INTERNATIONAL STANDARD

ISO/IEC 5338

First edition
2023-12

Information technology — Artificial intelligence — AI system life cycle processes

Technologies de l'information — Intelligence artificielle — Processus de cycle de vie des systèmes d'IA

Reference number
ISO/IEC 5338:2023(E)

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 42, *Artificial intelligence*.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Artificial intelligence (AI) systems in the fields of computer vision and image recognition, natural language processing, fraud detection, automated vehicles, predictive maintenance and planning have achieved remarkable successes. To build and maintain an AI system, it is an efficient approach to extend the life cycle processes for a traditional software system to include AI-specific life cycle characteristics.

An example of such a specific characteristic of an AI system life cycle is where a system employs machine learning (ML) using training data and it becomes necessary to retrain the ML model using new training data that is more representative of current production data.

ISO/IEC/IEEE 12207 describes software life cycle processes and ISO/IEC/IEEE 15288 describes system life cycle processes. While these life cycle processes are broadly applicable to AI systems, they require the introduction of new processes and the modification of existing processes to accommodate the characteristics of AI systems. This document extends the current generic life cycle process International Standards to make them applicable for AI systems so that the AI system life cycle can benefit from established models and existing practices. Some AI systems are in use in areas which are related to safety, such as health care or traffic control. Such safety critical AI systems need special attention and considerations as described in ISO/IEC TR 5469 [5].

Integrating the AI system life cycle into existing processes delivers efficiency gains, better adoption of AI and mutual understanding among AI system stakeholders as defined in ISO/IEC 22989. Such an integrated life cycle approach embraces the fact that AI systems typically are a combination of AI-specific elements and traditional elements such as source code and databases.

This document provides further details on AI system life cycle processes as discussed in ISO/IEC 42001[18].

Information technology — Artificial intelligence — AI system life cycle processes

1 Scope

This document defines a set of processes and associated concepts for describing the life cycle of AI systems based on machine learning and heuristic systems. It is based on ISO/IEC/IEEE 15288 and ISO/IEC/IEEE 12207 with modifications and additions of AI-specific processes from ISO/IEC 22989 and ISO/IEC 23053.

This document provides processes that support the definition, control, management, execution and improvement of the AI system in its life cycle stages. These processes can also be used within an organization or a project when developing or acquiring AI systems. When an element of an AI system is traditional software or a traditional system, the software life cycle processes in ISO/IEC/IEEE 12207 and the system life cycle processes in ISO/IEC/IEEE 15288 can be used to implement that element.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC/IEEE 15288:2023, *Systems and software engineering — System life cycle processes*

ISO/IEC/IEEE 12207:2017, *Systems and software engineering — Software life cycle processes*

ISO/IEC 22989:2022, *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*

ISO/IEC 23053, *Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 22989, ISO/IEC 23053, ISO/IEC/IEEE 15288, ISO/IEC/IEEE 12207 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
knowledge acquisition
process of locating, collecting, and refining knowledge and converting it into a form that can be further processed by a knowledge-based system

Note 1 to entry: Knowledge acquisition normally implies the intervention of a knowledge engineer, but it is also an important component of machine learning.

[SOURCE: ISO/IEC 2382:2015, 2123777, modified — Notes 2 to entry 3 to entry have been deleted.]

4 Abbreviated terms

AI artificial intelligence

ML machine learning

5 Key concepts

5.1 General

AI system life cycle consists of three types of processes:

— Generic processes: Processes that are identical to the processes defined in ISO/IEC/IEEE 15288 and ISO/IEC/IEEE 12207.

— Modified processes: Processes where elements are modified, added or removed from the ISO/IEC/IEEE 15288 and ISO/IEC/IEEE 12207 definition.

NOTE 1 The Clause for each of these “Modified processes” contains a subclause of AI-specific particularities that provide guidance to adapt the process to AI systems.

— AI-specific processes: Processes that are specific to characteristics of AI systems but are not based directly on any processes in ISO/IEC/IEEE 15288 and ISO/IEC/IEEE 12207.

AI system life cycle processes in Clause 6 are presented as generic, modified or AI-specific. Figure 1 shows the life cycle processes of AI system, grouped by type, and compared to ISO/IEC/IEEE 15288:2023, Figure 4.

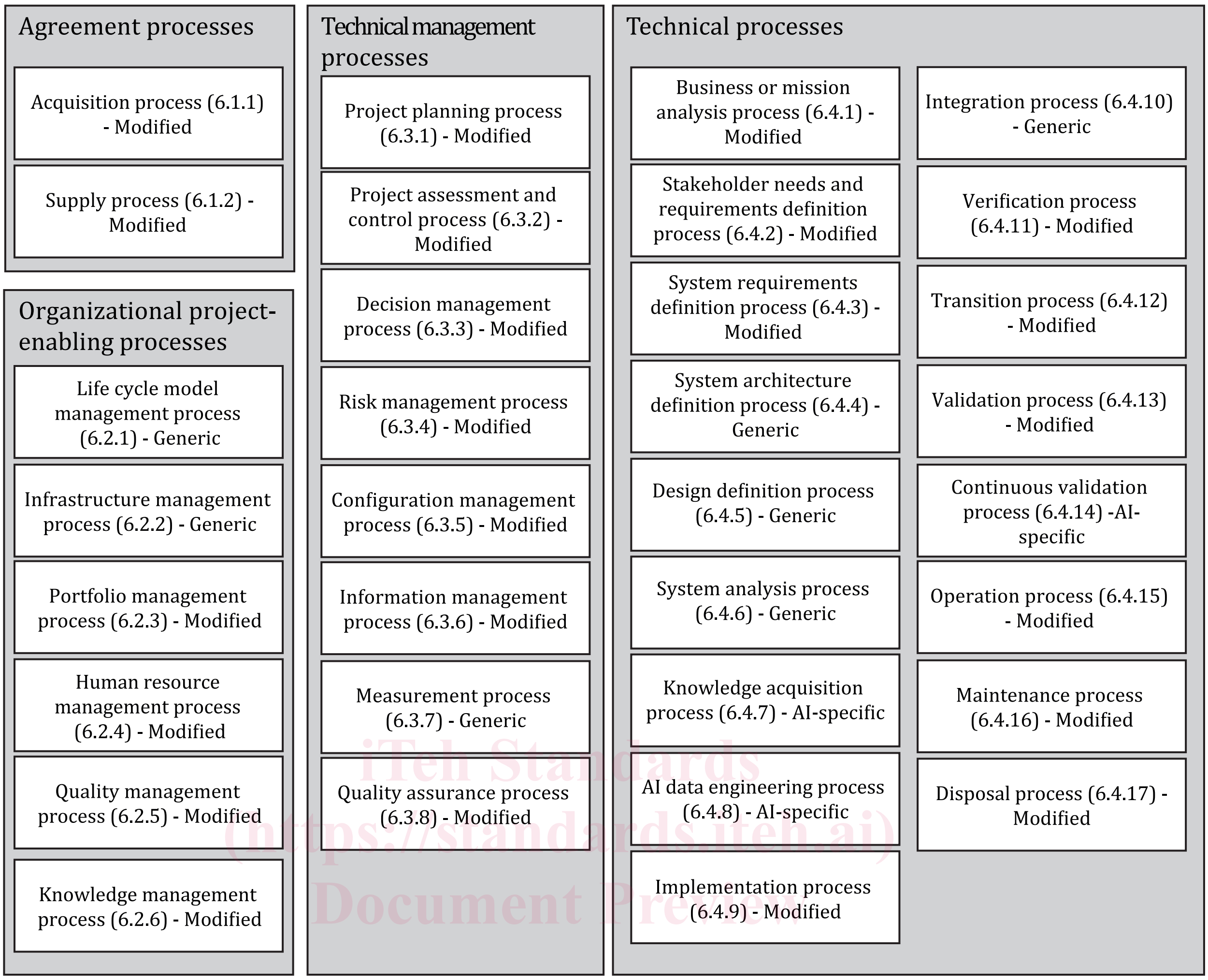

Figure 1 — AI system life cycle processes relative to ISO/IEC/IEEE 15288:2023, Figure 4

The following aspects of AI systems are key factors that differentiate the life cycle processes from those that are traditional systems:

— Measurable potential decay: Since AI models aim to model a desired behaviour which can change over time, measuring and monitoring any deviations of the production data (data drift) or deviations towards the desired output (concept drift) can be required. The changing of desired behaviour is not restricted to AI systems only, but for AI models this is uniquely measurable by validating input and output.

— Potentially autonomous: AI system's ability to make automated, complex and fast decisions creates the potential to replace actions or processes otherwise executed by humans. Consequently, AI systems can require extra attention to ensure fairness, security, safety, privacy, reliability, transparency and explainability, accountability, availability, integrity and maintainability. The more likely an AI system is able to do harm, the more important this extra attention becomes. See ISO/IEC TR 24368[14] for an overview of ethical and societal concerns in the development and deployment of AI systems. See ISO/IEC 23894[11] for more information of risk management of AI systems.

— Iterative requirements and behaviour specification: AI systems can be based on iterative and agile requirements specification, knowledge specification, behaviour modelling and usability design. AI system development can take place through cycles of requirements specification, prototype demonstration and requirements refinement. This aspect differs from traditional software applications based on fixed, well-defined requirements. Further, as AI systems are used, the requirements can also evolve as unseen situations arise and refined requirements, specifications and gaps are identified.

— Probabilistic: Decisions made by AI systems based on machine learning are inherently probabilistic. Therefore, it is important for stakeholders to recognize that decisions made by AI systems are not always correct. Formally testing the correctness of models has inherent limitations and uncertainties when it comes to guarantees.

— Reliant on data: AI systems based on machine learning rely on sufficient, representative data to train, test and validate models. The behaviour of machine learning models is not programmed but is instead learned from the data. Because of this, it is important that particular consideration be given to the data (e.g. data quality) that are required for an AI system for training, testing, verification and validation.

— Knowledge intensive: For heuristic models, knowledge acquisition is of relatively high importance, since the knowledge is coded explicitly in the model and determines its correctness.

— Novel: New knowledge and skills can be required for organizations designing, developing or using AI systems. Other stakeholders, such as AI system users, can be unfamiliar with AI. This can cause trust and adoption challenges. The novelty of AI can cause overconfidence and enthusiasm without fully accounting for AI system risks. The perception that AI systems can eventually replace humans or demonstrate intelligence can also impact how stakeholders view AI systems.

— Incomprehensible: In case of heuristic models or machine learning, model behaviour is emergent in the sense that it is not explicitly programmed but is instead the indirect result of knowledge engineering or derived from the training data. Stakeholders can find AI systems to be less predictable, explainable, transparent, robust and understandable than explicitly programmed systems. This can reduce trust in AI systems.

NOTE 2 A high-level overview of AI ethical and societal concerns can be found in ISO/IEC TR 24368.[14] More information on addressing ethical concerns during system design can be found in IEEE 7000-2021[20].

5.2 AI system concepts

A model can be a machine learning model which has learned how to compute based on data, or it can be a heuristic model engineered based on human knowledge. In a heuristic model, the computations are engineered explicitly (procedural), implicitly by specifying rules or probabilities (declarative), or both.

In the case of machine learning, the data are the primary input for the model. For a heuristic model, the primary input is knowledge. Regardless, both data and knowledge are required in either case. Data are needed to test heuristic models and to perform analysis to build the knowledge. Knowledge is required to understand the context in which a machine learning model operates and to help select and prepare data for training and testing.

For traditional systems, both knowledge and data are often important as well. Knowledge can be required to implement business logic. Data typically plays an important part in any data processing system and can be required for functional testing.

The differentiation between an AI system and an AI application is provided in ISO/IEC 5339[3]. The distinguishing characteristics of AI applications are also defined in ISO/IEC 5339[3].

5.3 AI system life cycle model

The AI system life cycle model describes the evolution of an AI system from inception through retirement. This document does not prescribe a specific life cycle. Instead, it concentrates on AI-specific processes that can occur during the system life cycle. AI-specific processes can occur during one or more of the life cycle stages and individual stages of the life cycle can be repeated during the system's existence. For example, during the re-evaluation stage development and deployment can be repeated multiple times to implement bug fixes and system updates.

A system life cycle model helps stakeholders build AI systems more effectively and efficiently. International Standards are useful in developing the life cycle model, including ISO/IEC/IEEE 15288 for systems as a whole, ISO/IEC/IEEE 12207 for software and ISO/IEC/IEEE 15289[10] for system documentation. These International Standards describe life cycle processes for traditional systems. Figure 2 is based on ISO/IEC 22989:2022, Figure 3. It provides an example of the stages and high-level processes that can be applied to the development and life cycle of AI systems. For details, see ISO/IEC 22989:2022, 6.1.

The AI system life cycle or any subset of its stages can be owned and managed by separate organizations or entities (e.g. acquisition and provision of data, ML model or the code for other components used for the AI system development or deployment). Additionally, an organization can depend on other organizations to establish the infrastructure or to provide the necessary capability of the AI system life cycle (e.g. infrastructure setup cutting across on-premise, cloud-based or hybrid). This document takes into account the implications, specifics and associated risks of the AI system supply chain to propose new processes, adapt and tailor existing processes to build an AI system across organizational boundaries.

In addition, certain domains have specific life cycle International Standards such as medical devices, where IEC 62304:2006+A1:2015[19] applies. Organizations should consider the AI specifics described in this document together with IEC 62304:2006+A1:2015[19] when implementing such domain specific standards.

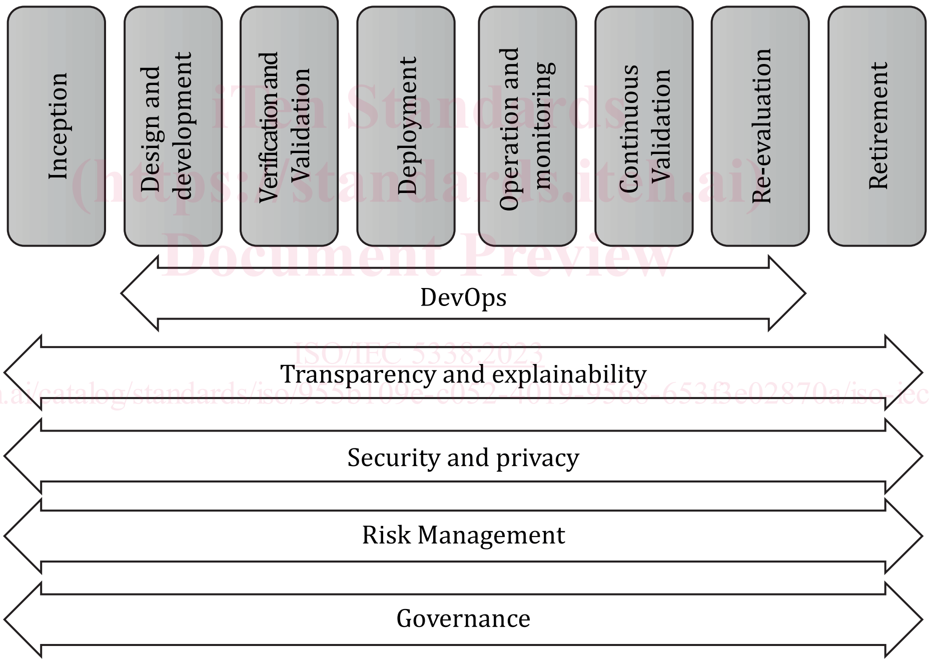

Figure 2 — Example of AI system life cycle model stages and high-level processes

The stages in Figure 3 are based on the stages described in ISO/IEC 22989 together with grouping of technical processes described in this document. The stage “continuous validation” is not marked as “in case of continuous learning”, in contrast with the example life cycle model in ISO/IEC 22989:2022, Figure 4. The continuous validation stage is also applicable in situations without continuous learning, for example, to detect data drift, concept drift or to detect any technical malfunctions.

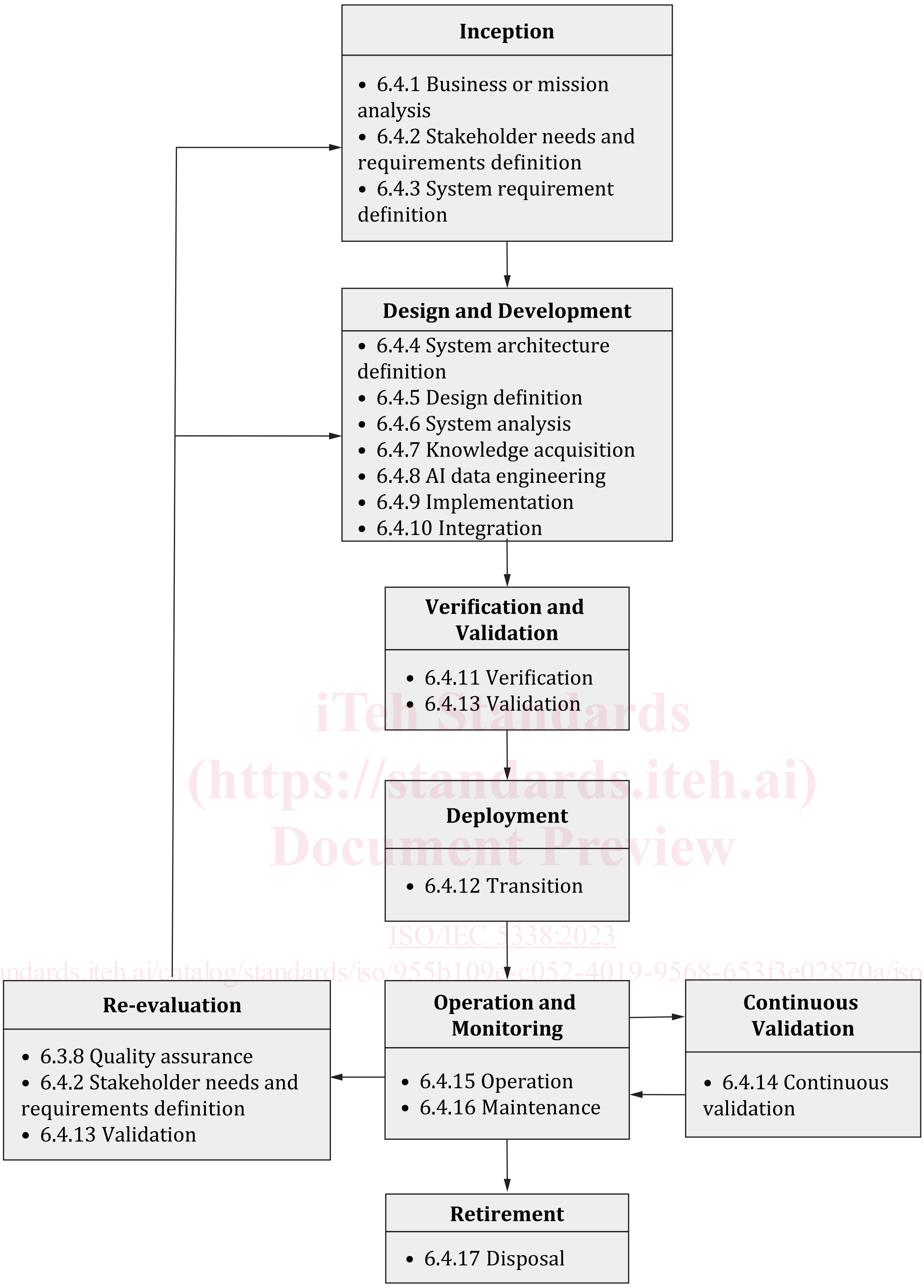

Figure 3 — AI system life cycle stages with technical processes

The concept of stages is meant to group activities that have a certain chronological order, to illustrate their dependency, but it does not suggest complete separation of activities in time or in the organization. For example, in agile software development, development and operation are distinct stages which are performed concurrently. Nevertheless, a piece of functionality first should be implemented before it can be verified and then deployed.

