INTERNATIONAL STANDARD

ISO 16642

Second edition
2017-11

Computer applications in terminology — Terminological markup framework

Applications informatiques en terminologie — Plate-forme pour le balisage de terminologies informatisées

This document is a preview generated by EVS

Reference number
ISO 16642:2017(E)

© ISO 2017

This document is a preview generated by EVS

COPYRIGHT PROTECTED DOCUMENT

© ISO 2017, Published in Switzerland

All rights reserved. Unless otherwise specified, 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
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

This document is a preview generated by EVS

© ISO 2017 – All rights reserved

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO documents 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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO 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).

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

For an explanation on 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 the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 3, *Computer applications for terminology*.

This second edition cancels and replaces the first edition (ISO 16642:2003), which has been technically revised.

The main changes compared to the previous version are as follows:

— The following formats are no longer actively used. Consequently, references to these formats have been removed (including Annex A, Annex B, and Annex C):

  — Martif with specified constraints (MSC);

  — Geneter;

  — Data category interchange format (DCIF);

  — Generic mapping tool (GMT).

— With the removal of Annex B and Annex C, this document no longer includes any comprehensive code examples of a TML. Examples of TMLs are now available in ISO 30042, TermBase eXchange, and also at the following Web site: www.tbxinfo.net.

— References to the former ISO/TC 37 Data Category Registry or ISOcat have been changed from normative to informative. In addition, the name has changed to DatCatInfo, now as an example of data category repositories.

— References to ISO 12620:1999 and ISO 12620:2009 have been removed. These previous standards have been withdrawn.

— The TypedValuedElement style has been added.

— Examples have been updated to reflect ISO 30042:2008 (TBX). TBX-Basic is mentioned as a TML.

This document is a preview generated by EVS

 © ISO 2017 – All rights reserved

— Some of the examples and tables have been moved to appropriate sections.

— As a consequence of the aforementioned changes, some historical, didactic, or duplicate information has been removed to adhere more closely to ISO editorial standards.

This document is a preview generated by EVS

© ISO 2017 – All rights reserved

Introduction

Terminological data are collected, managed and stored in a wide variety of systems, typically various kinds of database management systems, ranging from personal computer applications for individual users to large terminological database systems operated by major companies and governmental agencies. Terminology databases are comprised of various types of information, called data categories, and can adopt different structural models. However, terminological data often need to be shared and reused in a number of applications, and this sharing is facilitated when the data adheres to a common model. To facilitate co-operation and to prevent duplicate work, it is important to develop standards and guidelines for creating and using terminological data collections (TDCs) as well as for sharing and exchanging data.

This document presents a modular approach for analysing existing TDCs and designing new ones. It also provides a framework for defining terminological markup languages (TMLs) that are interoperable.

This document makes reference to DatCatInfo, an example of an available data category repository. DatCatInfo is an online database of information about the types of data that can be included in terminological data collections and other language resources. It is available at www.datcatinfo.net.

This document is a preview generated by EVS

© ISO 2017 – All rights reserved

Computer applications in terminology — Terminological markup framework

This document is a preview generated by EVS

1 Scope

This document specifies a framework for representing data recorded in terminological data collections (TDCs). This framework includes a metamodel and methods for describing specific terminological markup languages (TMLs) expressed in XML. The mechanisms for implementing constraints in a TML are defined, but not the specific constraints for individual TMLs.

This document is designed to support the development and use of computer applications for terminological data and the exchange of such data between different applications. This document also defines the conditions that allow the data expressed in one TML to be mapped onto another TML.

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 704, *Terminology work — Principles and methods*

ISO 1087-1, *Terminology work — Vocabulary — Part 1: Theory and application*

ISO 3166-1, *Codes for the representation of names of countries and their subdivisions — Part 1: Country codes*

ISO 26162, *Systems to manage terminology, knowledge and content — Design, implementation and maintenance of terminology management systems*

ISO 30042:2008, *Systems to manage terminology, knowledge and content — TermBase eXchange (TBX)*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 1087-1 and the following apply.

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

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

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

3.1
basic information unit
information unit (3.12) attached to a *component* (3.3) of the metamodel and that can be expressed by means of a single *data category* (3.6)

3.2
complementary information
CI
information supplementary to that described in *terminological entries* (3.22) and shared across the *terminological data collection* (3.21)

Note 1 to entry: Domain hierarchies, institution descriptions, bibliographic references and references to text corpora are typical examples of complementary information.

© ISO 2017 – All rights reserved