International Standard

ISO/IEC 11179-34

First edition
2024-05

Information technology — Metadata registries (MDR) —

Part 34: Metamodel for computable data registration

Technologies de l'information — Registres de métadonnées (RM) —

Partie 34: Métamodèle pour l'enregistrement des données calculables

Reference number
ISO/IEC 11179-34:2024(en)

© ISO/IEC 2024



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2024

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ... **iv**
Introduction ... **v**
1 Scope ... **1**
2 Normative references ... **1**
3 Terms and definitions ... **1**
4 Abbreviated terms ... **2**
5 Conformance ... **2**
5.1 Overview of conformance ... 2
5.2 Degree of conformance ... 2
5.2.1 General ... 2
5.2.2 Strictly conforming implementations ... 2
5.2.3 Conforming implementations ... 2
5.3 Conformance by feature ... 3
5.4 Registry conformance ... 3
5.4.1 Standard registry profiles ... 3
5.4.2 Conformance labels ... 3
5.5 Implementation conformance statement (ICS) ... 3
5.6 Obligation ... 3
6 Relationship to ISO/IEC 11179-3 ... **4**
6.1 Metamodel for a metadata registry ... 4
6.2 Specification of the metamodel ... 4
6.3 Use of UML class diagrams and textual description ... 4
6.4 Package dependencies ... 5
7 Computable_Data package ... **5**
7.1 Overview of the Computable_Data package ... 5
7.2 Computable_Data metamodel region ... 6
7.2.1 Overview of the Computable_Data metamodel region ... 6
7.2.2 Classes in the Computable_Data metamodel region ... 8
7.2.3 Associations in the Computable_Data metamodel region ... 22
7.2.4 Datatypes in the Computable_Data metamodel region ... 24
Annex A (informative) **Mapping between IEEE 2791-2020 and this document** ... **26**
Annex B (normative) **Consolidated Class Hierarchy** ... **31**
Annex C (informative) **Examples of computable data registration** ... **32**
Bibliography ... **41**

© ISO/IEC 2024 – All rights reserved

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 32, *Data management and interchange*.

This document is part of the fourth modularization of ISO/IEC 11179. This document brings into ISO/IEC 11179 the ability to register information about computable data.

A list of all parts in the ISO/IEC 11179 series can be found on the ISO and IEC websites.

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.

© ISO/IEC 2024 – All rights reserved

Introduction

Significant scientific discoveries are increasingly achieved through complex and distributed computations and data analyses. These computations and analyses always involve processing files through a series of steps and transformations, usually called a pipeline or a workflow.

Data pipelines typically use multiple pieces of software, each of which typically has multiple versions available, multiple input parameters, multiple outputs, and possibly platform-specific configurations. As with experimental parameters in a laboratory protocol, small changes in computational parameters can have a large impact on the scientific validity of the results.

To reproduce and verify scientific discoveries, details of pipelines need to be documented and shared, including the protocol, procedures, or metadata associated with them. The more analysis steps and the more complicated a pipeline, the greater the need for a standardized mechanism of communication. A detailed communication helps ensure responsibility, reproducibility, and the ability to verify protocol, track provenance information, and promote interoperability.

This document is inspired by the IEEE 2791-2020[1], which aims to improve communication of bioinformatics protocols and data to facilitate bioinformatics workflow related exchange and communication between regulatory agencies, pharmaceutical companies, bioinformatics platform providers and researchers. Although IEEE 2791-2020 has a bioinformatics background and application areas, the concepts and methods it expressed are applicable to a broader field of scientific research. A mapping table is included in Annex A, showing the relationship between domains and fields in IEEE 2791-2020 and classes and attributes in this document.

ISO/IEC 11179-3 specifies the structure of a Metadata Registry (MDR) and provides a metamodel for registry common facilities. That metamodel is intended to be extended by other parts of ISO/IEC 11179 for specific purposes.

This document provides a specification of the extensions to the registry metamodel specified in ISO/IEC 11179-3 to enable the registration of metadata about computable data. Registration of metadata about computable data are like a manifest describing all details related to input files, output files, and the pipeline used to process these files. The intent is to facilitate efficient communication and interoperability among different platforms, industries, scientists, and regulators and to improve reproducibility and replicability.

In Clauses 6 and 7, this document uses **bold** font to highlight terms which represent metadata objects specified by the metamodel.

EXAMPLE **Computable_Data** (see 7.2.2.1) is a class each instance of which models computable data.

© ISO/IEC 2024 – All rights reserved

Information technology — Metadata registries (MDR) —

Part 34:
Metamodel for computable data registration

1 Scope

This document provides a specification for an extension to a metadata registry (MDR), as specified in ISO/IEC 11179-3, in which metadata that describe computable data can be registered.

The specification in this document, together with the relevant clauses of the specification in ISO/IEC 11179-3, provides the ability to record metadata about computable 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 11179-3, *Information technology — Metadata registries (MDR) — Part 3: Metamodel for registry common facilities*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 11179-3 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
computable data
data that are computed for the purpose of making the data reproducible

3.2
pipeline
full set of logically connected steps needed to transform input data into a result, for which one or more outputs of one data transformation step can be the input of one or more subsequent steps

3.3
computation step
individual tool (or a well defined and reusable script) used in a *pipeline* (3.2) that can be executed sequentially or in parallel with other computation steps

3.4
computation execution environment
information of the environment needed to execute a *computation step* (3.3) to the extent to which consistent results are obtained when produced repeatedly, including platform, deployment, software configuration and running applications, etc.

© ISO/IEC 2024 – All rights reserved