INTERNATIONAL STANDARD

ISO/IEC 21838-1

First edition
2021-08

Information technology — Top-level ontologies (TLO) —

Part 1:
Requirements

Reference number
ISO/IEC 21838-1:2021(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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see patents.iec.ch).

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 32, *Data management and interchange*.

A list of all parts in the ISO/IEC 21838 series can be found on the ISO website.

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

This document was developed in response to the demand from many quarters for ontology-based solutions to the problem of semantic interoperability across networks of information systems. The demand arises particularly from large organizations and consortia of organizations in areas such as bioinformatics, healthcare, the manufacturing industry and military and government administration, where independently created information systems need to exchange data in such a way that meaning is preserved.

An ontology is on the one hand an artefact for human use, built out of terms and relations expressed using natural language. On the other hand, it is an artefact for use by computers, which requires that these terms and relations are captured in a formal language that is machine readable and has well-defined (typically, model-theoretic) semantics. Multiple languages have been developed for the purposes of ontology formalization, of which Common Logic (CL) and the Web Ontology Language (OWL) – specifically OWL 2 with direct semantics – are normatively referenced in this document.

An ontology can help to achieve sharing of meaning because its terms are associated with formal definitions specifying their meanings in a way that can be processed computationally. If an ontology can be shared across participating organizations, then data can be exchanged in such a way that meaning is preserved if the data can be associated with corresponding shared ontology terms.

CL and OWL 2 serve different ends. CL is a logical framework with the full expressivity of first-order logic (FOL), the unifying framework for all semantic web applications. Formalization in a language with the expressivity of FOL is required for the purposes of this document since weaker expressivity would not allow the ontology to capture in a formal way the implications of axioms in areas such as mereology and theories of location and change.

Formalization in a language like OWL 2 is needed, even though it is less expressive than CL, since it is decidable and this means that it can be used effectively by computer systems for purposes of logical reasoning and ontology quality assurance.

Where heterogeneous bodies of data need to be exchanged or manipulated, some have adopted approaches that involve the creation of a suite of ontologies incorporating a distinction of levels, with a single very general ontology at the top, governing one or more specific ontology modules at lower levels (Annex A provides examples). This document addresses the need that arises for those communities that have adopted such multi-level approaches. Specifically, its purpose is to specify what is required of a top-level ontology if it is to serve the needs of those building or re-engineering ontologies or other legacy systems at lower levels in a way that will support semantic interoperability among them.

To be fit for purpose, a top-level ontology needs to have appropriate content that is well documented and be available in machine-readable forms providing support for computational reasoning. This document specifies these requirements in terms of coverage, documentation and representation.

Information technology — Top-level ontologies (TLO) —

Part 1: Requirements

1 Scope

This document specifies required characteristics of a domain-neutral top-level ontology (TLO) that can be used in tandem with domain ontologies at lower levels to support data exchange, retrieval, discovery, integration and analysis.

If an ontology is to provide the overarching ontology content that will promote interoperability of domain ontologies and thereby support the design and use of purpose-built ontology suites, then it needs to satisfy certain requirements. This document specifies these requirements. It also supports a variety of other goals related to the achievement of semantic interoperability, for example, as concerns legacy ontologies developed using heterogeneous upper-level categories, where a coherently designed TLO can provide a target for coordinated re-engineering.

This document specifies the characteristics an ontology needs to possess to support the goals of exchange, retrieval, discovery, integration and analysis of data by computer systems.

The following are within the scope of this document

— Specification of the requirements an ontology needs to satisfy if it is to serve as a top-level hub ontology.

— Specification of the relations between a top-level ontology and domain ontologies.

— Specification of the role played by the terms in a top-level ontology in the formulation of definitions and axioms in ontologies at lower levels.

The following are outside the scope of this document:

— Specification of ontology languages, including the languages OWL 2 and CL, used in ontology development with standard model-theoretic semantics.

— Specification of methods for reasoning with ontologies.

— Specification of translators between notations of ontologies developed in different ontology languages.

— Specification of rules governing the use of IRIs as permanent identifiers for ontology terms.

— Specification of the principles of ontology maintenance and versioning.

— Specification of how ontologies can be used in the tagging or annotation of data.

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 24707, *Information technology — Common Logic (CL) — A framework for a family of logic-based languages*

WORLD WIDE WEB CONSORTIUM *W3C Recommendation — OWL 2 Web Ontology Language Document Overview* (Second Edition), https://www.w3.org/TR/2012/REC-owl2-overview-20121211/

WORLD WIDE WEB CONSORTIUM *W3C Recommendation — OWL 2 Web Ontology Language Direct Semantics*, https://www.w3.org/TR/owl2-direct-semantics/

WORLD WIDE WEB CONSORTIUM *W3C Recommendation — OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax* (Second Edition), http://www.w3.org/TR/2012/REC-owl2-syntax-20121211/

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

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

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

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

NOTE The following terms and definitions are not intended as a substitute for existing technical vocabularies used in ontology development and maintenance, for example, as defined by the W3C. To reduce the possibility of confusion, expressions used in describing a W3C recommended usage are capitalized.

3.1
entity
object
item that is perceivable or conceivable

Note 1 to entry: The terms ‘entity’ and ‘object’ are catch-all terms analogous to ‘something’. In terminology circles ‘object’ is commonly used in this way. In ontology circles, ‘entity’ and ‘thing’ are commonly used. See B.3.3.

[SOURCE: ISO 1087-1:2000]

3.2
class
general *entity* (3.1)

Note 1 to entry: In some ontology communities, all general entities are referred to as classes. In other ontology communities, a distinction is drawn between classes as the extensions of general entities (for example, as sets of instances) and the general entities themselves, sometimes referred to as ‘types’, ‘kinds’, or ‘universals’. The expression ‘class or type’ is used in this document in order to remain neutral regarding these different usages.

3.3
particular
individual *entity* (3.1)

Note 1 to entry: In contrast to classes or types, particulars are not exemplified or instantiated by further entities.

3.4
relation
way in which *entities* (3.1) are related

Note 1 to entry: Relations can hold between particulars (this leg *is part of* this lion); or between classes or types (mammal *is a subclass of* organism); or between particulars and classes or types (this lion *is an instance of* mammal). On some views, identity is treated as a relation connecting one entity to itself.

Note 2 to entry: On the difference between ‘relation’ and ‘relational expression’ see 3.6, Note 1 to entry.

Note 3 to entry: ‘Relation’ is a primitive term. See 4.1.1, NOTE 1.

3.5
expression
word or group of words or corresponding symbols that can be used in making an assertion

Note 1 to entry: Expressions are divided into natural language expressions and expressions in a formal language.

3.6
relational expression
expression (3.5) used to assert that a *relation* (3.4) obtains

EXAMPLE 'is a' (also known as 'subtype' or 'subclass'), 'part of', 'member of', 'instantiates' 'later than', 'brother of', 'temperature of'.

Note 1 to entry: The term 'relational expression' is introduced in order to remove any confusion that can arise if a person uses 'relation' to refer to the real-world link or bond between entities (as in 3.4), while another person uses 'relation' to refer to the linguistic representation of this real-world link or bond.

Note 2 to entry: In OWL 2, relational expressions are referred to as Properties. 'Expression' is used to connote logical composition: a Class Name in OWL 2 is logically simple, a Class Expression is logically complex. In FOL, 'n-ary predicate' is often used as a synonym of 'relational expression'.

3.7
term
expression (3.5) that refers to some *class* (3.2) or to some *particular* (3.3)

Note 1 to entry: An ontology will typically contain a unique 'preferred term' for the entities within its coverage domain. Preferred terms may then be supplemented with other terms recognized by the ontology as synonyms of the preferred terms.

3.8
definition
concise statement of the meaning of an *expression* (3.5)

Note 1 to entry: For the purposes of this document, definitions can be of two sorts: (1) those formulated using a natural language such as English, supplemented where necessary by technical terms or codes used in some specialist domain; (2) those formulated using a computer-interpretable language such as OWL 2 or CL.

3.9
axiom
statement that is taken to be true, to serve as a premise for further reasoning

Note 1 to entry: Axioms may be formulated as natural language sentences or as formulae in a formal language. In the OWL community, 'Axiom' is used to refer to statements that say what is true in the domain that are 'basic' in the sense that they are not inferred from other statements.

3.10
formal language
language that is machine readable and has well-defined semantics

Note 1 to entry: Well-defined semantics will typically be model-theoretic semantics.

3.11
formal theory
collection of *definitions* (3.8) and *axioms* (3.9) expressed in a *formal language* (3.10)

Note 1 to entry: In some formal theories, definitions are expressed by means of axioms.

3.12
axiomatization
result of expressing a body of knowledge or information as a *formal theory* (3.11)

3.13
logical interpretability
ability to derive each and every *axiom* (3.9) of one *formal theory* (3.11) from another

Note 1 to entry: One formal theory is logically interpretable in a second formal theory if the language of the first can be translated into the language of the second so that the translation of every axiom in the first is derivable from the second.

3.14
ontology
collection of *terms* (3.7), *relational expressions* (3.6) and associated natural-language *definitions* (3.8) together with one or more *formal theories* (3.11) 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 Annex B.

3.15
signature
set of non-logical symbols of a *formal language* (3.10) or *formal theory* (3.11)

Note 1 to entry: The signature of an ontology consists of a set of *terms* (3.7) and *relational expressions* (3.6).

3.16
knowledge base
combination of an *ontology* (3.14) with a collection of data which *terms* (3.7) in the ontology have been used to describe, classify or connect.

3.17
domain
collection of *entities* (3.1) of interest to a certain community or discipline

EXAMPLE The domain of agriculture, the domain of cell biology, the domain of aircraft maintenance, the domain of philately.

Note 1 to entry: 'Entities of interest' can include both particulars and classes or types. The definition is to be interpreted as meaning that a domain is a collection of entities that is narrow in scope. Thus, there is no universal domain, to which everything would belong. Compare with ISO/IEC 2382[21], which defines 'domain model' in the context of artificial intelligence as: model of a specific field of knowledge or expertise.

3.18
domain ontology
ontology (3.14) whose *terms* (3.7) represent *classes* (3.2) or *types* and, optionally, certain *particulars* (3.3) (called 'distinguished individuals') in some *domain* (3.17)

3.19
category
general *class* (3.2) or type that is shared across many different *domains* (3.17) and is represented by a domain-neutral *term* (3.7)

EXAMPLE Process, attribute, event, region, information entity.

3.20
top-level ontology
TLO
ontology (3.14) that is created to represent the *categories* (3.19) that are shared across a maximally broad range of *domains* (3.17)

Note 1 to entry: Top-level ontologies are 'reference ontologies' in the sense of ISO/IEC 19763-3[5], A top-level ontology is sometimes referred to as a 'formal ontology', 'foundational ontology', 'upper level ontology', or 'domain-neutral ontology'.

3.21
ontology suite
collection of *ontologies* (3.14) developed in such a way as to be mutually consistent and non-redundant

Note 1 to entry: See Annex A.

3.22
ontology reuse
importing an *ontology* (3.14), or part of an ontology, into a second ontology in such a way as to preserve the meaning of the imported content

EXAMPLE Terms from a tool ontology are reused in a power tool ontology; the latter is a specialization of the former.

Note 1 to entry: Terms from the existing ontology will typically be reused in the new ontology and appear together with the newly created terms.

3.23
ontology conformance
relation (3.4) between two *ontologies* (3.14) when one consistently extends the other

EXAMPLE A power tool ontology stands in the *relation* of *ontological conformance* to a tool ontology if the former is a consistent ontology that results from adding new content (terms, definitions, axioms) to the latter.

Note 1 to entry: 'Extension' means semantically that any element in a model of the extending ontology which satisfies the conditions for being an instance of a class in the starting ontology must be an instance of that class in the extending ontology.

Note 2 to entry: This is a narrowly defined usage of 'conformance' that is intended to be used only in contexts in which relations between ontologies are at issue. Where conformance in the sense of *fulfilment of a requirement or satisfaction of a criterion* is intended in this document, the term 'conformity' is used.

4 Requirements for a top-level ontology

4.1 TLO as textual artefact

4.1.1 Overview

A TLO shall include a textual artefact represented by a natural language document providing: (1) a list of domain-neutral terms and relational expressions, incorporating identification of primitive terms, and (2) definitions of the meanings of the terms and relational expressions listed. Natural-language definitions may incorporate semi-formal elements if these are needed for readability.

NOTE 1 In the case of primitive terms, definitions can take the form of elucidations of meaning supplemented by examples of use.

EXAMPLE An example of a definition with semi-formal elements is:

transitivity =def. relation R is transitive if whenever *a* stands in R to *b* and *b* stands in R to *c* it follows that *a* stands in R to *c*.

Given the nature of a TLO, a portion of its terms and relational expressions will be so basic in their meaning that there will be no logically simpler, and thus more easily intelligible, expressions on the basis of which they can be defined in a non-circular way. Ontology terms and relational expressions for which this is the case are called 'primitives', and they have definitions in the sense of 3.8, but these are circular or are mere paraphrases.

A TLO shall specify which of its terms and relational expressions are primitive in this sense. For all other terms and relational expressions in the TLO, definitions shall be provided which satisfy the conditions that:

a) they are non-circular;

b) they form a consistent set;

c) they are concise.

NOTE 2 Concise signifies that the definition contains no redundant elements (for example, lists of examples, explanations of usage, and so on).

These requirements apply both to the natural language definitions and also to the definitions provided in the OWL 2 and CL axiomatizations referenced in 4.2 and 4.3.

Non-circularity excludes not only immediate circularity (where the defined term or a term with equivalent meaning is used in the definition) but also mediated circularity (for example, where a term is used in the definition of a second term, which is itself used in the definition of the first term). To ensure non-circularity it is recommended that definitions are formulated as statements of singly necessary and jointly sufficient conditions for the correct application of the defined term.

EXAMPLE Triangle = def. closed figure that lies in a plane and consists of exactly three straight lines.

Consistency of the collection of natural language definitions is shown through the development of an axiomatization that is proven consistent, as described in 4.2 and 4.3.

NOTE 3 Consistency, non-circularity and conciseness of definitions are features that distinguish ontologies from traditional dictionaries and other lexical resources.

4.1.2 Relations between textual artefact and axiomatizations of the TLO

The terms and relational expressions in the textual artefact shall be converted into symbols in the axiomatizations. These symbols together form the signature of the resultant logical theory. They may incorporate textual strings.

EXAMPLE The text string 'is a' is converted into the symbol 'is_a'.

Terms and relational expressions in the textual artefact should have counterparts in the OWL 2 axiomatization wherever this is feasible, given the expressivity of OWL.

Each definition in the textual artefact whose content is expressible in OWL 2 shall correspond in the OWL 2 axiomatization to a group of one or more axioms with a corresponding logical content.

All terms in the textual artefact shall correspond to terms in the CL axiomatization.

All definitions of non-primitive terms in the textual artefact shall correspond to axioms in the CL formalization.

4.2 Axiomatization in the Web Ontology Language (OWL 2 with direct semantics)

4.2.1 General

The TLO shall be made available via at least one machine-readable axiomatization in OWL 2 with the direct semantics or in some description logic that is designated by W3C as a successor of OWL 2. The signature of the OWL axiomatization shall be identical, modulo the conversion from strings into symbols and modulo the conversion of ternary into binary relational expressions, to the set of natural language terms and relational expressions of the TLO as specified under 4.1. The axioms should represent the content of the natural language definitions described in 4.1 to the extent that this is possible given the expressivity of OWL 2. The axiomatization shall satisfy the conformity criteria in *W3C Recommendation — OWL 2 Web Ontology Language Direct Semantics*. The axiomatization shall be

