International Standard

ISO/IEC 5392

Information technology — Artificial intelligence — Reference architecture of knowledge engineering

First edition
2024-03

Reference number
ISO/IEC 5392:2024(en)

Contents

Page

Foreword v
Introduction vi
1 Scope 1
2 Normative references 1
3 Terms and definitions 1
4 Abbreviated terms 5
5 Knowledge engineering system-of-interest 6
5.1 General 6
5.2 Important elements of knowledge engineering 6
5.3 Relationship between KE and AI systems 8
6 KE stakeholders 10
7 Concerns of KE stakeholders 12
7.1 Safety and security 12
7.2 Reliability 13
7.3 Availability 13
7.4 Construction quality 13
7.5 Responsibility 13
7.6 Bias reduction 14
8 Reference architecture of KE 14
8.1 General 14
8.2 User view of KE 14
8.2.1 Data supplier 14
8.2.2 Fundamental technology supplier 15
8.2.3 Algorithm supplier 15
8.2.4 System coordinator 16
8.2.5 Knowledge service provider 16
8.2.6 Knowledge applier 17
8.2.7 Knowledge ecosystem partner 17
8.3 Functional view of KE 17
8.3.1 Functional architecture of KE 17
8.3.2 KE infrastructure layer 19
8.3.3 KE construction layer 19
8.3.4 KE platform layer 19
8.3.5 KE application layer 20
8.3.6 Multi-layer functions 20
8.4 KE distribution architecture 20
8.4.1 General 20
8.4.2 Distributed architecture with semantic web services 21
9 Key technologies of KE and computational methods 22
9.1 Knowledge representation 22
9.1.1 General 22
9.1.2 Knowledge representation quality 23
9.2 Knowledge modelling 23
9.3 Knowledge acquisition 24
9.4 Knowledge storage 24
9.5 Knowledge fusion 24
9.6 Knowledge computing 24
9.7 Knowledge visualization 25
9.8 Knowledge maintenance 25
9.9 Knowledge exchange 25
10 Enabling technologies and digital infrastructure of KE 25

10.1 Enabling technologies 25
10.1.1 Machine learning 25
10.1.2 Natural language processing 25
10.1.3 Speech processing 26
10.2 Digital infrastructure 26
10.2.1 Big data 26
10.2.2 Cloud computing 26

Annex A (informative) **Examples of fundamental KE tools** **27**

Annex B (informative) **Specifications related to KE** **28**

Annex C (informative) **Characteristics of typical KE applications** **30**

Annex D (informative) **KE life cycle** **32**

Annex E (informative) **Building a solution architecture integrating ISO/IEC/IEEE 42010** **34**

Bibliography **41**

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

Knowledge-driven AI applications have gradually gained attention. In knowledge engineering (KE), knowledge is automatically or semi-automatically acquired from information sources, which in turn are generated by processing huge-scale multi-source heterogeneous data. The knowledge is integrated into knowledge-based systems and used to provide intelligent knowledge-driven services. One of the objectives of KE is to represent and transfer human knowledge within industries such as finance, medical care, transportation and manufacturing to machine knowledge with representations understandable by both humans and AI systems. Now, KE, along with big data, deep learning, natural language processing etc., has become one of the core driving forces of AI development.

Key technologies of KE include knowledge representation, knowledge modelling, knowledge acquisition, knowledge storage, knowledge fusion, knowledge calculation, knowledge maintenance, knowledge visualization, etc. In addition, many knowledge service platform products and solutions have been developed to permit KE implementations to be more agile in organizations. The distributed KE systems can be integrated and deployed through knowledge exchange and knowledge maintenance among the systems. The distributed, autonomous agent systems and their collaboration across system of systems can further generate the necessary intelligence and knowledge driven behaviours for collaboration and cooperation.

Resource description framework (RDF),[1] resource description framework schema (RDFS),[2] RDFS-PLUS, ontology web language (OWL),[3] SPARQL protocol and RDF query language (SPARQL)[4] and ontology-related theories and standards[5-7] provide a solid foundation of tools and theories in the aspects of knowledge representation and knowledge modelling. Other related KE standards have been developed.

KE has been successfully applied to many industries including financial fraud identification, remote operation and maintenance of equipment, user profile and product recommendations, research focus tracking and forecasting, smart credit analysis, legal dispute and case prediction based on similar cases, intelligent distribution of news, intelligent computer-aided diagnosis and treatment, etc. Many organizations regard platforms or systems based on KE as important knowledge infrastructures. However, KE vocabularies, basic KE constructional components, KE processes and their relationships are not yet clearly defined. This causes misunderstandings and unnecessary communication and deployment costs amongst the data supplier, fundamental technology supplier, algorithm supplier, system coordinator and other stakeholders of KE systems.

To facilitate collaboration amongst KE stakeholders, KE characteristics and applications can be comprehensively described and categorized. Expected use of the document is to guide the construction of KE systems.

Information technology — Artificial intelligence — Reference architecture of knowledge engineering

1 Scope

This document defines a reference architecture of knowledge engineering (KE) in artificial intelligence (AI). The reference architecture describes KE roles, activities, constructional layers, components and their relationships amongst themselves and other systems from systemic user and functional views. This document also provides a common KE vocabulary by defining KE terms.

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

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 22989 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
architecture
fundamental concepts or properties of an entity in its environment and governing principles for the realization and evolution of this entity and its related life cycle processes

[SOURCE: ISO/IEC/IEEE 42010:2022, 3.2]

3.2
architecture view
information part comprising portion of an architecture description

EXAMPLE An Information or Data View addresses information-relevant concerns framed by an Information viewpoint. It contains as view components, a conceptual data model, a data management model and a data access model and correspondences linking those components together.

[SOURCE: ISO/IEC/IEEE 42010:2022, 3.7]

3.3
data
reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing

Note 1 to entry: Data can be processed by humans or by automatic means.

[SOURCE: ISO/IEC 20546:2019, 3.1.5]

3.4
information
data that are processed, organized and correlated to produce meaning

Note 1 to entry: Information concerns facts, concepts, objects, events, ideas, processes, etc.

[SOURCE: ISO/IEC 20547-3:2020, 3.3]

3.5
knowledge engineering
KE
discipline concerned with acquiring knowledge from domain experts and other knowledge sources and incorporating it into a knowledge base

Note 1 to entry: The term "knowledge engineering" sometimes refers particularly to the art of designing, building, and maintaining knowledge-based systems.

[SOURCE: ISO/IEC 2382:2015, 28.01.07, modified — replaced notes to entry.]

3.6
concept
<terminology> unit of thought differentiated by a unique combination of characteristics

Note 1 to entry: Concepts are not necessarily bound to particular languages. They are, however, influenced by the social or cultural background which often leads to different categorizations of concepts.

[SOURCE: ISO 1087:2019, 3.2.7, modified — replaced "knowledge" with "thought".]

3.7
entity
object of the environment or domain (real-world objects and events, abstract concepts, documents, etc.)

EXAMPLE In the case of a knowledge graph, entity descriptions forming a network and provides context for each other entity interpretation.

3.8
attribute
property of an entity with respect to a defined characteristic

EXAMPLE "Entity X has 5 kg mass" is an attribute, but "having mass" is a characteristic and "5 kg mass" is a property, and neither individually are attributes.

3.9
ontology
collection of terms, relational expressions, and associated natural-language definitions together with one or more formal theories designed to capture the intended interpretations of these definitions

Note 1 to entry: Background materials on the sources, rationale and interpretation of this definition are provided in ISO/IEC 21838-1:2021, Annex B.

[SOURCE: ISO/IEC 21838-1:2021, 3.14]

3.10
schema
formal description of a model

[SOURCE: ISO 19101-1:2014, 4.1.34]

3.11
relation
association amongst entities

[SOURCE: ISO/IEC 15938-5:2003, 3.3.2.29]

3.12
rule
statement in the form of a condition- action sentence that describe the logical inferences that can be drawn from an assertion in a particular form

EXAMPLE A rule can be constructed in the form of "IF-THEN" statements where the IF portion defines a context, and the THEN portion states a provision (which is applicable if the context is true or present).

3.13
structured knowledge
knowledge that are organized based on a pre-defined (applicable) set of rules

3.14
knowledge graph
graph representation of structured knowledge on concepts and relationships between them

Note 1 to entry: A knowledge graph can comprise an ontology and data related to the ontology.

Note 2 to entry: A knowledge graph can be represented as a collection of triples, with each triple (head, tail, relation) denoting the fact that relation exists between head entity and tail entity.

3.15
activity
specified pursuit or set of tasks

[SOURCE: ISO/IEC 22123-1:2023, 3.3.8]

3.16
conceptual model
description of common concepts and their relationships, particularly in order to facilitate exchange of information between parties within a specific domain

[SOURCE: ISO/TS 18864:2017, 3.6, modified — deleted "healthcare".]

3.17
knowledge representation
KR
process or result of encoding knowledge for communication or storage in a knowledge base

Note 1 to entry: As an analogy: data IS-TO code sets IS-TO data engineering AS knowledge IS-TO knowledge representation IS-TO knowledge engineering.

[SOURCE: ISO/IEC 2382:2015, 2123776, modified — replaced "and storing knowledge" with "knowledge for communication or storage"; replaced notes to entry.]

3.18
knowledge modelling
process that establishes and maintains the conceptual model for a knowledge base

3.19
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 via human learning involves a human learner participating in a learning experience. Knowledge acquisition within knowledge engineering typically implies the intervention of a knowledge engineer. Knowledge acquisition is also an important component of machine learning, both with and without human intervention.

[SOURCE: adapted from ISO/IEC 2382:2015, 28.01.09; replace notes]

3.20
knowledge fusion
process that merges, combines and integrates knowledge from different resources into a coherent form

3.21
knowledge storage
process that designs underlying storage methods based on the types of knowledge representation, utilizes hardware and software infrastructure to store, code and make indexes of the knowledge

3.22
knowledge computing
process that obtains new knowledge based on existing knowledge and their relationships

3.23
knowledge exchange
process that transfers, shares and fuses knowledge amongst multiple knowledge bases

3.24
knowledge visualization
process that visually represents knowledge to support human understanding

3.25
safety
freedom from risk which is not tolerable

[SOURCE: ISO/IEC Guide 51:2014, 3.14]

3.26
reliability
property of consistent intended behaviour and results

[SOURCE: ISO/IEC 27000:2018, 3.55]

3.27
availability
property of being accessible and usable on demand by an authorized entity

[SOURCE: ISO/IEC 27000:2018, 3.7]

3.28
accountable
answerable for actions, decisions and performance

[SOURCE: ISO/IEC 38500:2015, 2.2]

3.29
accountability
state of being accountable

[SOURCE: ISO/IEC 38500:2015, 2.3]

3.30
life cycle
evolution of a system, product, service, project or other human-made entity, from conception through retirement

[SOURCE: ISO/IEC/IEEE 15288:2023, 4.1.23]

3.31
data processing
DP
automated data processing
ADP
systematic performance of operations upon data

EXAMPLE Arithmetic or logic operations upon data, merging or sorting of data, assembling or compiling of programs, or operations on text, such as editing, sorting, merging, storing, retrieving, displaying, or printing.

Note 1 to entry: The term data processing is not a synonym for information processing. Information processing includes data communication (e.g. computer networks) and office automation (e.g. satisfying the business needs of an entity), whereas data processing does not include data communication and office automation.

[SOURCE: ISO/IEC 2382:2015, 01.01.06]

3.32
knowledge engineering system
KE system
system that acquires knowledge from domain experts and other knowledge sources and incorporates it into a knowledge base

3.33
knowledge engineering process
KE process
set of activities that acquires knowledge from domain experts and other knowledge sources and incorporates it into a knowledge base

4 Abbreviated terms

AI artificial intelligence

IoT internet of things

KE knowledge engineering

KERA knowledge engineering reference architecture

RDF resource description framework

RDFS resource description framework schema

OWL web ontology language

SPARQL SPARQL protocol and RDF query language

ML machine learning

NLP natural language processing

SHACL shapes constraint language

SKOS simple knowledge organization system

URL uniform resource locator

URI uniform resource identifier

5 Knowledge engineering system-of-interest

5.1 General

KE attempts to emulate the judgment and behaviour of a human expert in a given field. With the growing popularity of knowledge-based systems in recent years, there is a need for a systematic approach for building such systems, similar to methodologies used in software engineering. KE involves acquiring knowledge from domain experts, available data and other knowledge sources and incorporating it into a knowledge base. In addition, the rapid development of big data, cloud computing, natural language processing, computer vison among others have improved the capability of collecting and processing data, which also encourages enterprises and people to put more effort into knowledge-intensive applications based on the discipline of KE. KE began in the late 1980s and has a substantial history, including: knowledge interchange format,[21] knowledge query and manipulation language [22] (Knowledge Sharing Effort, early 1990s), knowledge acquisition data system (KADS) or COMMON KADS (mid 1990s), Cyc [23] (on-going).

5.2 Important elements of knowledge engineering

Important elements of KE involve concepts of:

— deployment;

— infrastructure;

— system;

— system operation restriction;

— demand;

— data;

— knowledge;

— construction;

— knowledge operating.

Figure 1 shows how these element concepts can be structured, decomposed and inter-related:

— The AI system associated with the KE process or KE system is supported through a construction process, which is based on data and information, a knowledge operating process and fundamental infrastructures under system operating restrictions.

— System operating restrictions are extracted from the KE system, such as application scenarios, performance requirements.

— After the KE system is developed, the deployment process is triggered, including integration, deployment and promotion of the KE system.

— During construction and knowledge operating, knowledge is acquired through extracted information from original data, including structured data, semi-structured data and unstructured data.

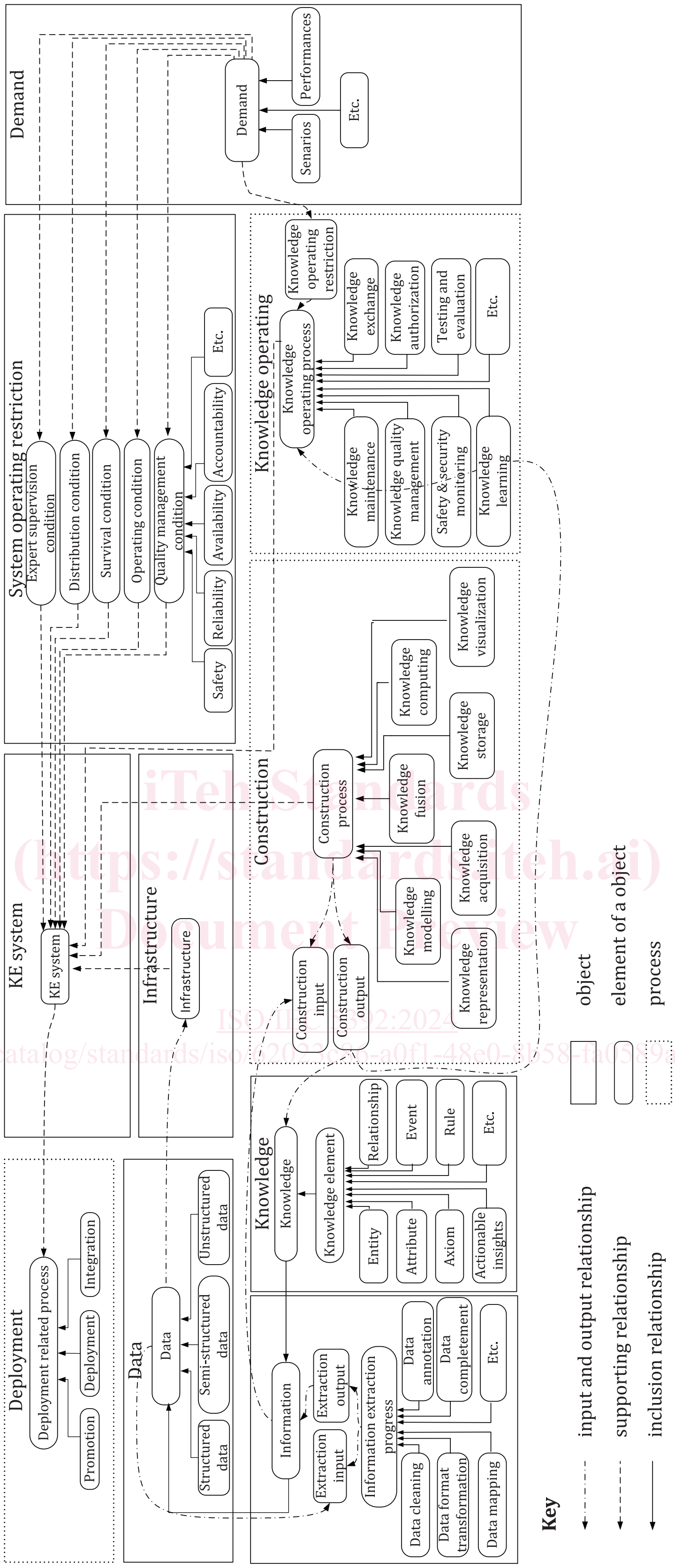

Figure 1 — Important elements of knowledge engineering

5.3 Relationship between KE and AI systems

According to the AI system functional view given in ISO/IEC 22989:2022, reproduced on the left of Figure 2, AI systems leverage existing information, or learning from the past, to build a model that approximates the behaviour of an environment to make recommendations on future behaviours of that environment. Through training data and continuous learning with the help of human in the loop, the machine learning model can be curated and regularly evaluated, updated and approved. The relationship of KE with respect to AI systems is depicted on the right of Figure 2. KE provides the further capability to acquire data, process the data to extract information, and store and exchange data, information or knowledge.

The AI system with KE can acquire knowledge directly from the information extracted from the data and further construct the knowledge base. During the process of data processing, the knowledge in the knowledge base can be applied to inspect and assist the process. At the same time, the knowledge base can maintain, update and verify itself as follows:

— by computing and reasoning new knowledge based on existing knowledge;

— through revisions and updates approved through the curation and synthesis into existing knowledge by an engineer;

— through discovering new knowledge during the data and information processing.

In addition, the knowledge in the knowledge base can be used to:

— govern the input data, such as transferring the data format, cleaning the error in data, supplementing relations among data;

— supervise and explain the learning process or the learning result;

— participate in the learning process as training set.

In the constructed knowledge base of the AI system, there are two types of knowledge.

— Knowledge of methods: for example, machine learning models and other models driven by approaches that include data driven and subject expertise captured from an expert.

— Knowledge of contents: the knowledge about subject area in the form of concepts, relationships, entities drawn from texts, videos, and so on, which is acquired from the input data or information and from first principles like physics-based models or biology-based models. The knowledge of contents can be used to supervise and explain the learning process and results as well as to assist to improve the quality of input data and information. At the same time, the knowledge of contents can be used to improve understanding, human knowledge and insight and can be transferred to or have an impact on other AI systems.