TECHNICAL REPORT

ISO/IEC TR 20943-1

First edition
2003-08-01

Information technology — Procedures for achieving metadata registry (MDR) content consistency —

Part 1: Data elements

Technologies de l'information — Procédures en vue d'obtenir la cohérence du contenu d'un registre de metadonnées (RM) —

Partie 1: Éléments de données

Reference number
ISO/IEC TR 20943-1:2003(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

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. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

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.

ISO/IEC TR 20943-1:2003, which is a Technical Report of type 3, was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 32, *Data management and interchange*.

ISO/IEC 20943 consists of the following parts, under the general title *Information technology — Procedures for achieving metadata registry (MDR) content consistency*:

Note: Parts 2 and 3 are currently under development.

— *Part 1: Data elements*

— *Part 2: XML structured data*

— *Part 3: Value domains*

Introduction

The exchange of metadata between metadata registries based on International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 11179 *Information technology — metadata registries (MDR)*, depends not only on registry software that conforms to the standard, but also on metadata contents that are comparable between registries. While the standard has provisions for data specification and registration, there are pragmatic issues pertaining to populating the registries with content. Based on the experiences of organizations that are implementing the standard, a Technical Report to explore content issues will help current and future users.

Metadata registries can be used to register data elements, value domains, and associated attributes for many kinds of organizational data resource collections. Metadata registries can store information on data elements used on forms, represented in enterprise data models, contained in EDI message sets, and described in documents and standards, as well as those data elements that are part of computer system applications. Some organizations use the registry to record essential facts about how data elements are used in existing applications, while other organizations use the registry as a repository of standard data elements to be used as models for data elements in application development. ISO/IEC 11179-6 specifically addresses the development and population of metadata registries.

ISO/IEC 11179-3 models a data element and its associated data element concept. Conceptualization and articulation of rules and relationships are needed in the creation of data element concepts, data elements, and value domains. Explication of the various possible levels of data elements and data element concepts and their relationships would greatly assist in the creation of shareable, well-formed data. Relationship and inheritance from the most generalized data element to the most specialized application data element need to be specified. Reuse of data value domains should be enabled and regularized.

While metadata registries can be used for storing information about a variety of metadata entities, this report addresses only data elements and associated metadata items. The goal of this paper is to ensure that there is a common understanding of the content of the data element attributes so that metadata can be shared between registries, despite their differences.

This Technical Report is based ISO/IEC 11179-3 of the six-part ISO/IEC standard that describes the organization of a registry for managing the semantics of data. The standard specifies the structure of a registry in the form of a conceptual model. The conceptual model is not intended to be a logical or physical data model for a computer system.

Information technology — Procedures for achieving metadata registry (MDR) content consistency —

Part 1: Data elements

1 Scope

1.1 Background

An ISO/IEC 11179-based metadata registry (MDR) (hereafter referred to as a "registry") is a tool for the management of shareable data; a comprehensive, authoritative source of reference information about data. It supports the standardization and harmonization processes by recording and disseminating data standards, which facilitates data sharing among organizations and users. It provides links to documents that refer to data elements and to information systems where data elements are used. When used in conjunction with an information database, the registry enables users to better understand the information obtained.

A registry does not contain data itself. It contains the metadata that is necessary to clearly describe, inventory, analyse, and classify data. It provides an understanding of the meaning, representation, and identification of units of data. The standard identifies the information elements that need to be available for determining the meaning of a data element (DE) to be shared between systems.

1.2 Purpose

The purpose of ISO/IEC TR 20943-1:2003 is to describe a set of procedures for the consistent registration of data elements and their attributes in a registry. ISO/IEC TR 20943-1:2003 is not a data entry manual, but a user's guide for conceptualizing a data element and its associated metadata items for the purpose of consistently establishing good quality data elements. An organization may adapt and/or add to these procedures as necessary.

1.3 Scope

The scope of ISO/IEC TR 20943-1:2003 is limited to the associated items of a data element: the data element identifier, names and definitions in particular contexts, and examples; data element concept; conceptual domain with its value meanings; and value domain with its permissible values.

1.4 Registration approach — data elements and value domains

There is a choice when registering code sets and other value domains in an ISO/IEC 11179 metadata registry. Some Registration Authorities treat these sets as value domains, and others treat them as data elements. For the purposes of ISO/IEC TR 20943-1:2003, the choice will always be to treat the sets as data elements unless explicitly stated. This choice is made to help illustrate the way to register many different kinds of data elements, including examples for registering standard code sets as data elements.

2 Normative references

The following referenced documents are indispensable for the application 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 11179-1:—[1)], *Information technology — Metadata registries (MDR) — Part 1: Framework for the specification and standardization of data elements*

ISO/IEC 11179-2:—[1)], *Information technology — Metadata registries (MDR) — Part 2: Classification for data elements*

ISO/IEC 11179-3:2003, *Information technology — Metadata registries (MDR) — Part 3: Registry metamodel and basic attributes*

ISO/IEC 11179-4:—[1)], *Information technology — Metadata registries (MDR) — Part 4: Rules and guidelines for the formulation of data definitions*

ISO/IEC 11179-5:—[1)], *Information technology — Metadata registries (MDR) — Part 5: Naming and identification principles for data elements*

ISO/IEC 11179-6:—[1)], *Information technology — Metadata registries (MDR) — Part 6: Registration of data elements*

ISO/IEC TR 15452:2000, *Information technology — Specification of data value domains*

Standards from which examples have been drawn to be used in this document are listed in the Bibliography.

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 11179 and ISO/IEC TR 15452 apply.

4 Data element abstraction

This clause presents a conceptual framework for structuring data elements and data element contents in a registry. Data elements are ideally the result of a process of development involving several types of abstraction, producing a series of "layers" related to each other by the method of abstraction used to produce one from the other. Layers may progress from the more general to the more specific.

Depending on the type of abstraction, relationships among the members of each layer and between layers are meaningful in terms of defining the structure of the registry contents. This provides a means of comparison of the contents of different registries and of searching within a registry. In addition to the data element definition and other attributes, comparing the type and level of abstraction by which the data element was derived can ensure that content can be shared among registries.

One could use layers to structure development of a system, for instance, with the highest layers of definition contained in a business view, and development progressing to the implemented system layer. The number and granularity of layers are driven by user requirements. This clause will describe two ways to derive layers, neither of which are intended to be mandatory for any particular implementation, and will present examples of the types of abstraction most useful to registry implementations.

1) To be published.

