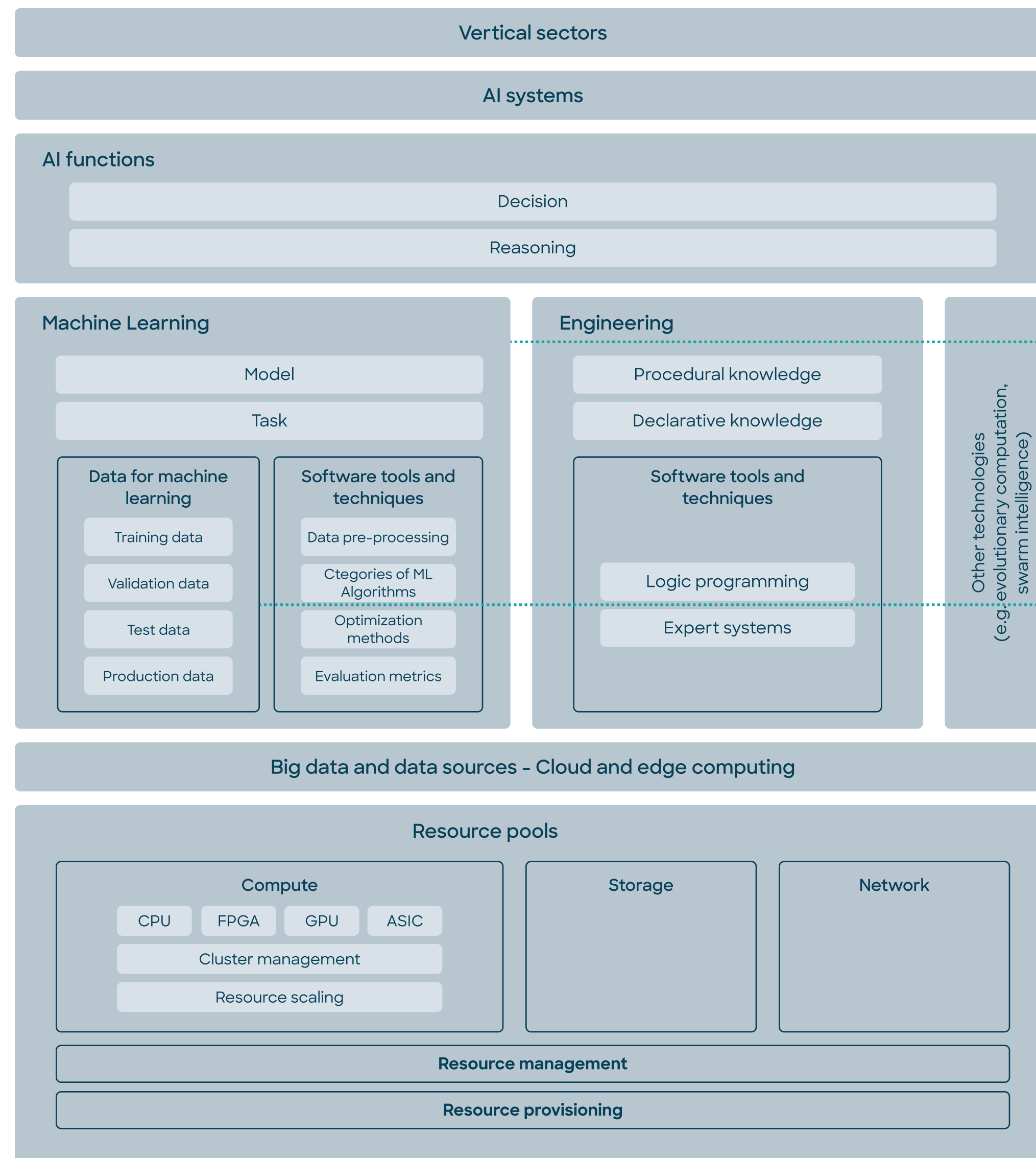
- Events that might lead the AI system to presenting a risk within the meaning of Article 65 (1)
- Events that might lead to a substantial modification of the AI system.
- Facilitation of the post-market monitoring referred to in Article 61
- Special requirements for biometric based systems

Article 3(19) Regulation on Market Surveillance and Compliance of Products: 'product presenting a risk' means a product having the potential to affect adversely health and safety of persons in general, health and safety in the workplace, protection of consumers, the environment, public security and other public interests, protected by the applicable Union harmonisation legislation, to a degree which goes beyond that considered reasonable and acceptable in relation to its intended purpose or under the normal or reasonably foreseeable conditions of use of the product concerned, including the duration of use and, where applicable, its putting into service, installation and maintenance requirements

'substantial modification' means a change to the AI system after its placing on the market or putting into service which is not foreseen or planned in the initial conformity assessment by the provider and as a result of which the compliance of the AI system with the requirements set out in Title III, Chapter 2 of this Regulation is affected or results in a modification to the intended purpose for which the AI system has been assessed

"...actively and systematically collect, document and analyse relevant data which may be provided by deployers or which may be collected through other sources on the performance of high-risk AI systems throughout their lifetime, and allow the provider to evaluate the continuous compliance of AI systems with the requirements set out in Title III, Chapter 2. Where relevant, post-market monitoring shall include an analysis of the interaction with other AI systems. This obligation shall not cover sensitive operational data of deployers which are law enforcement authorities"

- recording of the period of each use of the system (start date and time and end date and time of each use)
- the reference database against which input data has been checked by the system;
- the input data for which the search has led to a match;
- the identification of the natural persons involved in the verification of the results, as referred to in Article 14 (5).



! You're probably monitoring the model as a standard MLOps practice

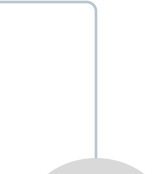
Concept drift: the actual outcome you seek to model, or the relationship between the data and the outcome, may fray.

Domain shift: if your dataset does not appropriately sample the production, post-deployment setting, the model's performance may suffer.



Data Drift: The underlying data can unexpectedly change. This might be because:

- Bugs in the data pipeline
- New data
- Malicious attacks



Service usage metrics, such as the total number of model calls, RPS (requests per second), and error rates.

System performance metrics, such as uptime and latency.

Resource utilization metrics, such as memory and GPU/CPU utilization.

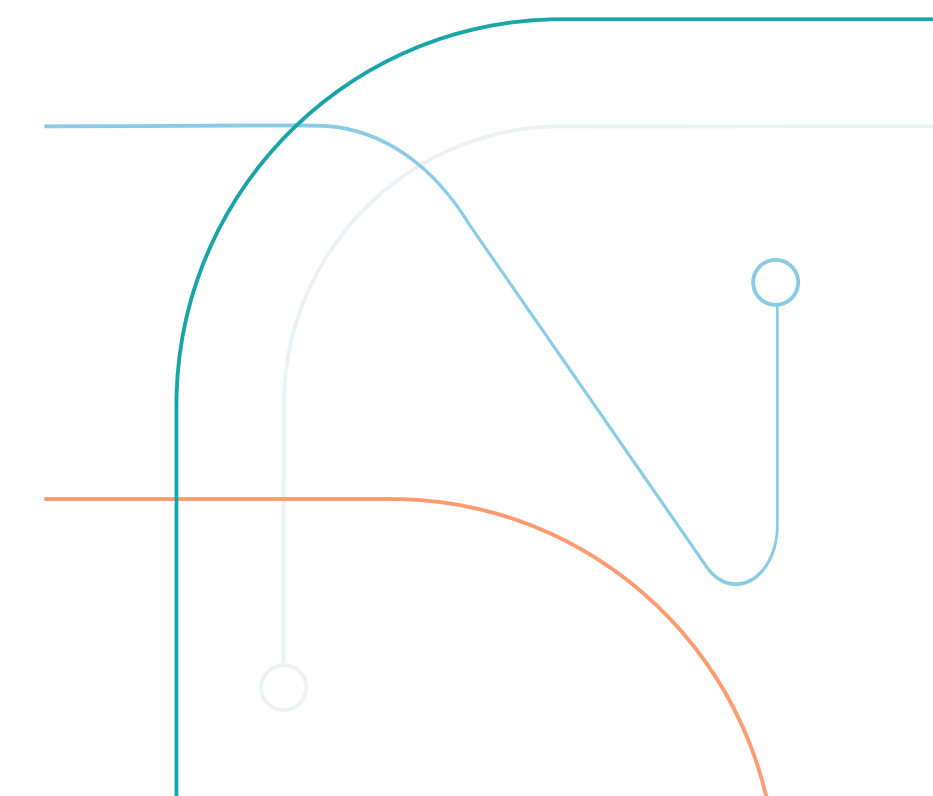


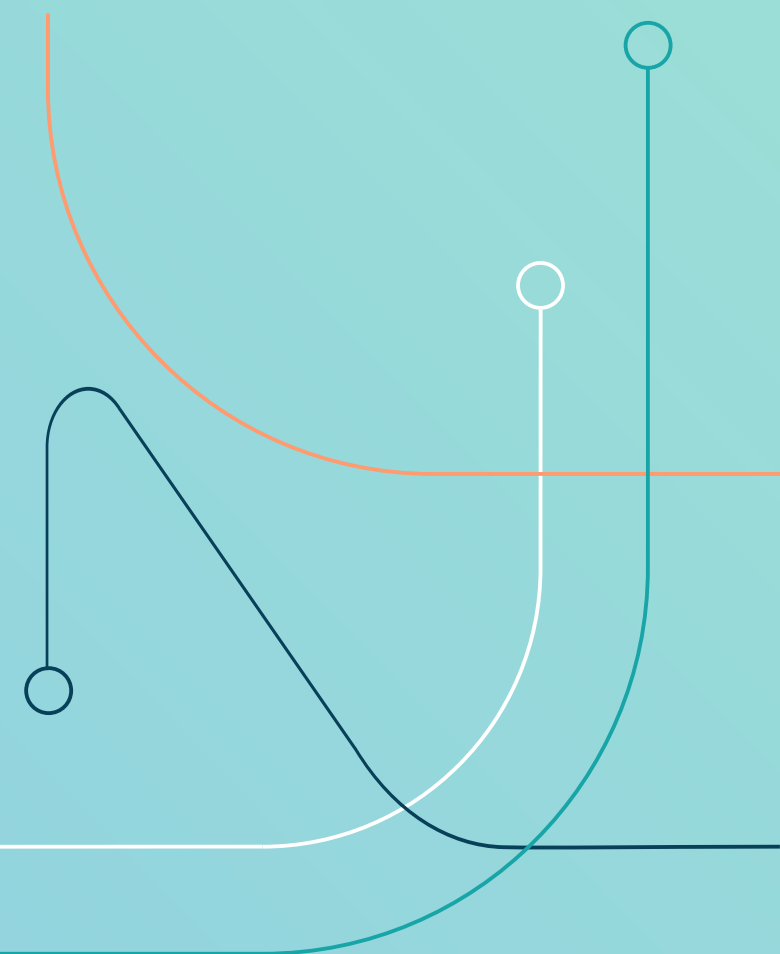
? Logging the appropriate metric or measurement depends on your system design

Figure 6 - AI ecosystem

What events do you already log?

What events would you need to log?





ARTICLE 13

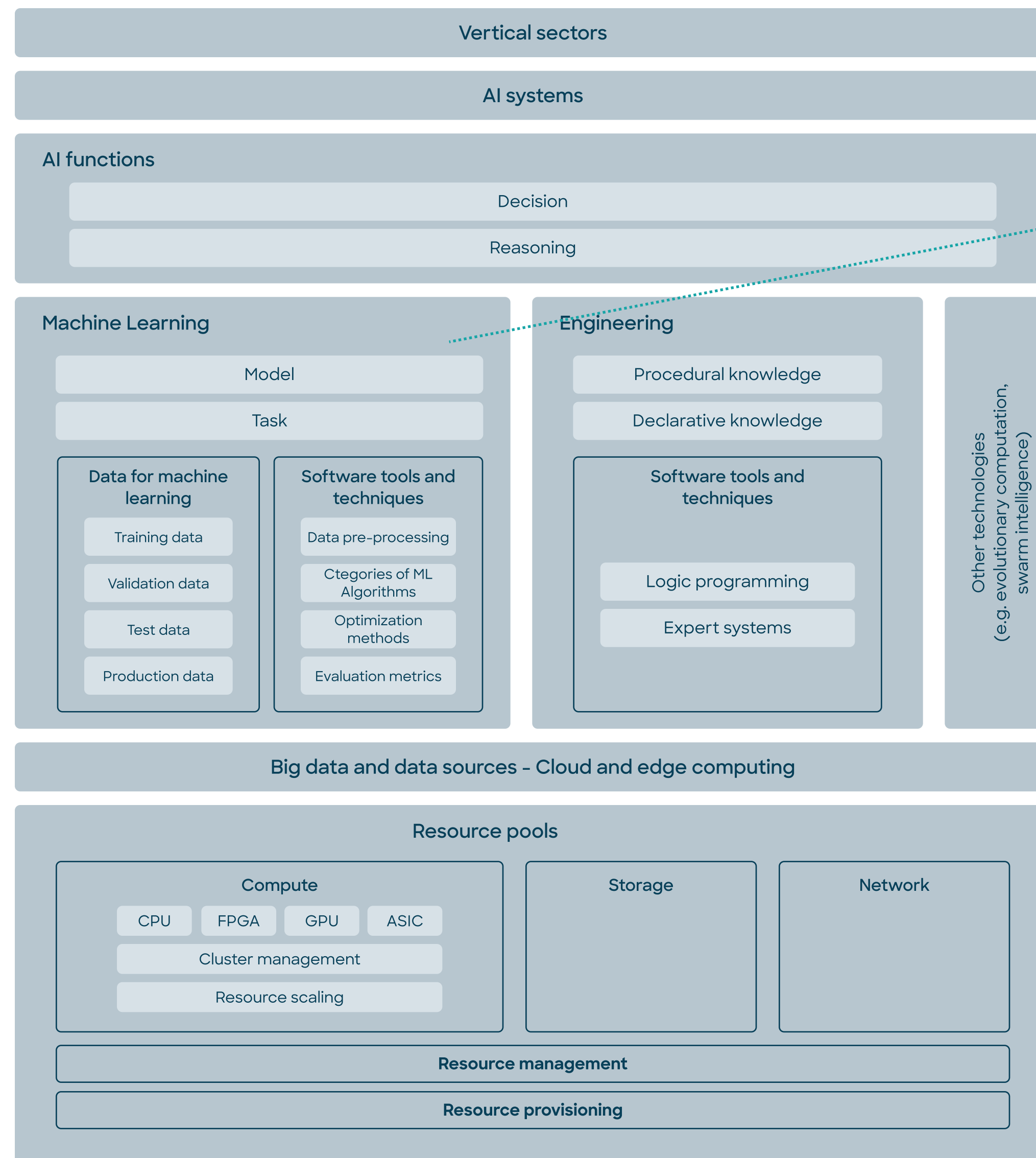
Provision of Information

Article 13

Transparency & Information to Users

Transparency shall thereby mean that, at the time the high-risk AI system is placed on the market, **all technical means** available in accordance with the generally acknowledged state of art **are used to ensure that the AI system's output is interpretable by the provider and the user**. The user shall be enabled to **understand and use the AI system appropriately** by generally knowing how the AI system works and what data it processes, allowing the user to explain the decisions taken by the AI system to the affected person pursuant to Article 68(c).

- designed and developed in such a way to ensure that their operation is sufficiently transparent to enable deployers to interpret the system's output and use it appropriately
- How should information should be provided?
 - intelligible instructions for use in an appropriate digital format
 - correct, clear and to the extent possible complete information that helps operating and maintaining the AI system as well as supporting informed decision-making by users
 - at least in the language of the country where the AI system is used
- the identity and the contact details of the provider and, where applicable, of its authorised representative
- the characteristics, capabilities, and limitations of performance of the high-risk AI system
 - Its intended purpose
 - the level of accuracy, robustness and cybersecurity
 - any known or foreseeable circumstance which may lead to risk
 - where applicable, the technical capabilities and characteristics of the AI system to provide information that is relevant to explain its output
 - when appropriate, its performance regarding specific persons or groups of persons on which the system is intended to be used
 - when appropriate, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used
 - where applicable, information to enable deployers to interpret the system's output and use it appropriately
- the changes to the high-risk AI system and its performance which have been predetermined by the provider at the moment of the initial conformity assessment
- the human oversight measures referred to in Article 14
- the computational and hardware resources needed, the expected lifetime of the high-risk AI system and any necessary maintenance and care measures, including their frequency, to ensure the proper functioning of that AI system, including as regards software updates
- where relevant, a description of the mechanisms included within the AI system that allows users to properly collect, store and interpret the logs



| | | Level of automation | Comments |
|------------------|----------------------------|--|---|
| Automated system | Autonomous | 6 - Autonomy | The system is capable of modifying its intended domain of use or its goals without external intervention, control or oversight. |
| | | Heteronomous | 5 - Full automation |
| | 4 - High automation | The system performs parts of its mission without external intervention | |
| | 3 - Conditional automation | Sustained and specific performance by a system, with an external agent being ready to take over when necessary | |
| | 2 - Partial automation | Some sub-functions of the system are fully automated while the system remains under the control of an external agent | |
| | 1 - Assistance | The system assists an operator | |
| | | 0 - No automation | The operator fully controls the system |

What is the **role** of the deployer & human operator?

What is the **expertise** of the deployer & human operator?



Have you addressed **all possible stakeholders**?



What are the **different formats to present information**?

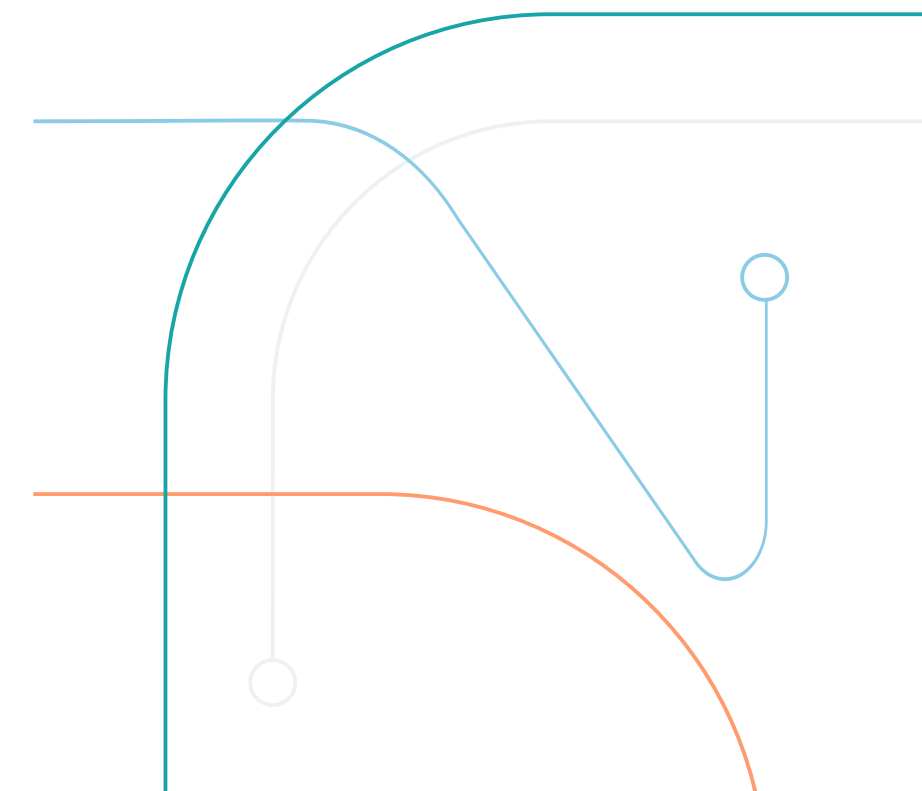
Instructions, guides, warnings, labels, etc.

Figure 6 - AI ecosystem

Which part of the system output requires transparency?

How can this be communicated to the deployer?

Which stakeholders require information or "transparency?"

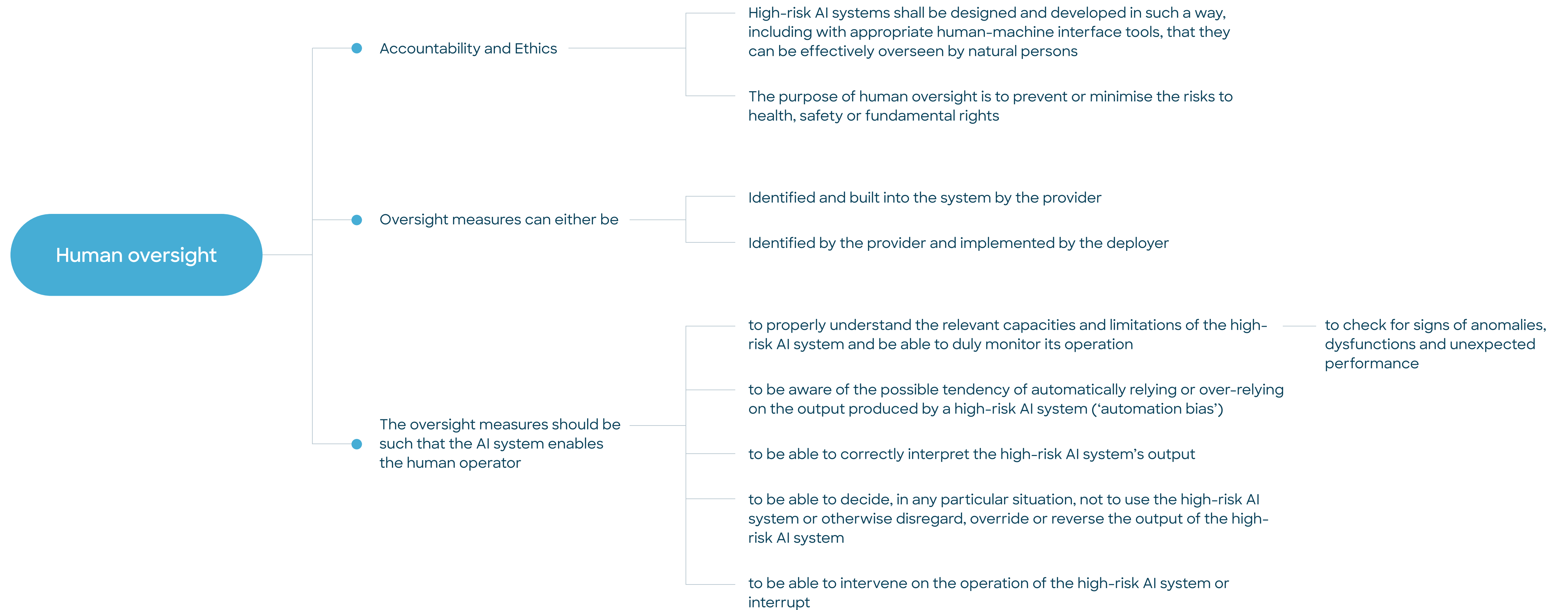


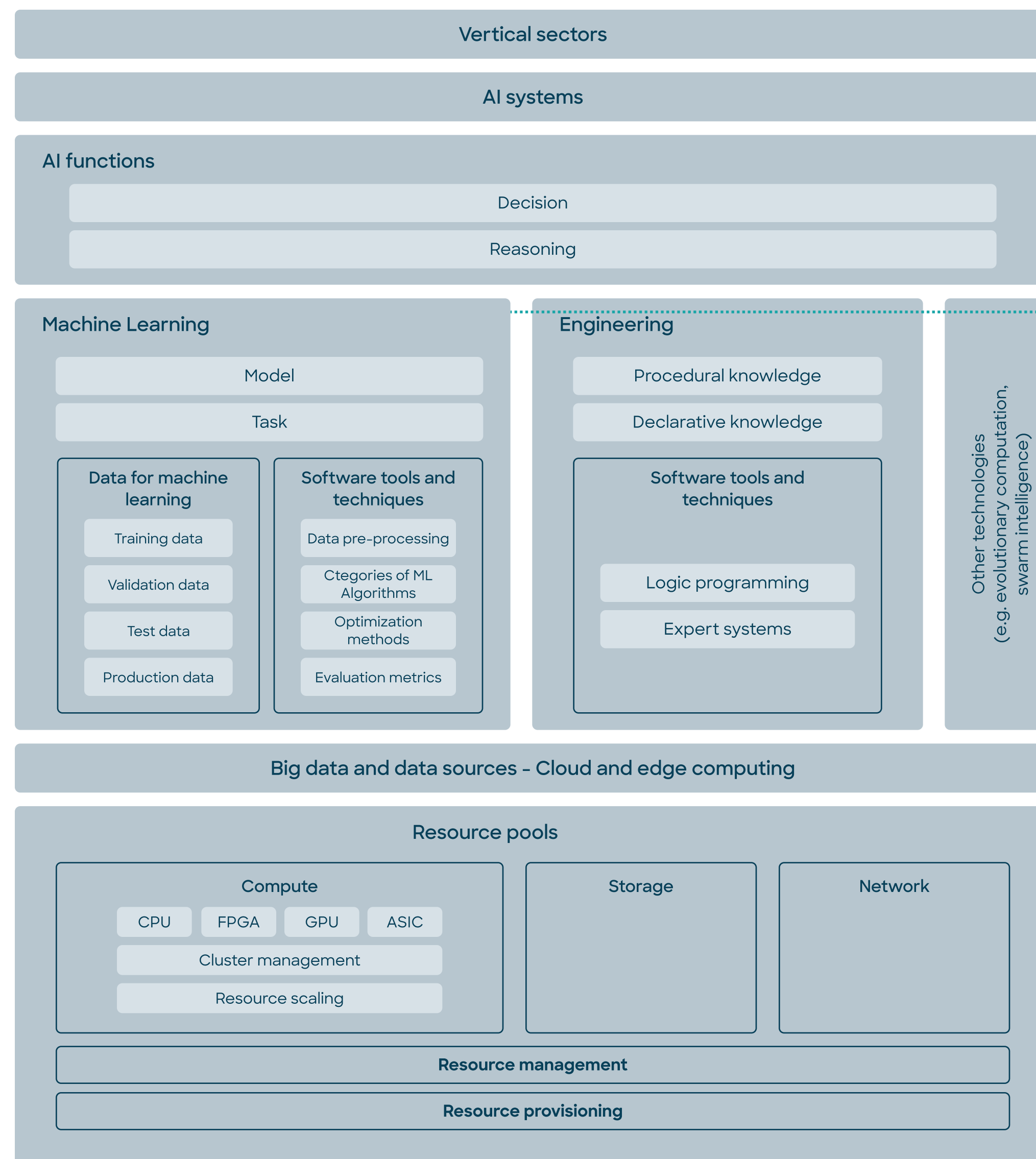


ARTICLE 14

Human Oversight

Article 14





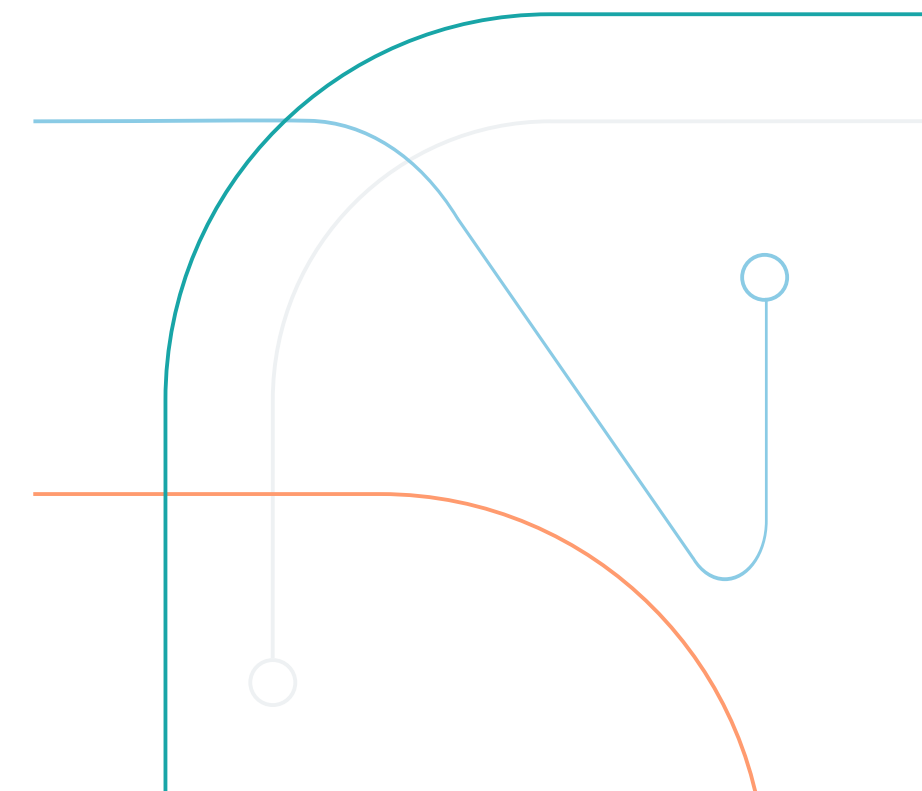
The **decision and reasoning capabilities** of the AI system need to be structured as such, that they can be actively monitored and are interpretable by a human operator

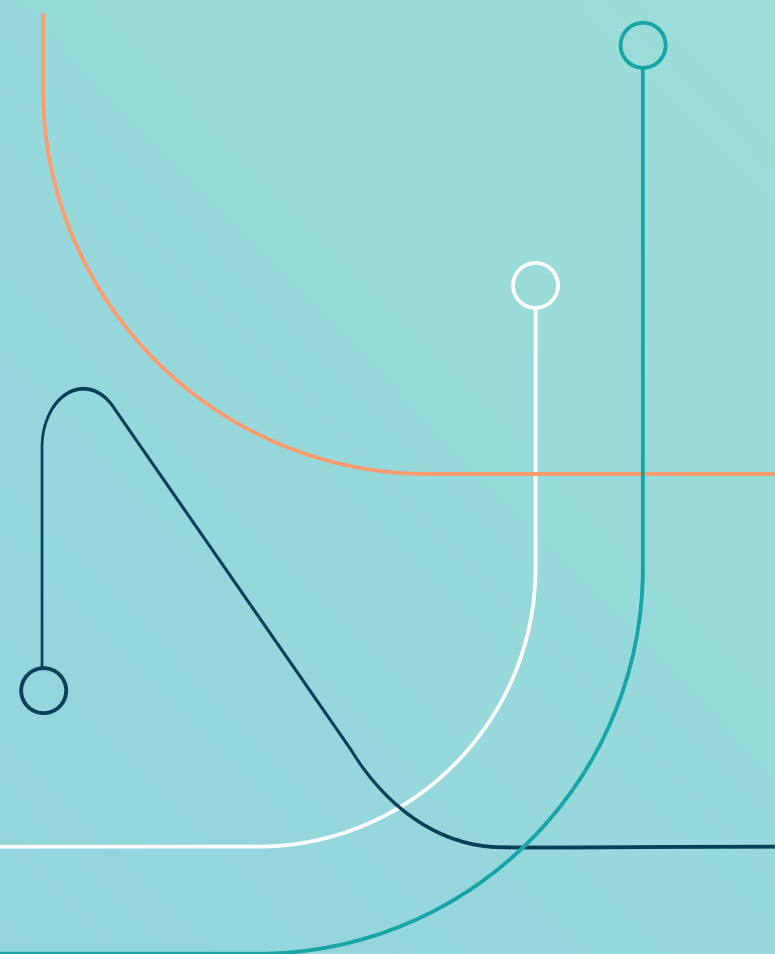
Machine learning models and tasks should be well-documented and transparent, allowing the **person responsible for oversight** to be aware of automation bias and recognise **faulty or otherwise incorrect performance**.

Figure 6 - AI ecosystem

Who is responsible for human oversight in this case?

What trigger conditions should lead to stopping the use of the system?

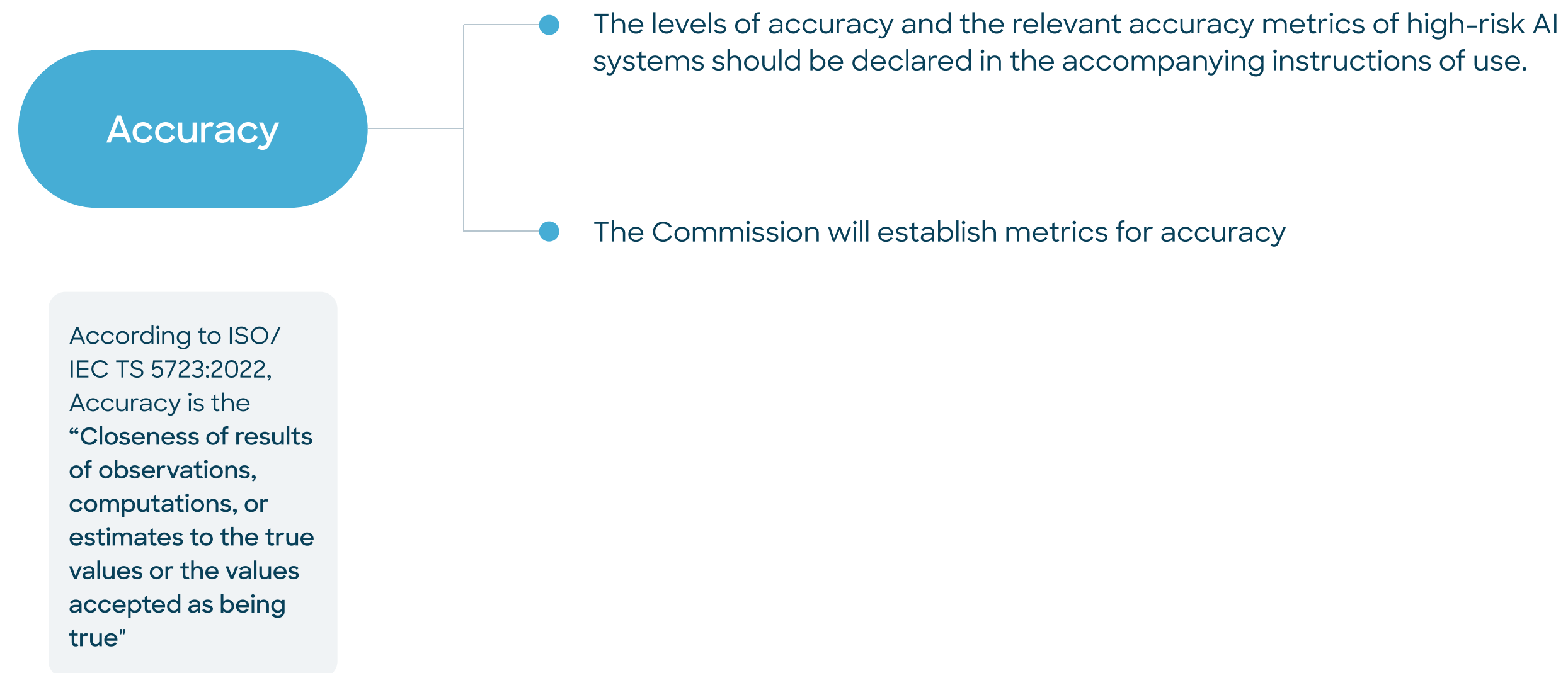


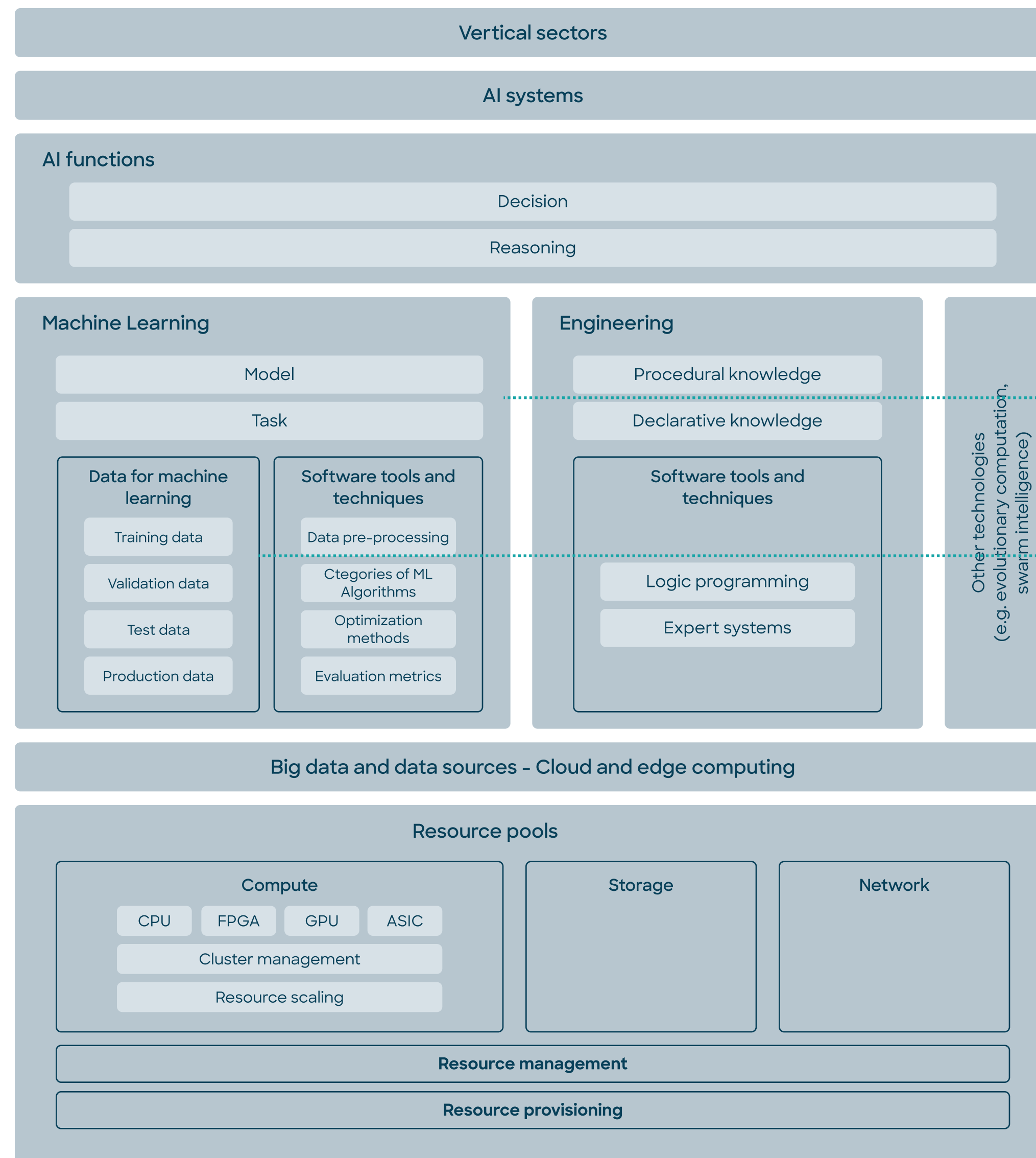


ARTICLE 15

Accuracy

Article 15





The **choice of algorithm** affects accuracy. Different algorithms have different strengths and are suited for different types of tasks

The **quality of the data, pre-processing activities**, and **validation data set** can impact the accuracy of the model



What **metrics** should be used & what are the baseline comparison?

How is **user-trust** ensured?

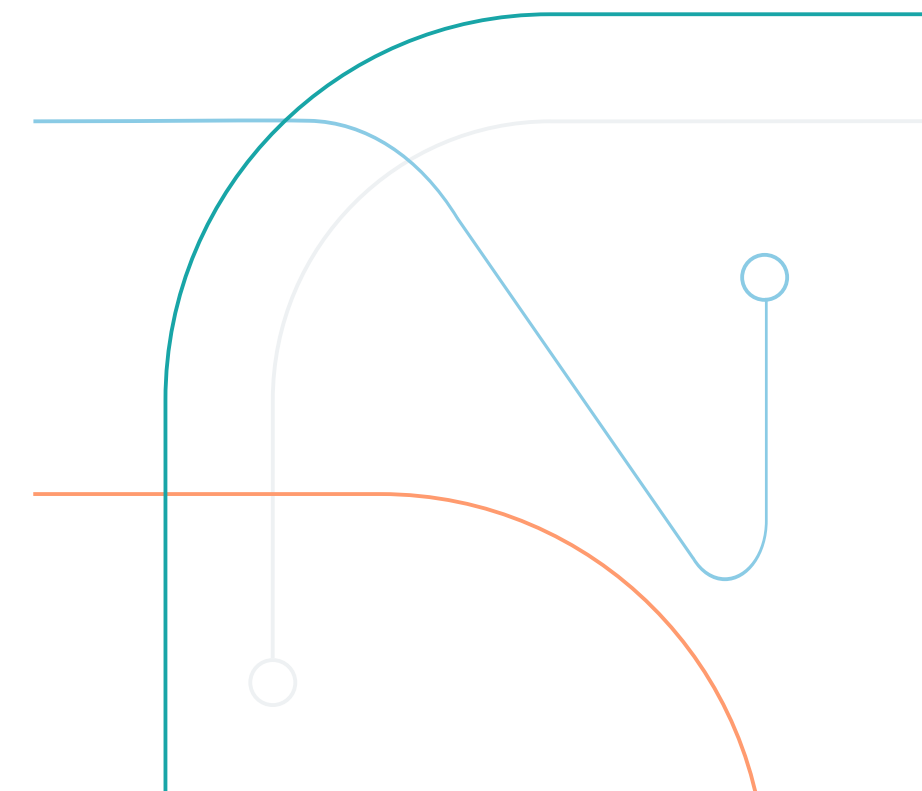
What **sector specific** requirements exist?

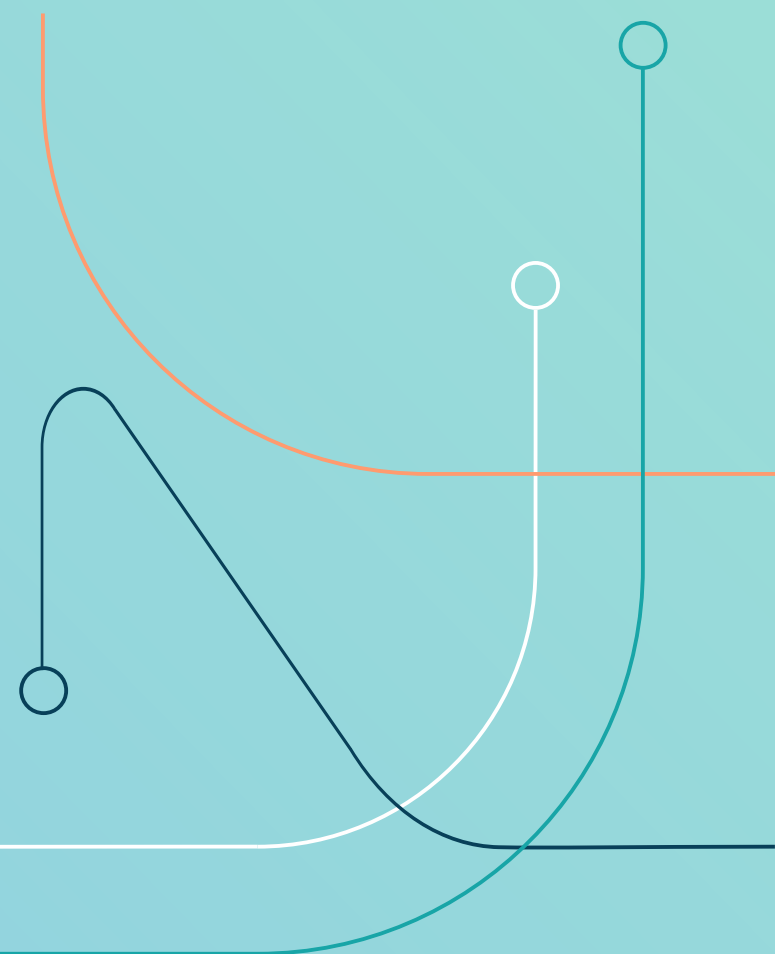
computational-centric measures (e.g., false positive and false negative rates), human-AI teaming, and external validity

Figure 6 - AI ecosystem

What is the appropriate measure for accuracy?

What performance level is satisfactory to your organization?

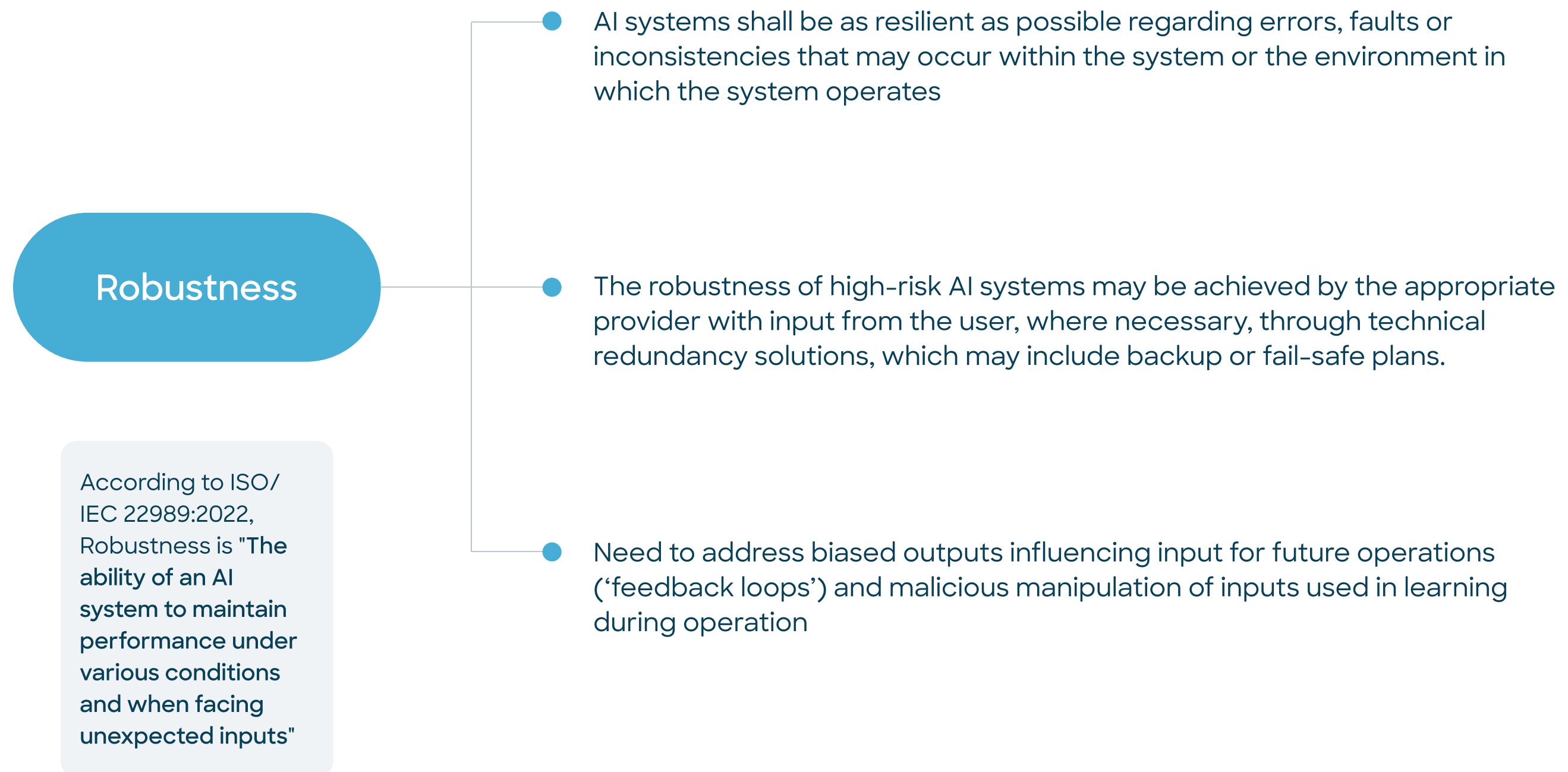




ARTICLE 15

Robustness

Article 15



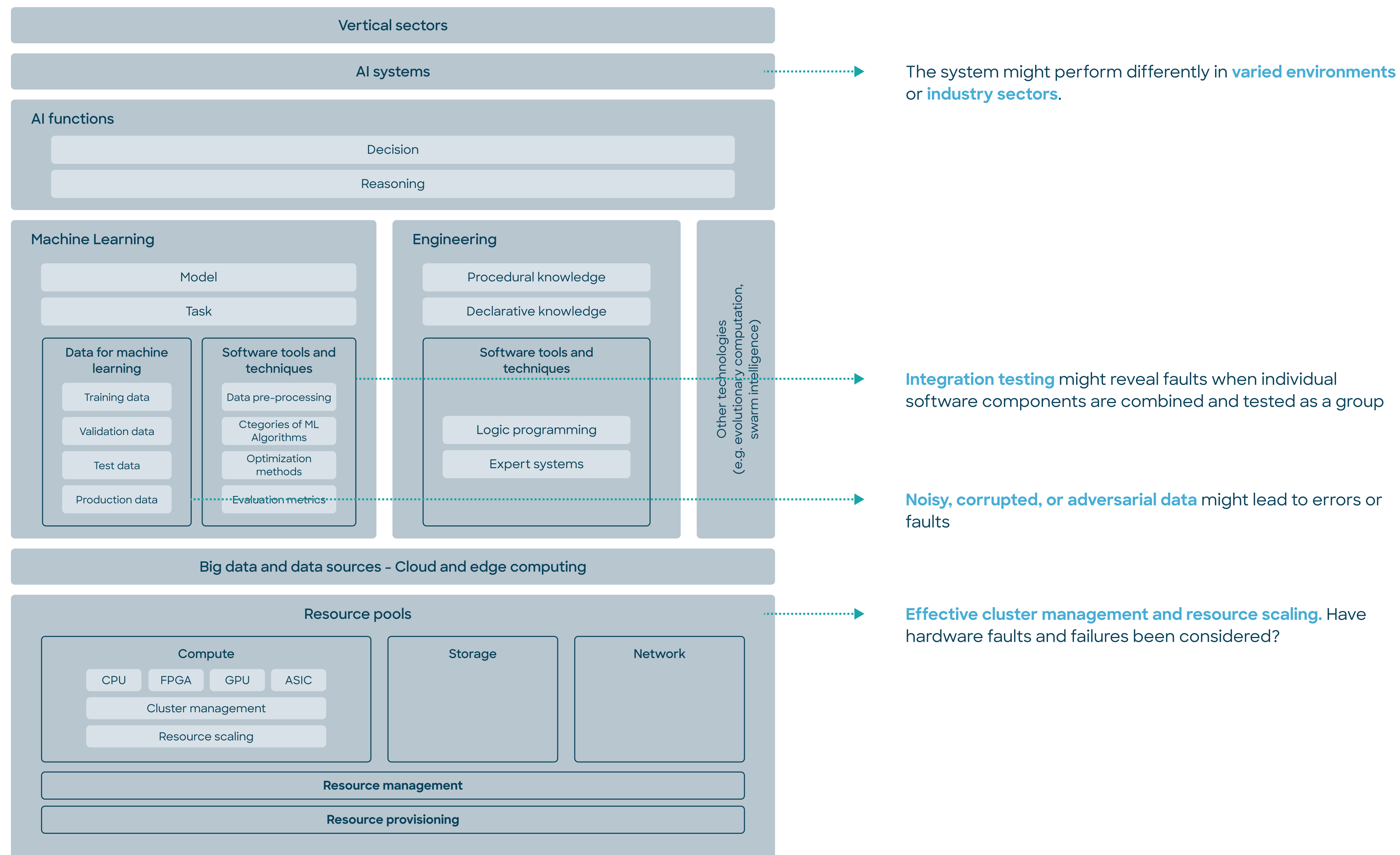
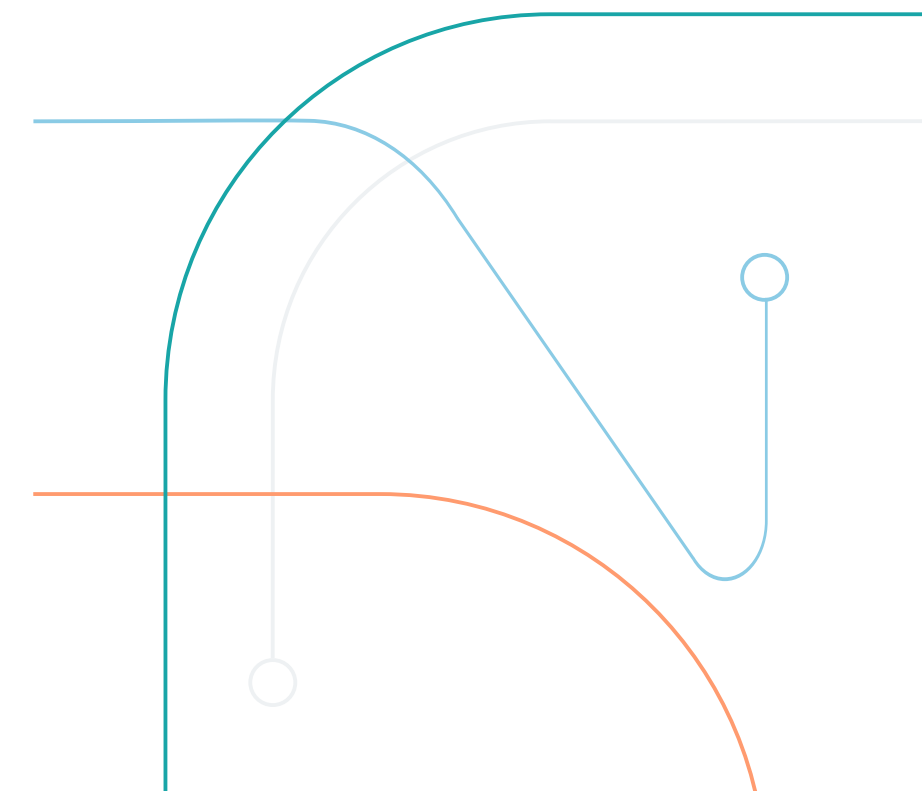
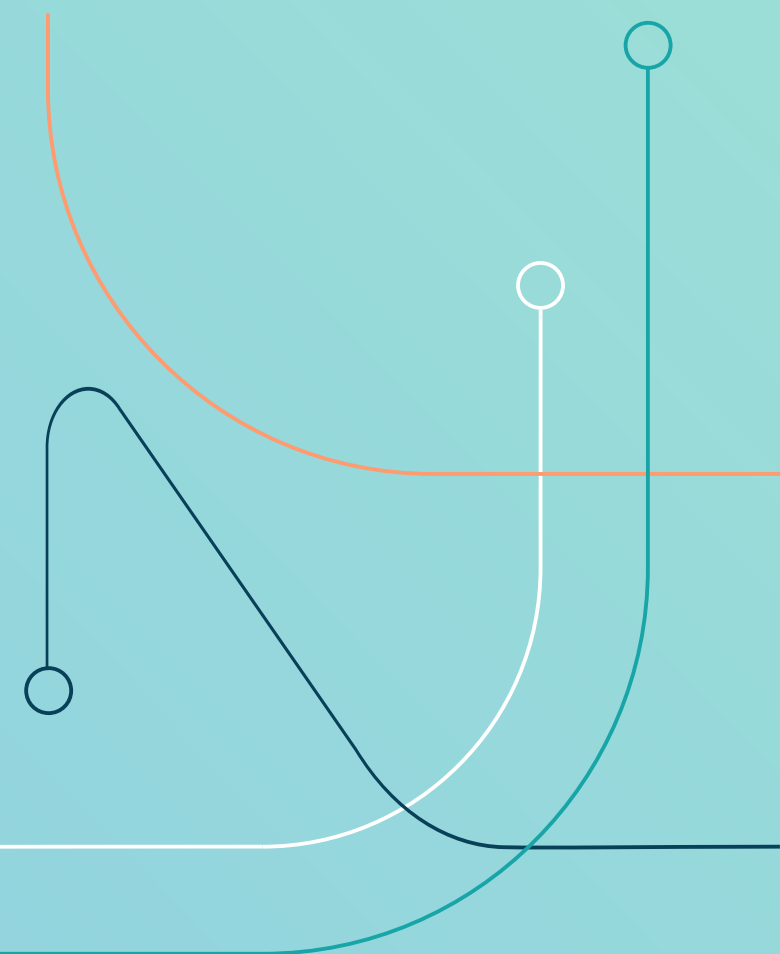


Figure 6 - AI ecosystem

What are the (unexpected) factors that may affect the performance of your system?

What fail-safe's can you design?

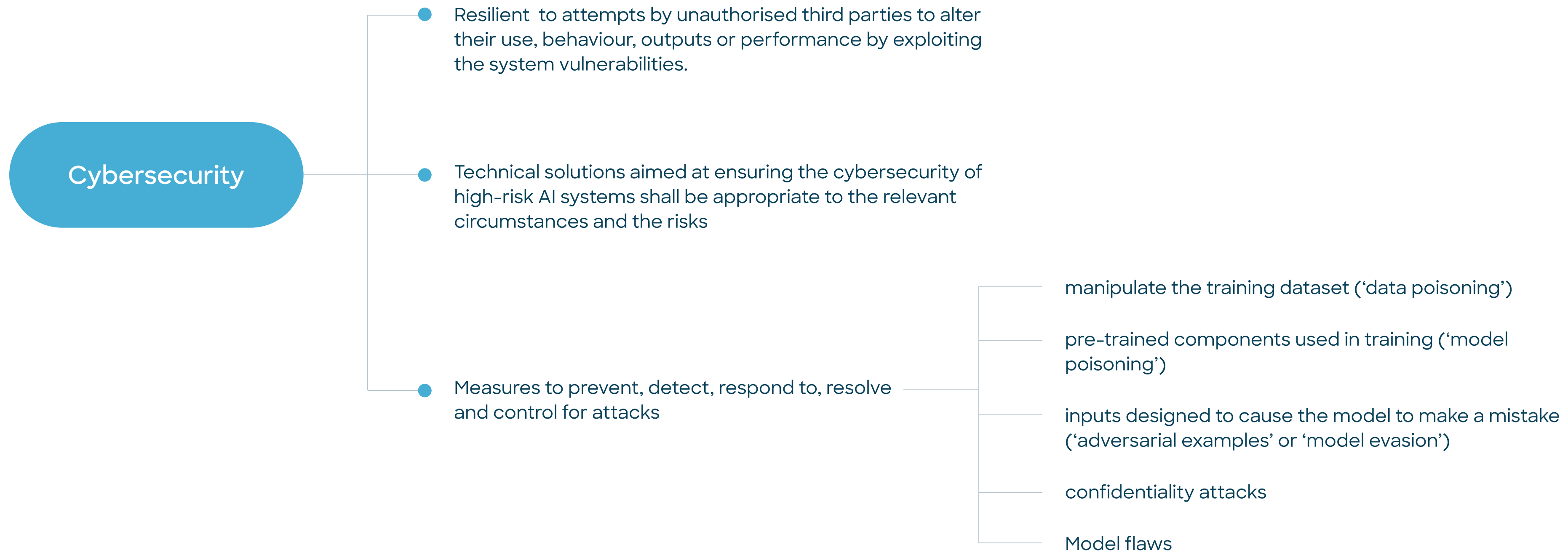




ARTICLE 15

Cybersecurity

Article 15



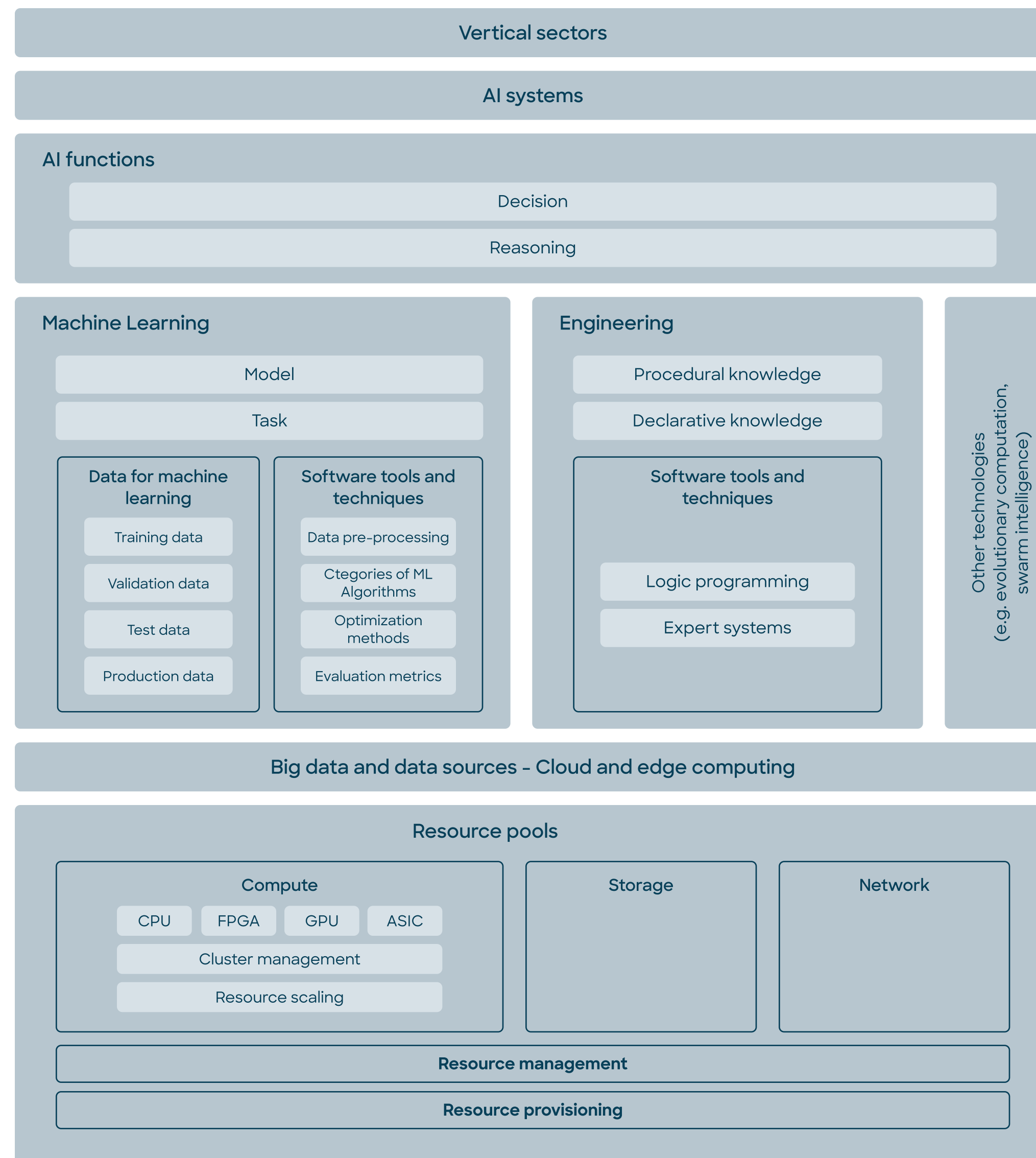


Figure 6 - AI ecosystem

European Union Agency for Cybersecurity Framework 2023

Layer III (Sector-specific). Various best practices that can be used by the sectoral stakeholders to secure their AI systems.

Layer II (AI-specific). Cybersecurity practices needed for addressing the specificities of the AI components with a view on their life cycle, properties, threats and security controls, which would be applicable regardless of the industry sector.

Layer I (cybersecurity foundations). The basic cybersecurity knowledge and practices that need to be applied to all ICT environments that host/operate/develop/integrate/maintain/supply/provide AI systems.

Shared responsibility?



1. Cloud Provider's Responsibility:

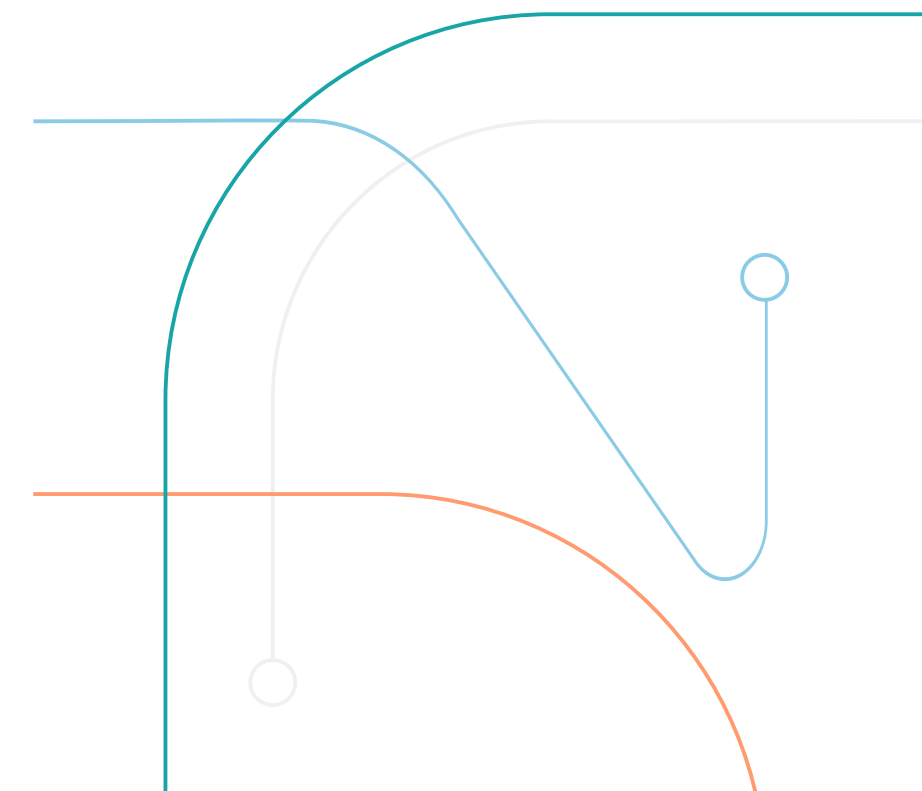
- a) Infrastructure Security
- b) Network Security
- c) Hardware and Software Maintenance

2. Cloud Customer's (User's) Responsibility:

- a) Data Security: ensure data is encrypted.
- b) Access Management: Implementing and managing user access, roles, and permissions.
- c) App Security: Users are responsible for ensuring that the apps they deploy on the cloud are secure, free from vulnerabilities, and regularly patched.
- d) End-point Protection

What infrastructure is your system running on?

What cyber-security law/standard are you already applying?



The appliedAI Institute for Europe aims to strengthen the European AI ecosystem by engaging in research, developing knowledge around AI, providing trusted AI tools, and creating educational as well as interactive formats around high-quality AI content.

As a non-profit subsidiary of the appliedAI Initiative, the Institute was founded in Munich in 2022. The appliedAI Initiative itself is a joint venture of UnternehmerTUM and IPAI. The Institute is managed by Dr. Andreas Liebl and Dr. Frauke Goll.

The appliedAI Institute for Europe focuses on the people in Europe. It pursues the vision of shaping a common AI community and providing high-quality content in the age of AI for the entire society. By promoting trustworthy AI, the Institute accelerates the application of this technology and strengthens trust in AI solutions.

With a focus on research, knowledge development, research and the provision of trusted AI tools, the appliedAI Institute for Europe provides a valuable resource for companies, organizations, and individuals looking to expand their knowledge and skills in AI. Through educational and interaction formats, the Institute enables an intensive exchange of expertise and fosters collaboration between stakeholders from different fields.

The appliedAI Institute for Europe invites companies, organizations, startups, and AI enthusiasts to benefit from the Institute's diverse offerings and resources. The appliedAI Institute for Europe is supported by the KI-Stiftung Heilbronn gGmbH.

For more information, please visit www.appliedai-institute.de