

BS ISO 24622-1:2015



BSI Standards Publication

**Language resource  
management — Component  
Metadata Infrastructure (CMDI)**  
Part 1: The Component Metadata Model

**bsi.**

...making excellence a habit.™

**National foreword**

This British Standard is the UK implementation of ISO 24622-1:2015.

The UK participation in its preparation was entrusted to Technical Committee TS/1, Terminology.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2015.  
Published by BSI Standards Limited 2015

ISBN 978 0 580 81379 5  
ICS 01.140.20

**Compliance with a British Standard cannot confer immunity from legal obligations.**

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 January 2015.

**Amendments/corrigenda issued since publication**

Date	Text affected
------	---------------

---

---

---

**Language resource management —  
Component Metadata Infrastructure  
(CMDI) —**

Part 1:  
**The Component Metadata Model**

*Gestion des ressources langagières — Composante infrastructure de  
métadonnées (CMDI) —*

*Partie 1: Composant modèle de métadonnées*





**COPYRIGHT PROTECTED DOCUMENT**

© ISO 2015

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  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [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 Terms and definitions</b> .....	<b>1</b>
<b>3 Metadata schema availability and reuse</b> .....	<b>5</b>
3.1 Overview.....	5
3.2 Metadata components and elements.....	5
<b>4 Semantics in the component metadata model</b> .....	<b>7</b>
4.1 Overview.....	7
4.2 Concept registries.....	8
4.3 Relation registries.....	8
<b>5 Metadata component and profile - compatibility and versioning</b> .....	<b>9</b>
<b>6 Expressiveness of the component metadata model</b> .....	<b>9</b>
<b>Annex A (informative) Abbreviations</b> .....	<b>10</b>
<b>Bibliography</b> .....	<b>11</b>

## 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 on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT), see the following URL: [Foreword — Supplementary information](#).

The committee responsible for this document is ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 4, *Language resource management*.

ISO 24622 consists of the following part, under the general title *Language resource management — Component metadata infrastructure (CMDI)*:

— *Part 1: The component metadata model*

A future part will address the component metadata specific language.

## Introduction

Component Metadata (CMD) is an approach to metadata modelling and metadata creation. It is being increasingly used these days to enable the metadata description of different types of Language Resources (LRs) with different metadata schemas, while still trying to maintain syntactic and semantic interoperability.

CMD<sup>1)</sup> is also the core of the Component Metadata Infrastructure (CMDI)<sup>[1]</sup>: this infrastructure contains not only the format specifications for this metadata modelling and creation approach, but also a set of registries and tools for metadata modelling and creation work.

The advantages of having such a unified approach to metadata descriptions for LRs, an approach that will be usable by many projects and initiatives, are obvious: firstly, there is a better chance of obtaining interoperability between metadata descriptions from different sources, and secondly, it will be possible to develop and share tools that work much more efficiently in this metadata framework.

The challenge of designing and organizing a comprehensive and unified approach to metadata description for the very varied set of LR types, and one that also can satisfy a sufficiently large section of the LR community, should not be underestimated. The landscape of metadata for LRs has been, and continues to be, fragmented. Until recently, it was the practice in creating the metadata descriptions for LRs to choose a specific metadata schema from a (small) existing set derived either from widespread traditions or from other disciplines; for example, OLAC<sup>[2]</sup> is an adapted version of DCMI,<sup>[3]</sup> which in turn originates in the library world. Additionally, there are, for the purposes of LR metadata description, specifically developed metadata schemas that can be limited in application to specific types of LR (e.g. IMDI<sup>[4]</sup>), or they can be of a proprietary nature (cf. the catalogues of the LR agencies such as LDC<sup>2)</sup> and ELRA<sup>3)</sup>). The result is a domain of LR metadata that is far from interoperable. Although some progress has been made in developing dedicated bridges for “translating” metadata from one specific schema to another and in providing a consolidated catalogue, this practice does not scale well since it depends on specific translations for each pair of different metadata schemas.

For some recent projects, founding principles have included the unification and consolidation of practices and the need to produce efficient and sufficiently specific metadata descriptions.

It follows that a number of international, European, and national projects and infrastructure initiatives such as CLARIN<sup>[5]</sup> and META-SHARE<sup>[6]</sup> now share the CMD approach to metadata for LRs. This International Standard will both standardize the fundamentals of this approach in order to achieve interoperability based on solid documentation, and foster cooperation between the various initiatives and projects that work on, and with, this International Standard.

The model description is the first part of an infrastructure that forms a complete package for the creation of metadata schemas. As stated in the Foreword, the complete infrastructure standard contains, in addition to this component metadata model specification (ISO 24622-1), one or more metadata component specification languages (planned), and a number of recommended metadata components and profiles (planned). Since this part of ISO 24622 specifies an abstract model, we will rely mainly on UML<sup>[7]</sup> to describe it.

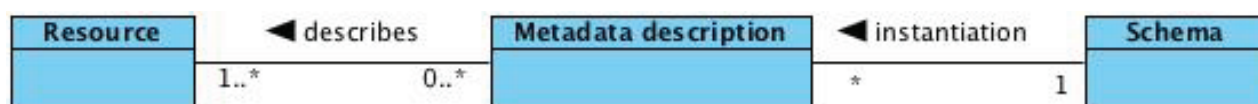


Figure 1 — Describing resources with metadata

- 1) Abbreviations are explained in [Annex A](#).
- 2) Linguistic Data Consortium, <http://www.ldc.upenn.edu/>
- 3) European Language Resources Association, <http://www.elra.info/>

This part of ISO 24622 addresses the basic need to provide a model that makes it easy for metadata modellers (e.g. researchers and resource description experts) to create new metadata schemas, which can in turn be used either to describe new types of resources or to enable a more appropriate description for resources in specific circumstances. The metadata schema is instantiated into metadata records [i.e. the metadata descriptions that describe the actual resource(s)] (see [Figure 1](#)).

The context of this desire for flexible metadata modelling is that for scientific work there are usually various requirements for the proper description of LRs, and these requirements can derive from the specific needs of a project or from the facility or repository that will be used to store the resource for future use. This variation requires a flexible framework that enables the easy creation of new metadata schemas for different purposes, but is also a framework (i) in which the instantiations have a strictly defined format so that at least syntactic correctness can be checked, and (ii) which provides explicit semantics for the metadata schema elements for interpretation of the metadata record content.

The metadata descriptions generated by schemas compliant with this model will also be compliant with other TC 37 International Standards, for example, those requiring that references to the described resources and resource parts use ISO 24619:2011 PISA-compatible persistent identifiers (PIDs)<sup>[9]</sup>.

The definition of a resource in this context is very broad. This part of ISO 24622 takes a pragmatic view: for example, an image can be a resource in itself when it is associated with a PID and can be referenced as such, or it can be part of a document where it lacks an identity of its own. In addition, a reference can point to a part of this image. An individual resource can stand alone in one environment and be treated as part of a collection in another environment. Also, metadata descriptions describe resources, but they, too, are a resource in different contexts. This part of ISO 24622 needs to support all such cases, and the model needs to provide descriptions at all levels of granularity.

This part of ISO 24622 takes two types of collections into account:

- a) A complex resource may have been created as a collection originally and, versioning aside, it will exist as such in a rather static published form. Its specification will be treated as an independent entity by the responsible archiving institution that also provides a PID for such a collection. In the context of this part of ISO 24622, the metadata for the collection is the collection specification. The archiving institution is responsible for maintaining the metadata representing the collection.
- b) In contrast, a different type of collection is one that was not planned and designed as a collection by its creators or by the holding archive, but achieves its status as a federated resource based on research that needs to be verifiable. Such collections, although purposefully constructed by the researcher, may not have any significance outside the context of the research for which they were created. Referring from the research documents to the collection may also become tedious if the collection contains hundreds of individual resources. It follows that there is a need to capture these types of collection with a metadata record that is associated with all its constituent resources and appropriate metadata, but only as the incarnation of this collection. There is no natural responsible party to maintain this metadata record. It is unlikely that the researcher who created the “virtual” collection (VC) has any way of consistently maintaining and curating this metadata record in the long term. There may be special registries maintained by digital archives or publishers where researchers can register such virtual collections.

Both types of collection are identified with the PID that refers to the collection metadata.



# Language resource management — Component Metadata Infrastructure (CMDI) —

## Part 1: The Component Metadata Model

### 1 Scope

The scope of this part of ISO 24622 is to describe a model that enables the flexible construction of interoperable metadata schemas for Language Resources (LRs). The metadata schemas based on this model can be used to describe resources at different levels of granularity (e.g. descriptions both on the collection level and on the level of individual resources).

### 2 Terms and definitions

#### 2.1

##### archive

##### digital archive

*repository* (2.26) dedicated to the long-term preservation of the associated data

Note 1 to entry: The data in digital archives are also often available on-line. This highlights the need for reliable *PIDs* (2.22)

#### 2.2

##### cardinality

##### metadata component cardinality

##### metadata element cardinality

specification of the number of occurrences of a *metadata component* (2.14) or *metadata element* (2.12) in an instantiation

#### 2.3

##### citation

object containing information that directs a textual resource reader's or user's attention from one resource to another

#### 2.4

##### closed vocabulary

limited set of items that forms the mandatory value domain of a *metadata element* (2.12)

#### 2.5

##### concept reference

##### concept link

reference to the definition of a concept in a *concept registry* (2.6)

#### 2.6

##### concept registry

*registry* (2.25) for registering concepts enabling their identification with a unique identifier

## 2.7 collection

### resource collection

grouping of multiple, different constituting elements, each of which is independent of the others and may be accessed individually

Note 1 to entry: A collection can be a virtual collection if its constituent elements come from other different (virtual) collections, and possibly if the elements are distributed over different repositories.

## 2.8 fragment identifier

*identifier* ([2.9](#)) used to reference a *resource part* ([2.28](#)) in a web context

[SOURCE: ISO 12619:2011]

## 2.9 identifier digital identifier

compact sequence of characters associated with digital, non-digital, or abstract entities

[SOURCE: Adapted from ISO 12619:2011]

Note 1 to entry: Identifiers can apply to entities such as books, images, reports, metadata records, and events.

## 2.10 metadata record metadata description metadata

*record* ([2.23](#)) containing a description of a *resource* ([2.27](#))

## 2.11 metadata schema schema

specification of a format and structure for a *metadata record* ([2.10](#))

Note 1 to entry: In the context of this part of ISO 24622, a machine-readable and verifiable format specification usually defined by an XML schema language.

## 2.12 metadata element

resource property name that can be used in metadata and that can be given a value

Note 1 to entry: A metadata element is referred to as metadata attribute in other communities.

EXAMPLE The DCMI elements.<sup>[3]</sup>

## 2.13 metadata set metadata element set

collection of *metadata elements* ([2.12](#)) used within a particular discipline, tradition, or practice to describe *resources* ([2.27](#))

Note 1 to entry: A metadata set is more general than a metadata schema in that it does not additionally specify the syntax (e.g. the DCMI elements<sup>[3]</sup>).

## 2.14 metadata component

grouping of *metadata elements* ([2.12](#)) and *metadata components* ([2.14](#)) that can be used to describe a specific aspect of a *resource* ([2.27](#))

EXAMPLE The biographical data of a person or the contact information for an organization.

## 2.15

### **metadata component registry** **component registry**

*registry* (2.25) of *metadata components* (2.14) and *metadata profiles* (2.16) for their sharing

## 2.16

### **metadata profile**

set of *metadata components* (2.14) that can be used together to describe a *resource* (2.27) and be transformed into a *metadata schema* (2.11)

Note 1 to entry: A metadata profile can be transformed into different metadata schemas that are still logically equivalent (i.e. they give logically equivalent resource descriptions).

## 2.17

### **metadata editor**

actor that creates *metadata records* (2.10) to describe specific *resources* (2.27) or the tool that is used to edit the metadata record

## 2.18

### **metadata modeler**

actor that creates new *metadata schemas* (2.11) for new types of *resources* (2.27) or new applications

Note 1 to entry: In this part of ISO 24622, metadata schemas are created by producing *metadata profiles* (2.16), which in turn form specifications for a metadata schema.

## 2.19

### **metadata provider**

organization or software service that makes *metadata* (2.10) available

## 2.20

### **open vocabulary**

set of items forming part of the value domain of a *metadata element* (2.12) on the recommendation of the *metadata modeler* (2.18)

## 2.22

### **Persistent Identifier**

#### **PID**

unique *identifier* (2.9) that ensures permanent access to a digital object by providing access to it independently of its physical location or current ownership

[SOURCE: 24619:2011]

Note 1 to entry: In this context, “unique” means that the same PID will not be subsequently issued for other *resources* (2.27). However, the same PID may, at the *resource-provider’s* (2.30) discretion, reference different representations or incarnations of the *resource* (2.27)

## 2.23

### **record**

structured information that can be read by software services

## 2.24

### **reference**

object that links to data stored elsewhere

Note 1 to entry: The words “citation” and “reference” are commonly used as quasi-synonyms. However, for the purposes of this part of ISO 24622, “citations” provide information for human readers and users, while “references” include the precise location where the referenced resource can be found. References can be machine-readable, and can be configured as actionable given the required criteria.

## 2.25

### **registry**

central directory designed for the persistent provision of negotiated information that can be reliably accessed

Note 1 to entry: In this part of ISO 24622, a registry is a software service that allows registering and for the registry to be queried for information.

## 2.26

### **repository**

#### **digital repository**

facility that provides reliable access to managed digital *resources* ([2.27](#))

## 2.27

### **resource**

object on the web with a specific identity

Note 1 to entry: For the purpose of this part of ISO 24622, the identity is one that can be expressed using a *URI* ([2.32](#))

Note 2 to entry: In the context of this part of ISO 24622, a resource can also be a language resource that has an off-line representation

## 2.28

### **resource part**

identifiable, accessible entity embedded in an independent *resource* ([2.27](#)) or in a larger part thereof

Note 1 to entry: Resource parts can be part of larger resource parts. In dynamic web environments, subsetting into parts is subject to change and interpretation, and this in turn requires a certain level of user decision-making to designate and identify such sub-entities.

## 2.29

### **resource part identifier**

string of characters that refers to a *resource part* ([2.28](#)), and which can be identified by some means within a given resource type

EXAMPLE      Such means are time for a media file, area for an image or record in a data stream.

## 2.30

### **resource provider**

organization that makes a *resource* ([2.27](#)) available on-line

## 2.31

### **Unified Modeling Language**

#### **UML**

language for specifying, visualizing, constructing, and documenting the artifacts of software systems and abstract models in general

## 2.32

### **Uniform Resource Identifier**

#### **URI**

compact string of characters used to identify or name a *resource* ([2.27](#)) with a syntax defined in IETF RFC 3986

## 2.33

### **value scheme**

#### **metadata element value scheme**

specification of the value domain of a *metadata element* ([2.12](#))

### 3 Metadata schema availability and reuse

#### 3.1 Overview

There is heavy demand in science and industry for a flexible description of digital language resources, resource parts, and collections of language resources. Not only is it desirable to be able to store and manage resources with high quality metadata for later use, but it is also becoming increasingly useful to use the metadata itself as a scientific resource. This makes it essential for there to be well-designed metadata schemas that are able to capture the maximum relevant information.

Although LR's are marked by considerable variety, and resource types accordingly warrant the use of many different metadata schemas, there is often a considerable information content overlap between these schemas. To minimize the time spent on creating new schemas, this part of ISO 24622 specifies a model with a flexible component-based approach that allows metadata modellers to create new schemas by reusing existing metadata modelling work captured in so-called metadata components. Metadata component specifications are specified in a component specification language and can be saved for future use, for instance in metadata component registries. To describe a specific resource type, a metadata modeller selects a number of appropriate metadata components and bundles these in a metadata profile. The metadata profile is specified in the metadata component specification language and is also stored for future use or extension, for example, in a metadata component registry.

To be functional, a metadata profile needs to be transformed into a metadata schema before it can be used to generate metadata descriptions. This approach resembles that taken by the metadata application profile initiative [8]. However, the use of metadata components turns the creation of a new schema by metadata modellers into a conceptual exercise rather than one of (XML) schema building.

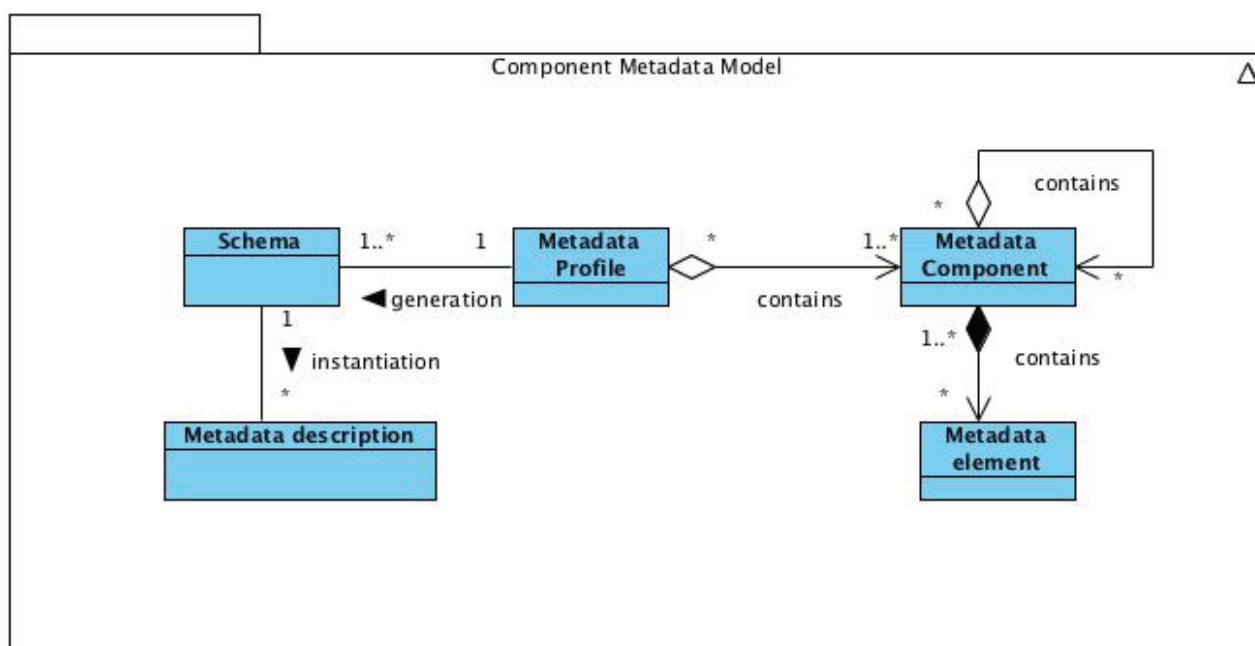


Figure 2 — The component metadata model

Figure 2 shows the relations between the metadata component, the metadata elements, and the metadata profiles. A metadata profile can be transformed into a schema specification from which the actual metadata descriptions can be instantiated.

#### 3.2 Metadata components and elements

Metadata components can contain metadata elements and other metadata components, and the component model is therefore recursive (see Figure 2). Metadata components have a name and an

optional concept reference, and can have cardinality that specifies how often the metadata component can occur in a metadata record. The name of the metadata component is the name that the metadata component creator thinks is appropriate for the component.

A metadata element describes a single atomic aspect of a resource; it has a name, a concept reference, a value range specification, and cardinality (see [Figure 3](#)). The metadata element name is the name of the resource aspect that the metadata component creator believes to be appropriate. The concept reference is a PID and refers to the semantic definition of the metadata element (see the Semantics section in [Clause 4](#) Semantics in the component metadata model). The value range or value scheme should provide an appropriate restriction of the values that the metadata element can take. The metadata element will, in the resultant instantiated metadata record, provide a real value for the resource aspect in the metadata description, for example:

*element name: Country*

*element concept reference: (URI for wiki def.<sup>4</sup>)*

*element value scheme: ISO 3166-2*

*element cardinality: 0..1*

The value scheme of a metadata element can take several forms:

- a) **data types** specific to the implementation language (to be defined in CMDI Part-2);
- b) **closed vocabulary**;
- c) **open vocabulary**;
- d) **regular expression**.

If the value scheme is an open or closed vocabulary, the items in this vocabulary need to have their own concept references (see [Clause 4](#)). The metadata component shall provide such a list of items in the metadata component itself or else a reference to a remote source for the vocabulary.

A metadata component can also be optionally associated with a concept reference; if so, a PID pointing to a concept in a concept registry is required.

The metadata element can have a cardinality specified; if a CMDI implementation supports cardinalities, it shall support at least the following options:

- a) **cardinality (number of occurrences): \*** (short for 0..\*) is the default and stands for zero or more;
- b) **cardinality: 0..1** stands for zero or exactly one;
- c) **cardinality: 1..\*** stands for minimal one;
- d) **cardinality: 1** stands for exactly one.

The last two cases are examples of obligatory elements, that is to say the metadata component or element shall occur at least once in every instantiation.

---

4) Some may not see a wiki as an acceptable concept registry.

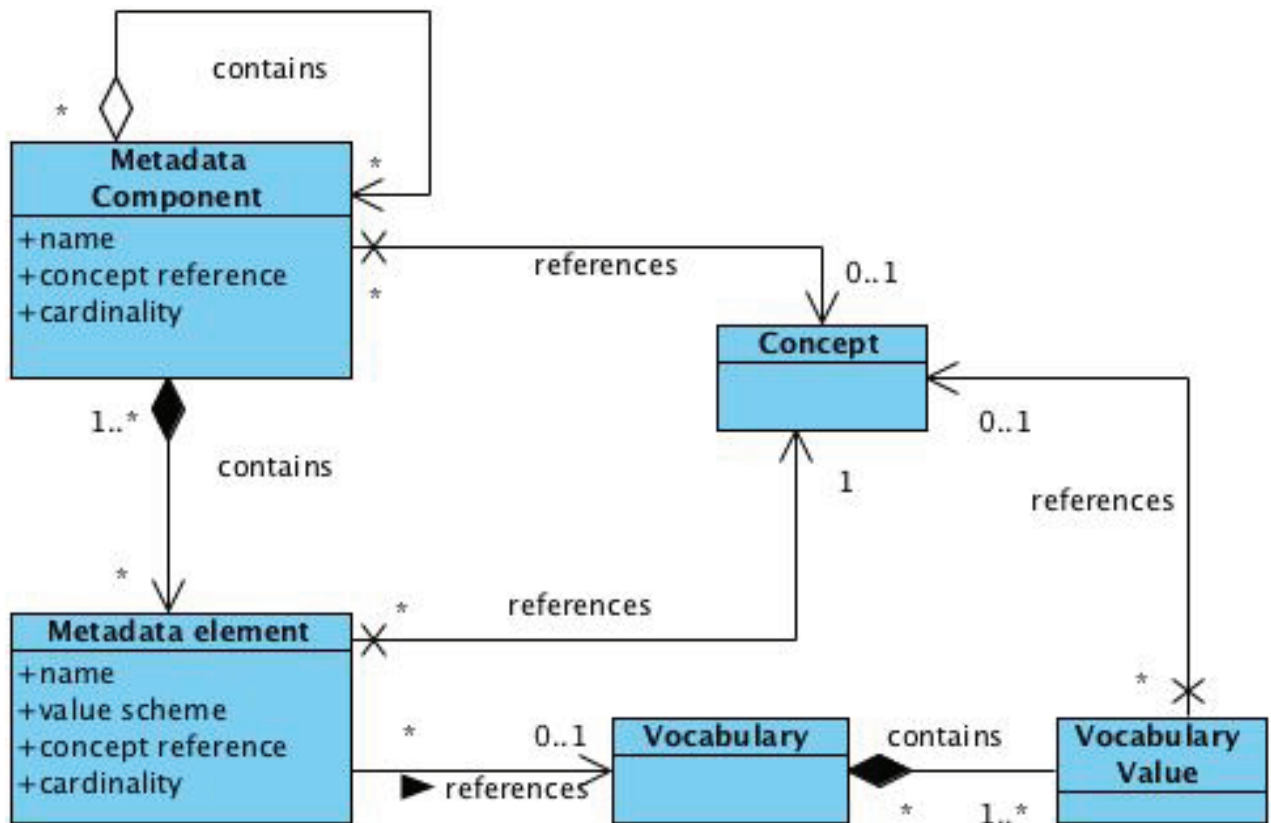


Figure 3 — Metadata component and element detail

## 4 Semantics in the component metadata model

### 4.1 Overview

The metadata component framework encourages the creation of new metadata components and profiles to facilitate accurate descriptions of resources.

The terminology used within metadata profiles and metadata components is determined by the metadata modeller; this enables communities and even individual modellers to use familiar names that match their traditions and which they feel comfortable using. However, to make semantic interoperability between metadata components possible, every metadata element must refer to a concept definition in a recognized concept registry. Metadata components may also be associated with a concept definition, although this is not required (see [Figure 4](#)). Using the concept references, semantic relations can be computed between metadata records originating from different metadata schemas. This can be used by metadata service providers that collect metadata records from communities using different metadata schemas and wish to show all the metadata in one consolidated metadata catalogue.

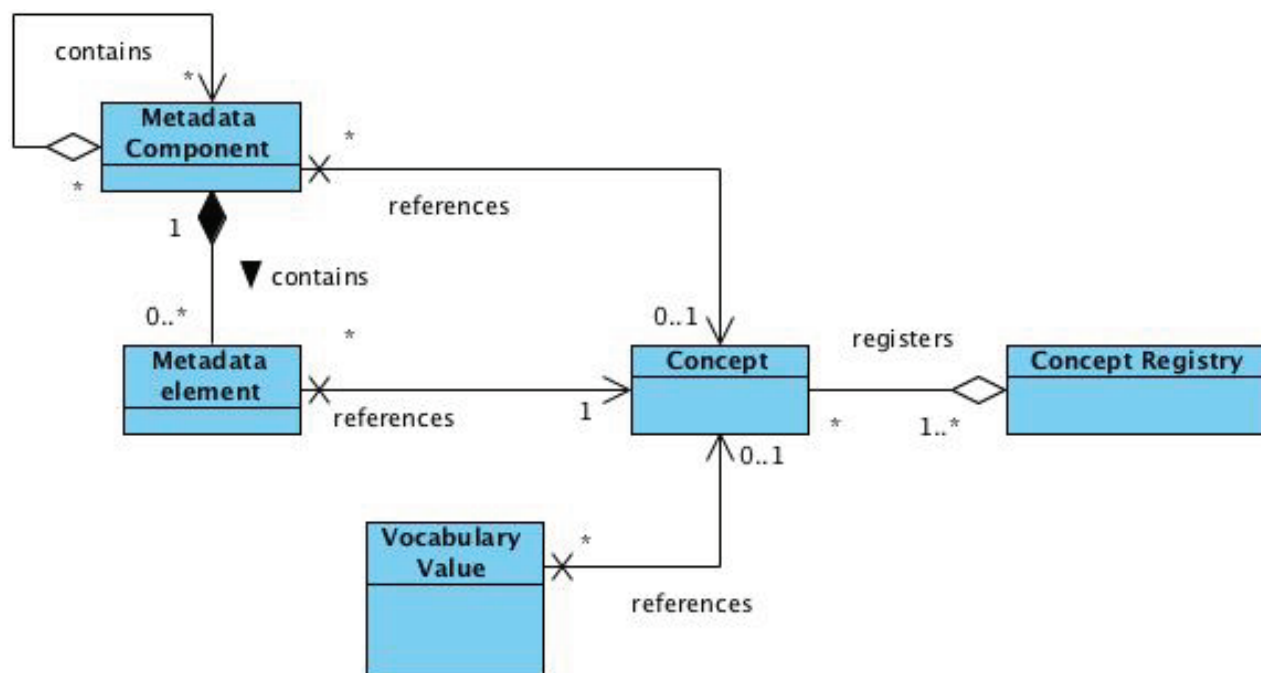


Figure 4 — Concept references in CMDI

## 4.2 Concept registries

A concept registry is a directory that stores definitions of concepts together with a unique identifier for each concept. The identifier is used to compute semantic relations between different metadata records.

The minimum requirements for an “accepted” concept registry are the following:

- stable and persistent accessibility;
- unique identifiers for registered concepts;
- at least an English language definition section;
- a transparent versioning and governance policy.

The ISOcat DCR, an implementation of ISO 12620:2009<sup>[10]</sup> under the control of TC 37, can be used as one such concept registry; DCMI is another.

## 4.3 Relation registries

A relation registry is a directory that stores semantic relations between concepts. Some concept registries are capable of registering relations between the registered concepts to varying degrees. If, in the semantic processing of the metadata records, the concept registry does not provide sufficient expressiveness, an additional relation registry can be introduced that is able to correct this deficiency. The relation registry is a directory that stores a relation type and the (concept registry) concept identifiers involved.



## 5 Metadata component and profile - compatibility and versioning

This part of ISO 24622 does not require any versioning or compatibility relations between metadata components and between profiles. However, if an implementation does provide one, it shall provide also the following versioning relations:

- component A is a *specialization* of component B if, and only if, all metadata elements present within A are also present within B;
- component A is a *simplification* of component B if B is a specialization of A;
- a component or profile can be specified as deprecated, and it shall not subsequently be used for new metadata modelling, although it shall remain available for reference purposes.

## 6 Expressiveness of the component metadata model

Figure 5 offers a more detailed version of the Component Metadata Model, and shows the model's expressive power. The serialization formats (as proposed in CMDI Part 2) shall provide the user with at least the same expressive power. In other words, instantiations of the model created with this format should not be lax by comparison with the model; for example, the constraint that a metadata component should contain at least one metadata component or one metadata element should always hold. However, depending on its application context, a serialization format may also be stricter: for example, a metadata component could be forbidden to contain itself, this is permitted by the metadata component model, but it may not be technically possible for a specific CMDI implementation.

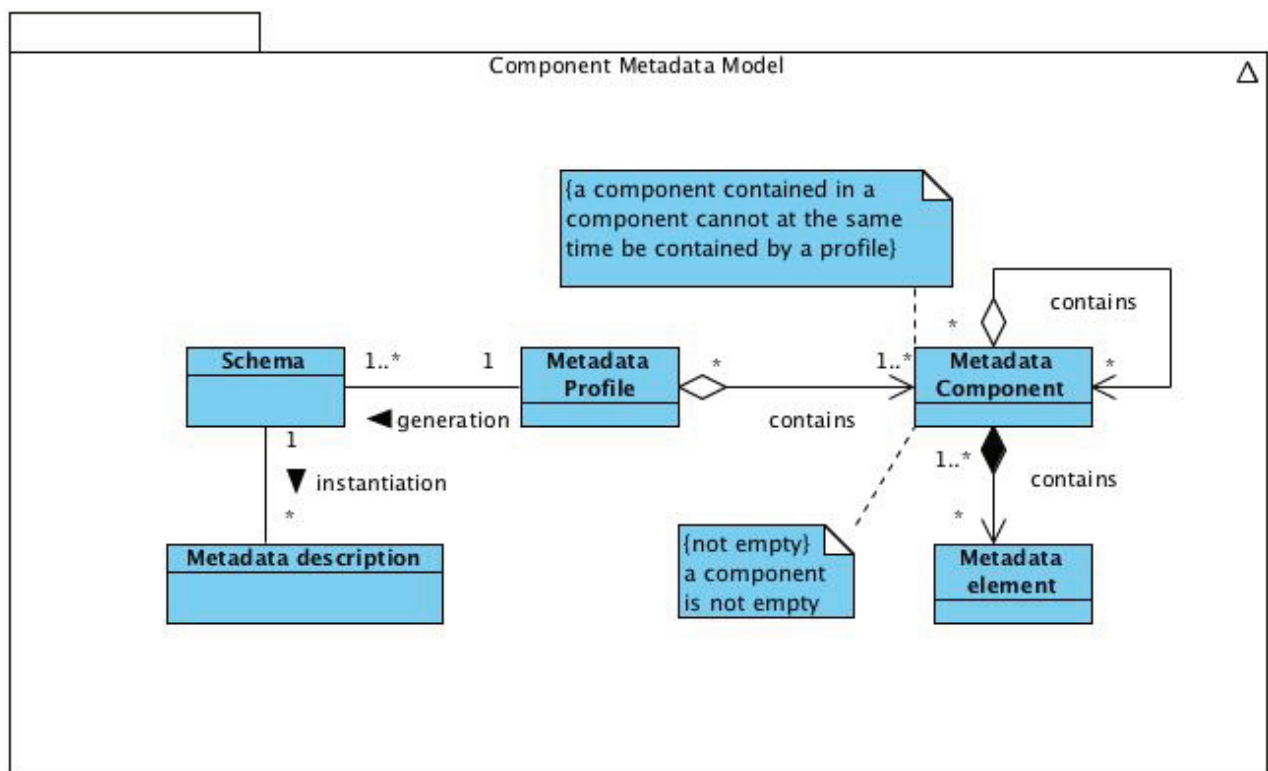


Figure 5 — A detailed version of the component metadata model

## Annex A (informative)

### Abbreviations

CMDI	Component Metadata Infrastructure	Metadata description framework consisting of the CMDI model, specification language, and tools
DCMI	Dublin Core Metadata Initiative	Set of metadata elements to describe a broad type of resource
DCR	Data Category Registry	Formally defined as set of linguistic categories for reference and use in annotated resources
IMDI	ISLE Metadata Initiative	A metadata framework for linguistic resources
IETF	Internet Engineering Task Force	Entity that develops Internet standards
IMDI	ISLE Metadata Initiative	XML-based metadata standard for language resources
LR	Language Resource	Sets of language data and descriptions in machine-readable form, used specifically for research or language technology development
OLAC	Open Language Archives Community	Extension of DCMI to describe Language Resources
PID	Persistent Identifier	Sustainable reference to an online resource
RDF	Resource Description Framework	Description language for the semantic web
RFC	Request for Comments	Standard proposal by the IETF intended for peer review
UML	Unified Modelling Language	Language for specifying and documenting software artefacts. Also used for abstract models.
URI	Uniform Resource Identifier	A compact string of characters used to identify or name a resource
XML	Extensible Markup Language	Formal language intended for the sharing of structured data

## Bibliography

- [1] BROEDER D., SCHONEFELD O., TRIPPEL T., VAN UYTVANCK D., WITT A. (2011). A pragmatic approach to XML interoperability — the component metadata infrastructure (CMDI). In *Balisage: The Markup Conference 2011*, volume 7.
- [2] OLAC. <http://www.language-archives.org/>, retrieved 1-5-2012
- [3] DUBLIN CORE METADATA INITIATIVE (DCMI). Terminology'. <http://www.ukoln.ac.uk/metadata/dcmi/abstract-model/2004-12-08/#sect-7>
- [4] BROEDER D., & WITTENBURG P. The IMDI metadata framework, its current application and future direction. *International Journal of Metadata. Semantics and Ontologies*. 2006, **1** (2) pp. 119–132. DOI:
- [5] VÁRADI T., WITTENBURG P., KRAUWER S., WYNNE M., KOSKENNIEMI K. (2008). CLARIN: Common language resources and technology infrastructure. LREC 2008.
- [6] GAVRILIDOU M., LABROPOULOU P., PIPERIDIS S., MONACHINI M., FRONTINI F., FRANCOPOULO G., ARRANZ V., MAPELLI V. (2011). A Metadata Schema for the Description of Language Resources (LRs), IJCNLP 2011.
- [7] RUMBAUGH J., JACOBSON I., BOOCH G. *The Unified Modeling Language Reference Manual*. Pearson Higher Education, Second Edition, 2004
- [8] APPLICATION PROFILES. <http://dublincore.org/documents/profile-guidelines/>, retrieved 1-5-2012.
- [9] ISO 24619:2011, *Language resource management — Persistent identification and sustainable access (PISA)*
- [10] ISO 12620:2009, *Terminology and other language and content resources — Specification of data categories and management of a Data Category Registry for language resources*





# British Standards Institution (BSI)

BSI is the national body responsible for preparing British Standards and other standards-related publications, information and services.

BSI is incorporated by Royal Charter. British Standards and other standardization products are published by BSI Standards Limited.

## About us

We bring together business, industry, government, consumers, innovators and others to shape their combined experience and expertise into standards-based solutions.

The knowledge embodied in our standards has been carefully assembled in a dependable format and refined through our open consultation process. Organizations of all sizes and across all sectors choose standards to help them achieve their goals.

## Information on standards

We can provide you with the knowledge that your organization needs to succeed. Find out more about British Standards by visiting our website at [bsigroup.com/standards](http://bsigroup.com/standards) or contacting our Customer Services team or Knowledge Centre.

## Buying standards

You can buy and download PDF versions of BSI publications, including British and adopted European and international standards, through our website at [bsigroup.com/shop](http://bsigroup.com/shop), where hard copies can also be purchased.

If you need international and foreign standards from other Standards Development Organizations, hard copies can be ordered from our Customer Services team.

## Subscriptions

Our range of subscription services are designed to make using standards easier for you. For further information on our subscription products go to [bsigroup.com/subscriptions](http://bsigroup.com/subscriptions).

With **British Standards Online (BSOL)** you'll have instant access to over 55,000 British and adopted European and international standards from your desktop. It's available 24/7 and is refreshed daily so you'll always be up to date.

You can keep in touch with standards developments and receive substantial discounts on the purchase price of standards, both in single copy and subscription format, by becoming a **BSI Subscribing Member**.

**PLUS** is an updating service exclusive to BSI Subscribing Members. You will automatically receive the latest hard copy of your standards when they're revised or replaced.

To find out more about becoming a BSI Subscribing Member and the benefits of membership, please visit [bsigroup.com/shop](http://bsigroup.com/shop).

With a **Multi-User Network Licence (MUNL)** you are able to host standards publications on your intranet. Licences can cover as few or as many users as you wish. With updates supplied as soon as they're available, you can be sure your documentation is current. For further information, email [bsmusales@bsigroup.com](mailto:bsmusales@bsigroup.com).

## BSI Group Headquarters

389 Chiswick High Road London W4 4AL UK

## Revisions

Our British Standards and other publications are updated by amendment or revision.

We continually improve the quality of our products and services to benefit your business. If you find an inaccuracy or ambiguity within a British Standard or other BSI publication please inform the Knowledge Centre.

## Copyright

All the data, software and documentation set out in all British Standards and other BSI publications are the property of and copyrighted by BSI, or some person or entity that owns copyright in the information used (such as the international standardization bodies) and has formally licensed such information to BSI for commercial publication and use. Except as permitted under the Copyright, Designs and Patents Act 1988 no extract may be reproduced, stored in a retrieval system or transmitted in any form or by any means – electronic, photocopying, recording or otherwise – without prior written permission from BSI. Details and advice can be obtained from the Copyright & Licensing Department.

## Useful Contacts:

### Customer Services

**Tel:** +44 845 086 9001

**Email (orders):** [orders@bsigroup.com](mailto:orders@bsigroup.com)

**Email (enquiries):** [cservices@bsigroup.com](mailto:cservices@bsigroup.com)

### Subscriptions

**Tel:** +44 845 086 9001

**Email:** [subscriptions@bsigroup.com](mailto:subscriptions@bsigroup.com)

### Knowledge Centre

**Tel:** +44 20 8996 7004

**Email:** [knowledgecentre@bsigroup.com](mailto:knowledgecentre@bsigroup.com)

### Copyright & Licensing

**Tel:** +44 20 8996 7070

**Email:** [copyright@bsigroup.com](mailto:copyright@bsigroup.com)



...making excellence a habit.™