



**SLOVENSKI STANDARD**  
**oSIST ISO/DIS 12620-2:2021**

**01-marec-2021**

---

**Upravljanje terminoloških virov - Podatkovne kategorije - 2. del: Repozitoriji**

Management of terminology resources -- Data categories - Part 2: Repositories

Titre manque

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

**Ta slovenski standard je istoveten z: ISO/DIS 12620-2**

[oSIST ISO/DIS 12620-2:2021](https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021)

<https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021>

**ICS:**

01.020	Terminologija (načela in koordinacija)	Terminology (principles and coordination)
35.240.30	Uporabniške rešitve IT v informatiki, dokumentiranju in založništvu	IT applications in information, documentation and publishing

**oSIST ISO/DIS 12620-2:2021**

**en**

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[oSIST ISO/DIS 12620-2:2021](https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021)

<https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021>

# DRAFT INTERNATIONAL STANDARD

## ISO/DIS 12620-2

ISO/TC 37/SC 3

Secretariat: DIN

Voting begins on:  
2021-01-20

Voting terminates on:  
2021-04-14

---

---

## Management of terminology resources — Data categories —

### Part 2: Repositories

ICS: 01.020; 35.240.30

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[oSIST ISO/DIS 12620-2:2021](https://standards.iteh.ai/catalog/standards/sist/b138ffac-30fb-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021)

<https://standards.iteh.ai/catalog/standards/sist/b138ffac-30fb-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021>

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENT AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

This document is circulated as received from the committee secretariat.



Reference number  
ISO/DIS 12620-2:2021(E)

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[oSIST ISO/DIS 12620-2:2021  
https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-  
8c6b188fc2f3/osist-iso-dis-12620-2-2021](https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021)



### **COPYRIGHT PROTECTED DOCUMENT**

© ISO 2021

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](mailto:copyright@iso.org)  
Website: [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

<b>Foreword</b> .....	<b>iv</b>
<b>Introduction</b> .....	<b>v</b>
<b>1 Scope</b> .....	<b>1</b>
<b>2 Normative references</b> .....	<b>1</b>
<b>3 Terms and definitions</b> .....	<b>1</b>
<b>4 Data category repositories</b> .....	<b>2</b>
4.1 General interoperability.....	2
4.2 Requirements for a DCR.....	3
4.3 Data category profiles.....	3
4.4 Data category selections.....	4
<b>5 Roles with governance responsibilities</b> .....	<b>4</b>
5.1 Community of practice.....	4
5.1.1 Description.....	4
5.1.2 Procedures for establishing a DCR.....	4
5.2 Profile experts.....	5
5.2.1 Description.....	5
5.2.2 Roles and responsibilities of profile experts.....	5
5.3 Profile managers.....	5
5.3.1 Description.....	5
5.3.2 Roles and responsibilities.....	5
5.4 DCR Management Board.....	6
5.4.1 Description.....	6
5.4.2 Roles and responsibilities.....	6
<b>6 Workflow for data category specification and maintenance</b> .....	<b>6</b>
<b>7 Public information and interactions</b> .....	<b>6</b>
<b>Bibliography</b> .....	<b>8</b>

iteh STANDARD PREVIEW  
 (standards.iteh.ai)  
 oSIST ISO/DIS 12620-2:2021  
<https://standards.iteh.ai/catalog/standards/sist/6158nac-5016-4693-8237-8-61188f33/osit-iso-dis-12620-2-2021>

## ISO/DIS 12620-2:2021(E)

### 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](http://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](http://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 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](http://www.iso.org/iso/foreword.html).

This document was prepared by Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 3, *Management of terminology resources*.

A list of all parts in the ISO 12620 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](http://www.iso.org/members.html).

## Introduction

ISO 12620-1, *Management of terminology resources — Data categories—Part 1: Specifications*, provides guidelines and requirements governing data category specifications for language resources. It specifies mechanisms for creating, documenting, harmonizing and maintaining data category specifications in a data category repository (DCR) with the goal of increasing the interoperability of language resources, such as terminological resources, lexicographical resources, and annotated text corpora. This document, *Part 2: Repositories*, specifies procedures and practices for the creation, management, and maintenance of DCRs.

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[oSIST ISO/DIS 12620-2:2021](https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021)

<https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021>

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[oSIST ISO/DIS 12620-2:2021](https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021)

<https://standards.iteh.ai/catalog/standards/sist/b138ffac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021>



# Management of terminology resources — Data categories —

## Part 2: Repositories

### 1 Scope

This document describes criteria for the management of a repository of data category specifications for use in the creation and maintenance of language resources within a given community of practice (CoP) in compliance with ISO 12620-1. The goal of this document is to define the roles and responsibilities associated with the creation and maintenance of such repositories. It also describes procedures whereby the community shall establish a governance structure for the management of a data category repository, as well as the addition of new data category specifications, and continuous quality assurance procedures.

### 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 12620-1<sup>1)</sup>, *Management of terminology resources — Data categories — Part 1: Specifications*  
<https://standards.iteh.ai/catalog/standards/sist/b138fac-30f6-4693-8237-8c6b188fc2f3/osist-iso-dis-12620-2-2021>

### 3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 12620-1 and the following 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/>

#### 3.1

##### **data category specification**

##### **DC specification**

complete descriptive record of a data category

#### 3.2

##### **data category repository**

##### **DCR**

digital collection of *data category specifications* (3.1)

Note 1 to entry: A DCR for language resources is available at [www.datcatinfo.net](http://www.datcatinfo.net), which points to <https://datcatinfo.termweb.se/termweb/app?service=page/Home> [3].

1) Under preparation. (Stage at the time of publication ISO/WD 12620-1.)

## ISO/DIS 12620-2:2021(E)

### 3.3 data category selection DC selection

set of *data category specifications* (3.1) selected from a *DCR* (3.2)

Note 1 to entry: A data category selection can represent the data categories used within a research discipline or a specific application or project.

### 3.4 community of practice CoP

group of people who share a concern or interest in a specific activity or set of activities-

EXAMPLE terminologists

Note 1 to entry: A CoP can be, but is not necessarily, a formal organization.

### 3.5 DCR Management Board

group of experts selected by a *community of practice* (3.4) to oversee the creation, maintenance, and administration of a *data category repository* (3.2) for a given subject field

Note 1 to entry: A DCR Management Board also serves as an information source for members of the community in publicizing the activities of the DCR.

### 3.6 profile

selection of *data category specifications* (3.1) within a *Data Category Repository* (3.2) that pertain to a specific subfield of the broader field of knowledge covered by that DCR

EXAMPLE terminology profile, lexicography profile, semantic annotation profile

### 3.7 Profile Management Group

group of experts responsible for managing *data category specifications* (3.1) assigned to a particular *profile* (3.6)

## 4 Data category repositories

### 4.1 General interoperability

Interoperability of language resources is a key factor for supporting innovation and progress in various focus areas of the language industry, such as terminology management, natural language processing, and annotation schemes. These areas support important sectors of our economy and social development such as global communication and trade, knowledge extraction, and content management.

Researchers and software developers working with language resources benefit greatly from being able to access a trusted source of information about data categories. Providing a precise description of the data categories that are used within a given data collection allows for a quick diagnosis of its compatibility with other data collections or its suitability for use in computer processes. A DCR containing vetted data category specifications provides users with the information they need in order to implement data categories in a manner that is consistent with other users. Consequently, the interoperability of language resources is greatly enhanced.

ISO 12620-1 describes requirements for the specification of data categories used in language resources. Data category specifications are normally stored in electronic format in a specially designed database. This database is called a *data category repository* (DCR). Today, it is essential for DCRs to be available on the internet. For instance, a DCR for language resource descriptions, named DatCatInfo, is available at [www.datcatinfo.net](http://www.datcatinfo.net), which links to <https://datcatinfo.termweb.se/termweb/app?service=page/Home>, see Reference [3] in the Bibliography.