The foundational standards for AI

ISO/IEC 22989 and ISO/IEC 23053

Paul Cotton, Milan Patel, Wei Wei

ISO JTC1 IEC
INFORMATION TECHNOLOGY STANDARDS
SC 42 – Artificial Intelligence

Introduction

Foundational standards are the fundamental building blocks for other standards and applications. 22989 and 23053 are the first two foundational standards developed by JTC1 SC42 (AI) since it's creation in 2018. Both standards provide common terminologies and key concepts in AI and ML field. In May 2022, both AI foundational standards have been approved and will be published shortly.

ISO JTC1 IEC
INFORMATION TECHNOLOGY STANDARDS
SC 42 – Artificial Intelligence

ISO/IEC 22989 AI concept & terminology

- ✓ FDIS 22989 has been approved in May 2022, will be published shortly
- ✓ Standard terminology provides common language in AI world which helps people with different backgrounds to understand each other
  - ✓ 117 terms in 7 categories: AI general terms (e.g., AI system), data terms (e.g., test data, validation data), ML terms (e.g., model parameter), neural network terms (e.g., deep learning), trustworthiness terms (e.g., bias, explainability, transparency), NLP terms (e.g., machine translation), computer vision terms (e.g., image recognition).
- ✓ Key concept in AI world helps people to develop standards which are interoperable and coherent
  - ✓ basic concepts in AI, ML, data, trustworthiness as well as AI system life cycle, stakeholder roles, functional overview, eco system, application fields etc.

22989: AI system definition

Blind men and the elephant

A group of blind men heard that a strange animal, called an elephant, had been brought to the town, but none of them were aware of its shape and form. Out of curiosity, they said: "We must inspect and know it by touch, of which we are capable". So, they sought it out, and when they found it they groped about it. The first person, whose hand landed on the trunk, said, "This being is like a thick snake". For another one whose hand reached its ear, it seemed like a kind of fan. As for another person, whose hand was upon its leg, said, the elephant is a pillar like a tree-trunk. The blind man who placed his hand upon its side said the elephant, "is a wall". Another who felt its tail, described it as a rope. The last felt its tusk, stating the elephant is that which is hard, smooth and like a spear.

- ✓ The most challenging task that SC42 WG1 experts spent most time and effort in 22898
- ✓ Hard to reach consensus, because different people with different background has different view
- ✓ Many contributions from experts and NBs including definitions from internal/external sources such as OBP, EU AI HLEG, EU AI Act, OECD, Oxford etc
- ✓ Current approach contains two parts: main body describes what AI system does + note describes how AI system works

AI system

engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives

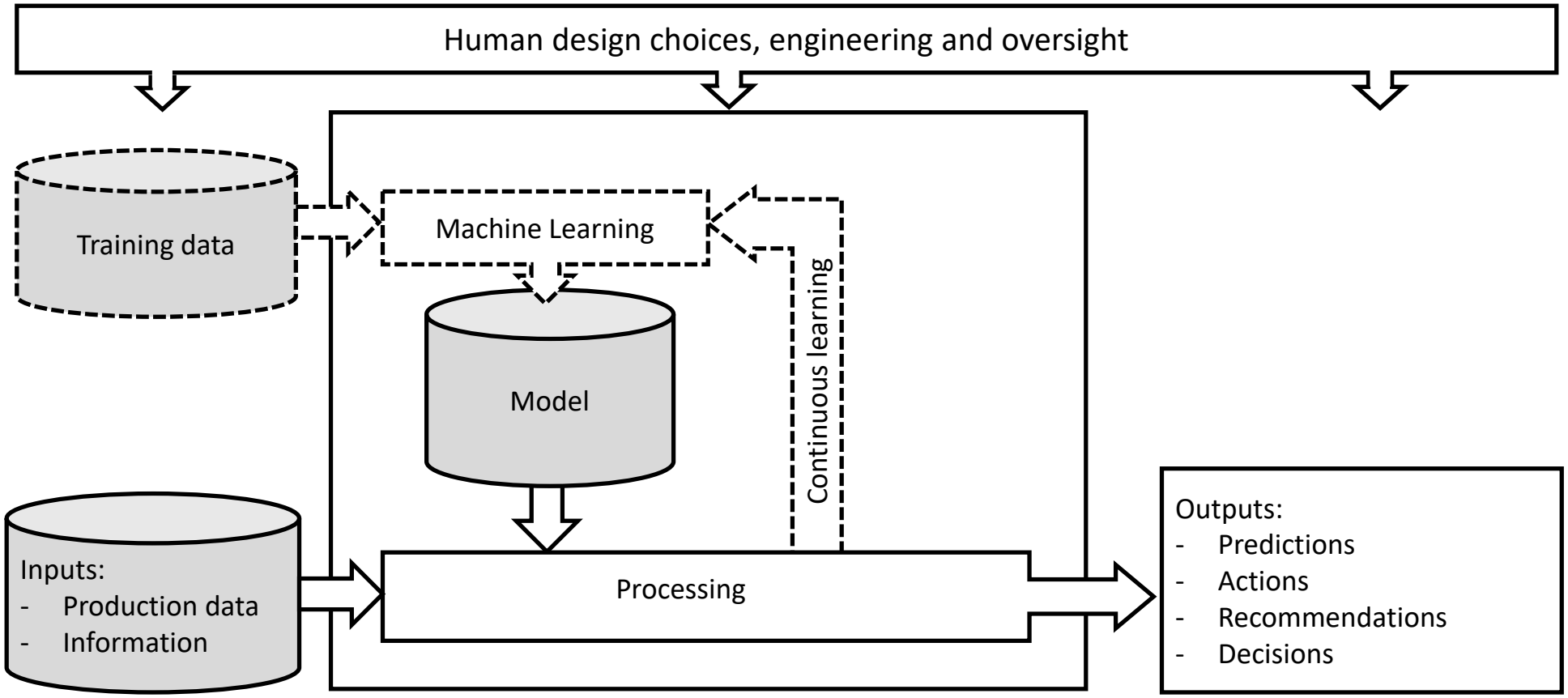
Note 1 to entry: The engineered system can use various techniques and approaches related to *artificial intelligence* (3.1.3) to develop a *model* (3.1.23) to represent data, *knowledge* (3.1.21), processes, etc. which can be used to conduct *tasks* (3.1.35).

Note 2 to entry: AI systems are designed to operate with varying levels of *automation* (3.1.7).

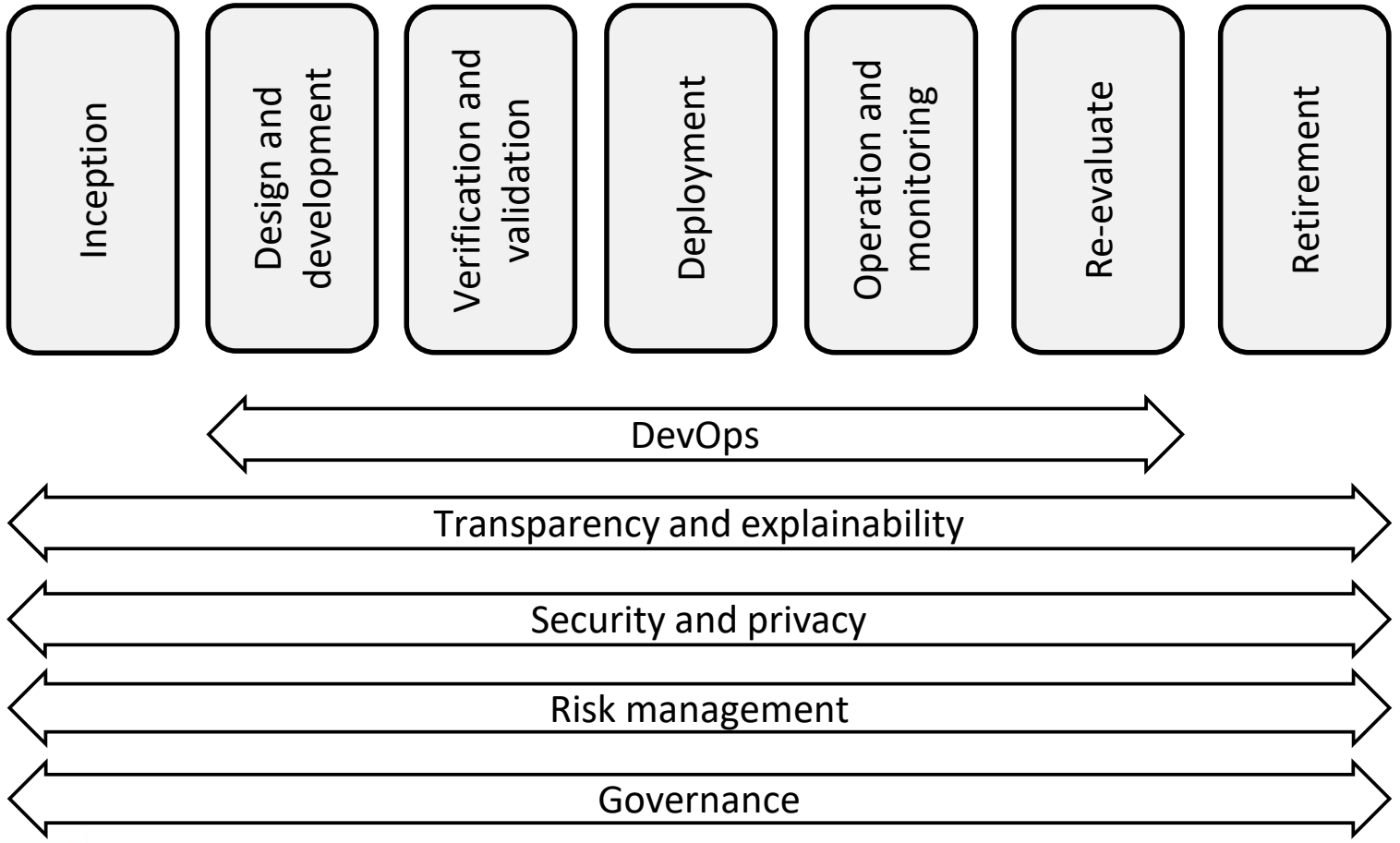
22989: Level of automations

		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	The system is capable of performing its entire mission without external intervention
		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

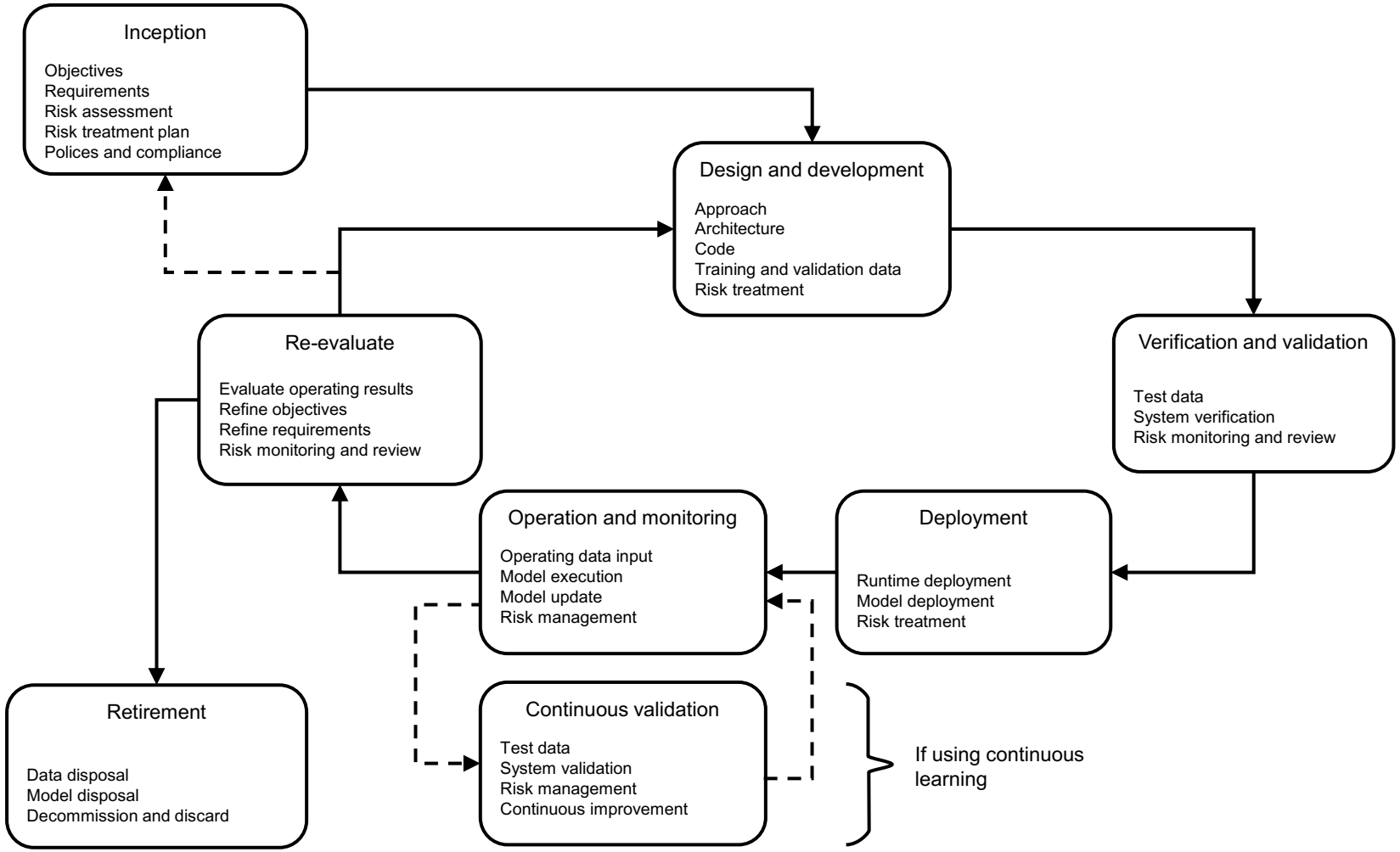
22989: AI functional overview



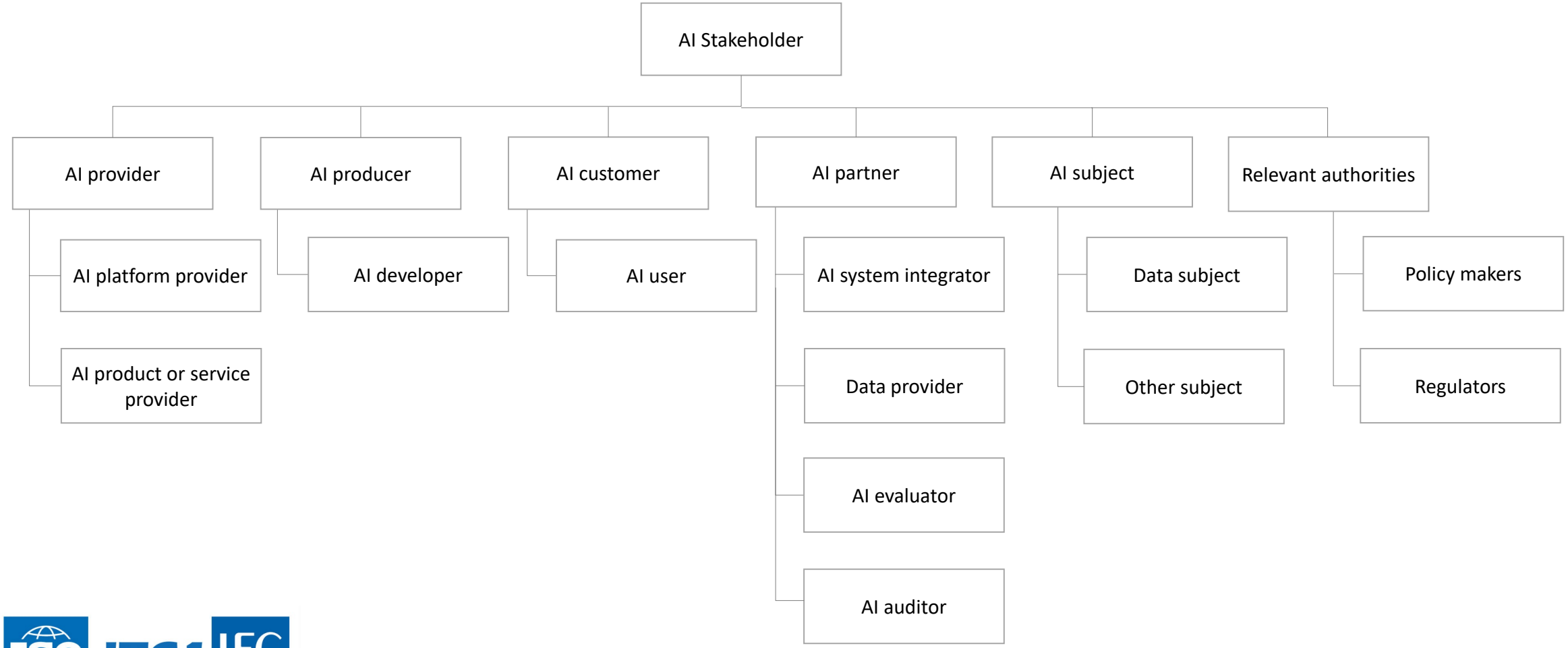
22989: AI system life cycle model



22989: AI system life cycle processes



22989: AI stakeholder roles



ISO/IEC 23053 Machine learning framework

Builds on ISO/IEC 22989

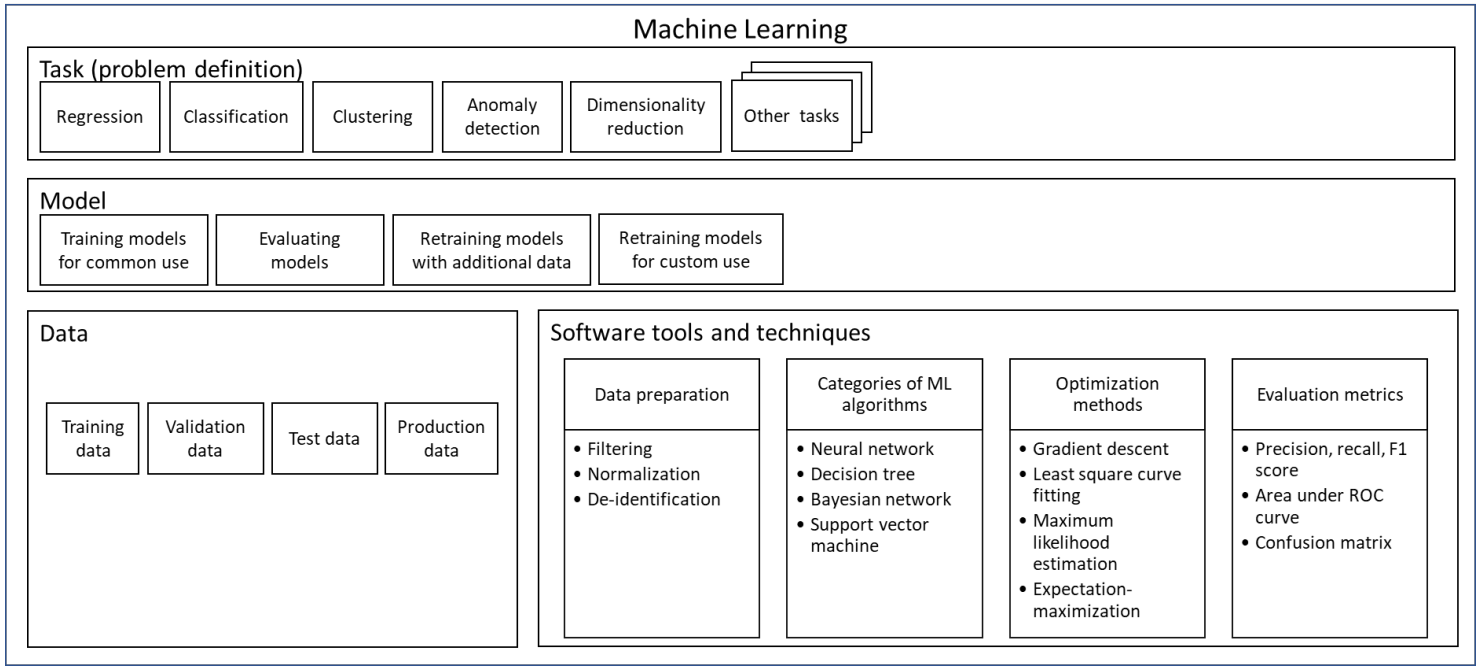
Deeper dive into concepts related to machine learning

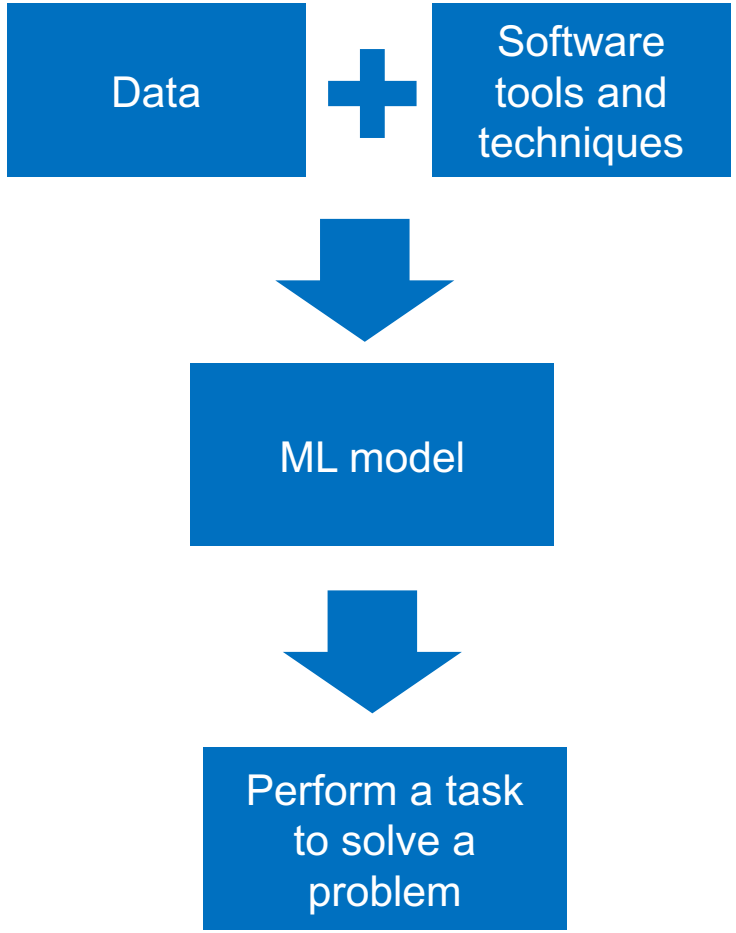
Arranging concepts into a framework to help clearly explain AI systems that use ML and how they are developed; includes descriptions of the data needed for the creation evaluation and use of an ML model

Aimed at a wide audience – experts and non-practitioners

machine learning:
process of optimizing model parameters through computational techniques, such that the model's behaviour reflects the data or experience

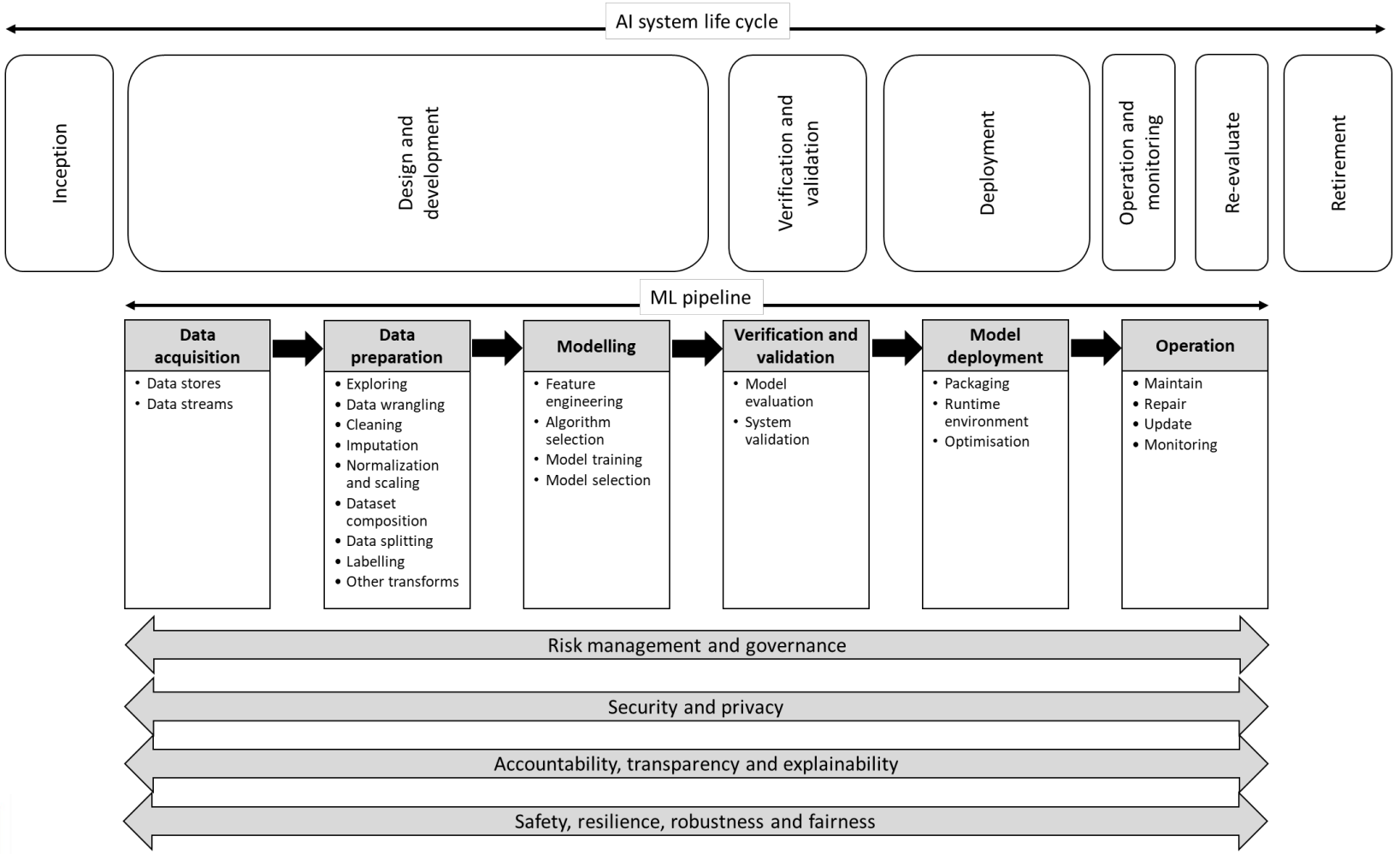
ISO/IEC 23053 Machine learning framework



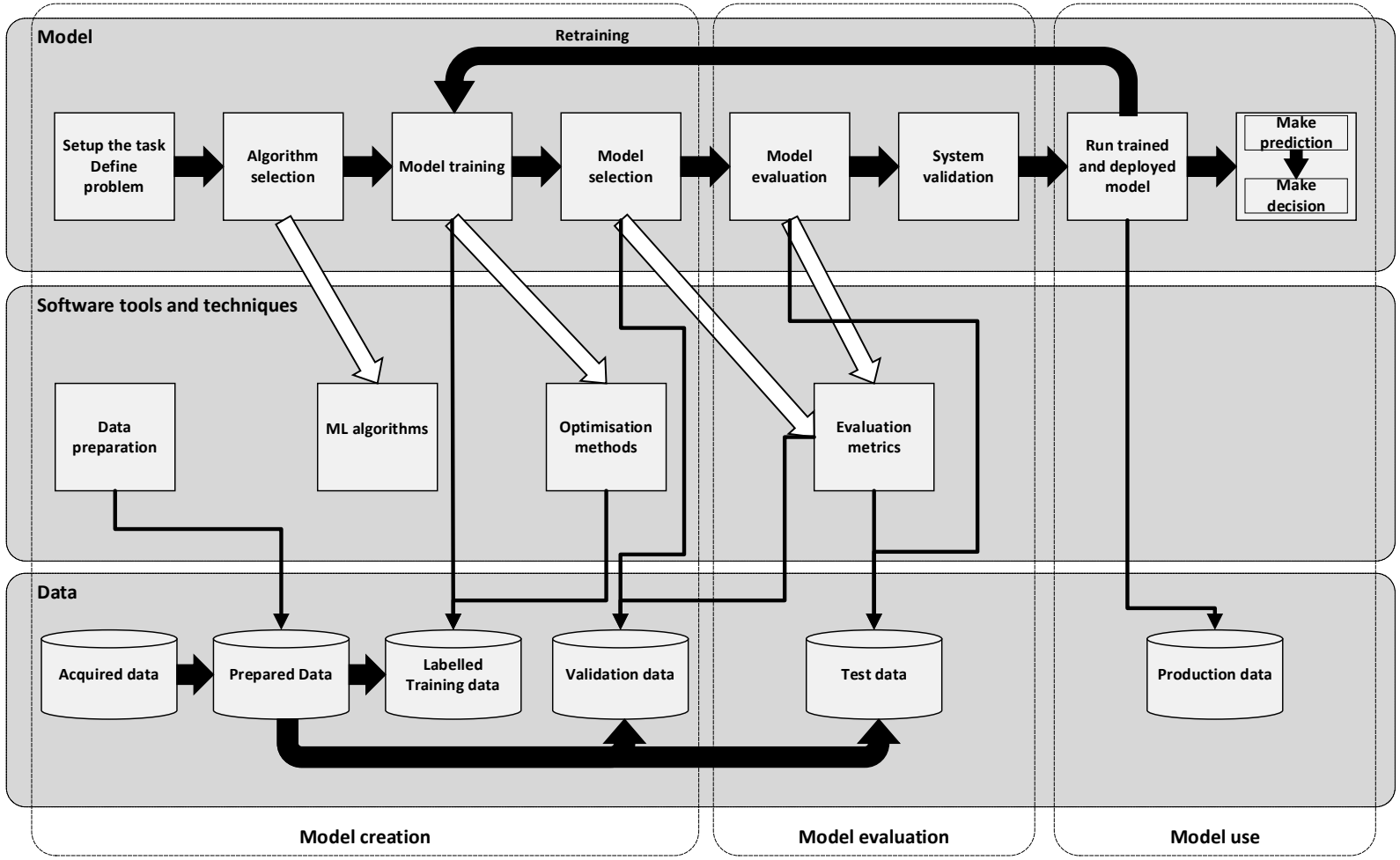


This framework is used for the structure of the document, but also to illustrate the machine learning process

Machine learning pipeline



Supervised machine learning process based on the ML pipeline



ISO/IEC 42001 AI management system

Requirements and guidance for establishing, implementing, maintaining and continually improving an AI management system.

Management processes to help an organization to develop or use AI systems responsibly in pursuing its objectives, help to meet regulatory requirements, obligations and expectations of interested parties.

Helps enable trust in the organization and the AI systems it develops or uses

ISO/IEC 42001 AI management system

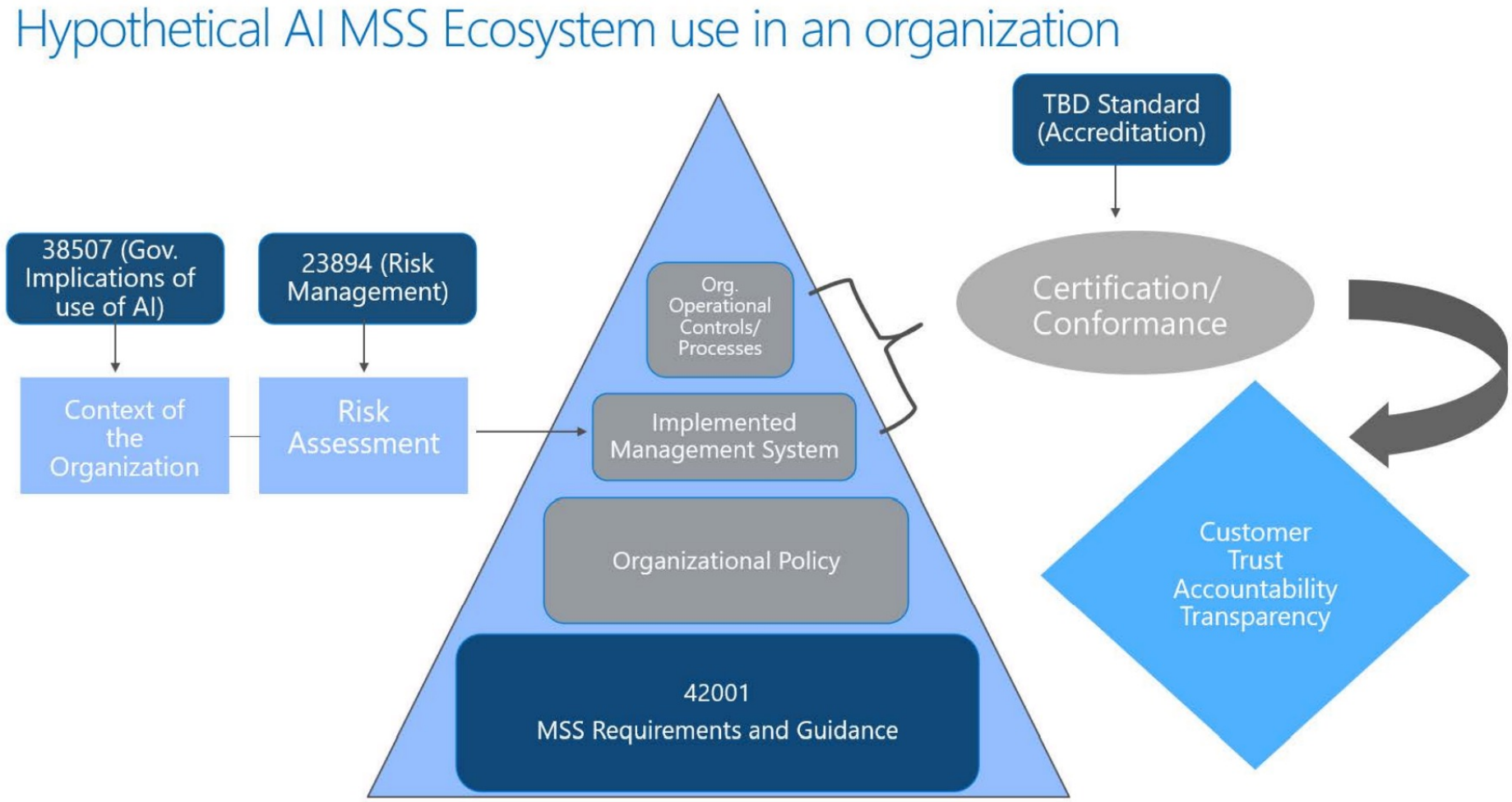
Risk based with controls and guidance to implement organizational and technical measures to mitigate risks

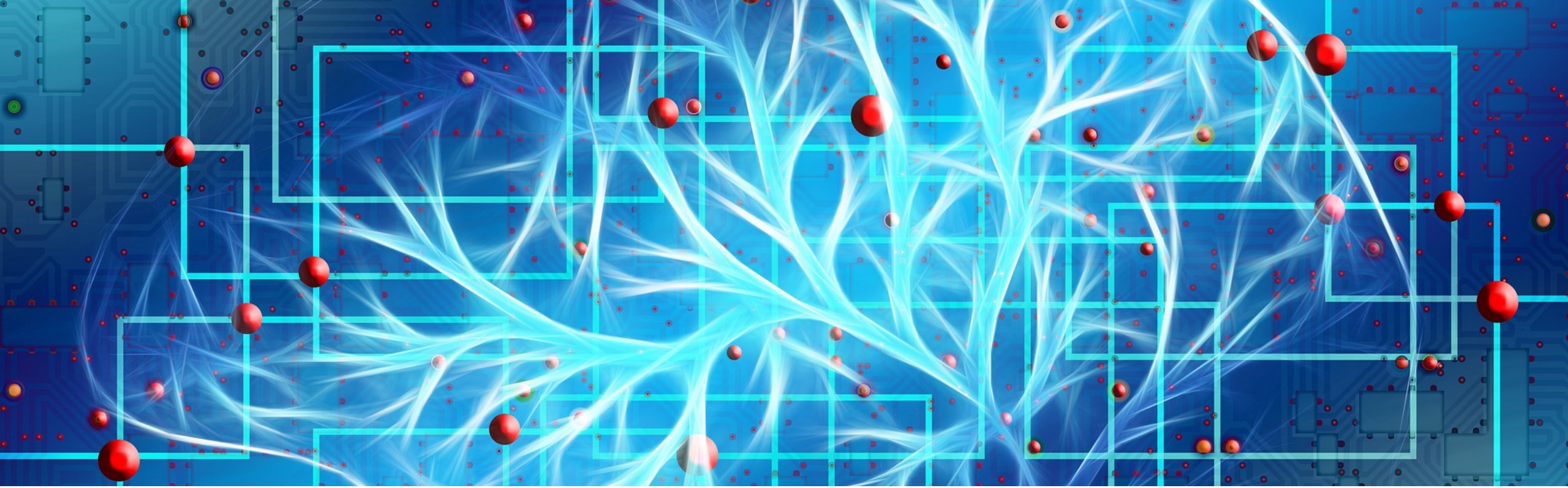
Incorporates the need for AI system impact assessments that can feed into organizational risk management processes

Can be integrated with existing management processes or other management system requirements

Can be used to underpin conformity assessment

Use of management system and related governance and risk standards





Summary

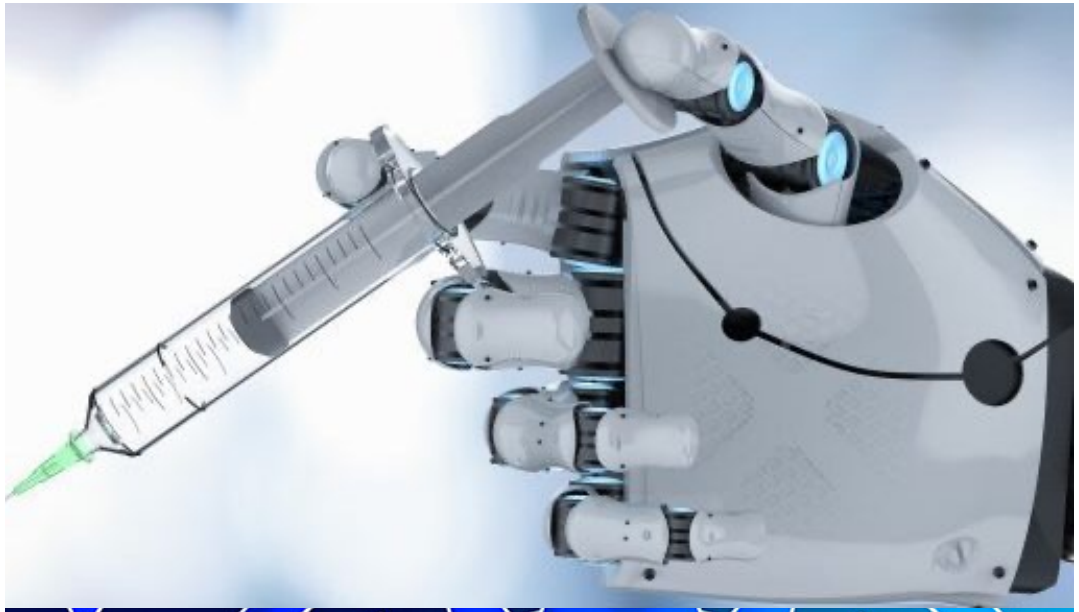
Foundational standards are the basis for development of other standards and applications in AI and its related field. For teaching and learning → it explains the fundamental part in AI. For communication → same terminology helps people with different background to understand each other better. For application → common concept and framework help people to develop interoperable and coherent applications or standards.

Thank you

Wei Wei (deweiwei@de.ibm.com)

Milan Patel (patel.milan@microsoft.com)

Paul Cotton (paul.cotton@outlook.com)



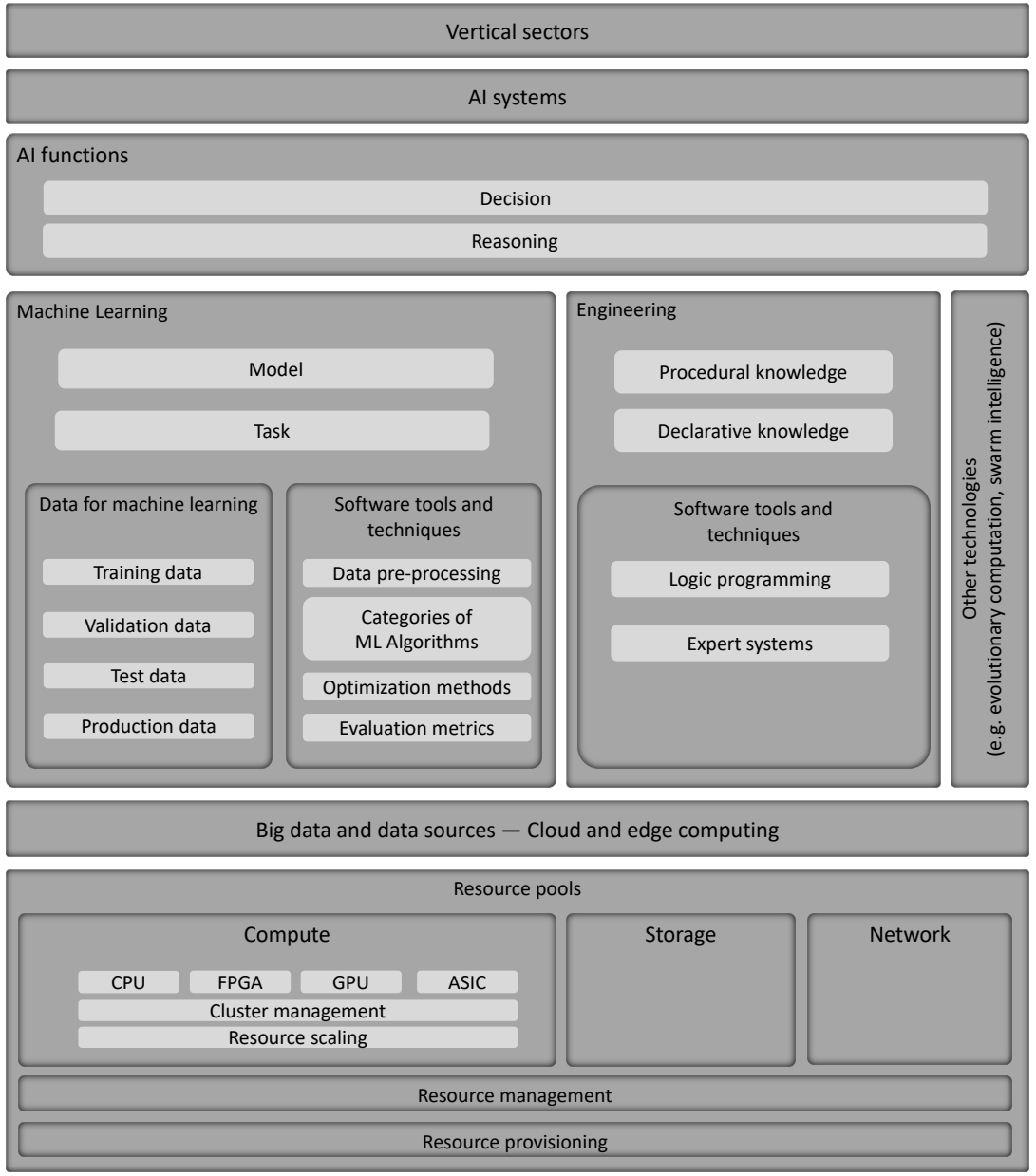
22989: Compare AI definition with EU AI Act

22989 AI System	AI system in EU AI Act
engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives Note 1 to entry: The engineered system can use various techniques and approaches related to artificial intelligence (3.1.3) to develop a model (3.1.23) to represent data, knowledge (3.1.21), processes, etc. which can be used to conduct tasks (3.1.35). Note 2 to entry: AI systems are designed to operate with varying levels of automation (3.1.7).	artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with ANNEX I List (a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; (c) Statistical approaches, Bayesian estimation, search and optimization methods.

22989: Compare AI definition with OECD

22989 AI System	AI system in OECD
engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives Note 1 to entry: The engineered system can use various techniques and approaches related to artificial intelligence (3.1.3) to develop a model (3.1.23) to represent data, knowledge (3.1.21), processes, etc. which can be used to conduct tasks (3.1.35). Note 2 to entry: AI systems are designed to operate with varying levels of automation (3.1.7).	An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. It does so by using machine and/or human-based inputs to: i. perceive real and/or virtual environments; ii. abstract such perceptions into models through analysis in an automated manner (e.g. with ML, or manually); and iii. use model inference to formulate options for information or action. AI systems are designed to operate with varying levels of autonomy.

22989: AI eco system



ISO JTC 1 IEC
INFORMATION TECHNOLOGY STANDARDS
SC 42 – Artificial Intelligence

22989: Mapping with OECD AI life cycle