Food and Agriculture Organization
of the United Nations

Linked Open Data Enabled Bibliographical Data (LODE-BD) 3.0

A practical guide on how to select appropriate encoding strategies for producing Linked Open Data Enabled Bibliographical Data

Linked Open Data Enabled Bibliographical Data (LODE-BD) 3.0

A practical guide on how to select appropriate encoding strategies for producing Linked Open Data Enabled Bibliographical Data

by
Imma Subirats
Food and Agriculture Organization of the United Nations
and
Marcia Lei Zeng
Kent State University

Food and Agriculture Organization of the United Nations
Rome, 2020

Copyright

Required citation:
Subirats, I. and Zeng, M.L. 2020. *Linked Open Data Enabled Bibliographical Data (LODE-BD) 3.0 – A practical guide on how to select appropriate encoding strategies for producing Linked Open Data Enabled Bibliographical Data*. Rome, FAO.
https://doi.org/10.4060/cb2209en

ISBN 978-92-5-133655-7

Contents

Acknowledgements

This version is based on LODE-BD 1.1 that was partially supported by the European Commission through the ICT PSP Grant #250525 for VOA3R (Virtual Open Access Agriculture & Aquaculture Repository: Sharing Scientific and Scholarly Research related to Agriculture, Food, and Environment). The authors also would like to thank Ioannis N. Athanasiadis, Nikos Manouselis, Ilias Hatzakis, Tom Baker, Gordon Dunsire, Hugo Besemer, Fernanda Peset, Xavier Agenjo, Francisca Hernández, MacKenzie Smith, Karen Coyle, Antoine Issac and the data providers of the VOA3R team for their support and advice throughout the completion of this publication, as well as Laurence Skirvin who assisted in the compiling of the crosswalk of Schema.org (see appendix). The authors also want to thank Gordon Dunsire, especially for his inputs into the LODE-BD 2.0 version.

Abbreviations and acronyms

AGLS	Australian Government Locator Service Metadata Standards
AGRIS	International System for Agricultural Science and Technology
BIBO	Bibliographic Ontology
DC	Dublin Core Metadata Element Set
DCAT	Data Catalog Vocabulary
DCTERMS	DCMI Metadata Terms
DOI	Digital Object Identifier
FAO	Food and Agriculture Organization of the United Nations
FRBR	Functional Requirements for Bibliographic Records
HTTP	Hypertext Transfer Protocol
IFLA	International Federation of Library Associations and Institutions
ISSN	International Standard Serial Number
LOD	Linked Open Data
LODE-BD	Linked Open Data Enabled Bibliographical Data
LRM	Library Reference Model
M2B	Meaningful Bibliographic Metadata
MIME	Media Type
MODS	Metadata Object Description Schema
ODRL	Open Digital Rights Language
RDF	Resource Description Framework
URI	Uniform Resource Identifier
USDA	United States Department of Agriculture
VOA3R	Virtual Open Access Agriculture & Aquaculture Repository
W3C	World Wide Web Consortium
XML	Extensible Markup Language

Executive summary

The Linked Open Data Enabled Bibliographical Data (LODE-BD) practical guide aims to support the selection of appropriate encoding strategies for producing meaningful LOD-BD (directly or indirectly). The guide provides recommendations applicable for structured data, describing bibliographic resources such as articles, monographs, theses, conference papers, presentation materials, research reports, learning objects, in print or electronic format. It also considers the inclusion of metadata that describes research datasets in a bibliographic service. The core component of LODE-BD contains a set of recommended decision trees for common properties used in describing a bibliographic resource instance. Each decision tree is delivered with various acting points and the matching encoding suggestions. The full range of options presented by this guide will enable data providers to make their choices according to their development stages, internal data structures, and the reality of their practices.

This practical guide is the latest version of the LODE-BD which was initially issued in 2011 and updated in 2015 (LODE-BD 2.0) with major changes, including a crosswalk of metadata terms used in LODE-BD and Schema.org (see Table 4 and Appendix). In this new version, authors have included metadata describing research data resources, based on the experience of the Food and Agriculture Organization of the United Nations' (FAO) International System for Agricultural Science and Technology (AGRIS) pilot project which integrates research datasets metadata from the United States Department of Agriculture (USDA), and guidance from Data Catalog Vocabulary (DCAT)-Version 2, a World Wide Web Consortium (W3C) Recommendation released in February 2020.

Background

The world's population is growing, the climate is changing, and people are increasingly migrating from rural areas to cities. To provide everyone with enough safe and nutritious food, global food and agriculture systems will need to transform. Research and information about food and agriculture is crucial for this transformation to happen sustainably, in a way that safeguards both livelihoods and the environment.

For decades, FAO has promoted the exchange of scientific research and technical knowledge related to all aspects of food and agriculture through AGRIS [1], which FAO and its member countries started in 1974. AGRIS was thought of as a one-stop access point for agricultural research, innovation and extension publications. At the beginning of the 1980s, a multi-lingual, collaborative thesaurus for agricultural concepts, AGROVOC [2], was developed to connect publications that were about identical or similar topics. With the advent of the Internet, and its rapid development in the 1990s, AGRIS partners became capable of publishing their own metadata. To enable the AGRIS network to share data without being tied to any internal data standard, FAO developed the AGRIS application profile.

With the paradigm of LOD and emerging technologies in the 21st Century, it has become a general strategy to liberate data from their silos that are framed by proprietary database schemas. However, simply transforming database schemas into Resource Description Frameworks (RDF) does not create Linked Data. There is a chance that data will get stuck at the 4th star in the 5-star classification that is promoted by Tim Berners-Lee (Berners-Lee, 2006). Automatic links between RDF triple stores on the web need to be able to be created; otherwise there is a risk that RDF silos will be created. The easiest way to facilitate automatic linking between datasets is the use of standards, including standard vocabularies for describing data/metadata elements and standard vocabularies for indicating values in the production of bibliographical data (metadata on document-like objects and beyond).

The idea of assisting information professionals in deciding what metadata terms and what standard vocabularies to use when encoding existing bibliographic data for the purpose of exchanging and sharing across data providers was born under the umbrella of Virtual Open Access Agriculture & Aquaculture Repository (VOA3R), an European research consortium project for sharing scientific and scholarly research related to agriculture, food, and environment, conducted during the early 2010s.

The VOA3R Federation was composed of 17 institutions from 13 countries which contributed bibliographic data to eight open repositories. The general objective of the VOA3R project was to improve the spread of European agriculture and aquaculture research results by using an innovative approach to sharing open access research products by connecting libraries, archives and other publication systems, while providing advanced search interfaces that include specific aspects of research work (methods, variables, measures, instruments, techniques, etc.) of each particular domain. From the initial idea of exchanging metadata using two different application profiles – the Dublin Core™ Metadata Element Set (DC) and the Metadata Object Description Schema (MODS) – the Federation expanded to include data harvesting. The VOA3R Federation also decided that the metadata encoding guidance would allow data providers to encode bibliographic data using properties from standardised namespaces, to use well-established authority data and controlled vocabularies that were available as Linked Data in agriculture and aquaculture, to publish data in RDF triples, and to submit the dataset to VOA3R. In doing so, VOA3R would act both as a service provider, enhancing the dissemination channel and accessibility of open access documents, and as an advocate to exchange and publish bibliographic data in RDF, to facilitate the use of Linked Data in agriculture and aquaculture. All these decisions led to the efforts of the creation of LODE-BD.

LODE-BD was initially issued in 2011 and updated in 2015 (LODE-BD 2.0) with a new crosswalk to Schema.org vocabulary which was founded by several search engines in 2011. Since the recommendations of LODE-BD 2.0 were based on DCMI Metadata Terms (DCTERMS) 2012, the LODE-BD version 3.0 is modified based on the DCTERMS' current version released in January 2020. As the movement of open research data has become more mainstream, LODE-BD 3.0 has been generated to include metadata describing research data resources, based on the experience of FAO's AGRIS pilot project of integrating research datasets metadata from USDA, which was successfully conducted in 2019, and the guidance from Data Catalog Vocabulary (DCAT) Version 2, a W3C Recommendation released in February 2020. The LODE-BD guide also includes properties from dc, dcterms, bibo, agls, eprint, prov and dcat namespaces.

In spite of the fact that the practical guide is geared towards the agriculture and aquaculture sectors through the VOA3R project, the LODE-BD guide is useful for any type of bibliographical data describing bibliographic resources in any subject domain.

1 http://agris.fao.org
2 http://www.fao.org/agrovoc/
3 W3C LD Glossary https://www.w3.org/TR/ld-glossary/#x5-star-linked-open-data
4 Databases for the storage and retrieval of RDF triples

1 Introduction

Introduction

With advances in the Internet, and a move towards open and Linked Data, the traditional approach of sharing data within silos seems to have reached its end. From governments and international organisations, to local cities and institutions, there is a widespread effort to open up and interlink data. This LODE-BD guide aims at providing bibliographic data providers of open repositories with a set of recommendations that will support the selection of appropriate encoding strategies for producing meaningful LOD-enabled bibliographical data. It follows the Linked Data principles stated by Tim Berners-Lee in his design note regarding the Semantic Web architecture, referring to a set of best practices for publishing, sharing, and interlinking structured data on the Web. The key principles that Linked Data builds on are: use URIs as names for things; use HTTP URIs so that people can look up those names; when someone looks up a URI, provide useful information, using the standards (RDF, SPARQL); and include links to other URIs so that they can discover more things (Berners-Lee, 2006).

1.1 Purpose of the LODE-BD guide

In the bibliographic universe there is a clear paradigm shift from fixed records to re-combinable metadata statements in the Semantic Web. For anyone who is contributing to an open bibliographic data repository as a data provider or service provider, the processes and strategies of providing data as Linked Data are practical issues. Guidelines and recommendations on what standards to follow and how to prepare LOD-ready metadata are essential.

There seems to be no one-size-fits-all approach because a great number of metadata-related standards have been developed during the last three decades. They have been created and maintained by different communities for specific purposes to guide the design, creation, and implementation of data structures, data values, data contents, and data exchanges in certain communities. The operational metadata standards for data structures form a whole spectrum, ranging from independent ones (which do not reuse any metadata terms from a known namespace) to integrated ones (which fully employ and incorporate existing metadata terms from other namespaces, usually seen in newly-developed metadata application profiles and ontologies). Decisions regarding what standard(s) to adopt will directly impact the degree of LOD-readiness of the bibliographic data.

The approach of employing well-accepted metadata element sets and value vocabularies has already shown great benefits and potentials in terms of resource discovery, data reuse, data sharing, and the creation of new content based on Linked Data. However, deciding to take this approach is only the first step for the data providers and service providers of an open bibliographic repository. In the context of producing LOD-enabled bibliographical data, data and service providers are likely to have many specific questions related to the encoding strategies, for example:

- What metadata standard(s) should be followed in order to publish any bibliographic data as Linked Data?
- What is the minimal set of properties that a bibliographic dataset should include to ensure meaningful data sharing?
- Is there any metadata model or application profile that can be directly adopted for producing bibliographical data (especially from a local database)?
- If the controlled vocabulary that has been used is available as Linked Data, what kind of values should be exchanged through the repository, specifically, the literal form representing a concept or the URI identifying the concept?
- How should data be encoded to move from a local database to a Linked Data dataset?

This guide was born with the purpose of assisting data providers in selecting appropriate encoding strategies for producing LOD-enabled bibliographical data (directly or indirectly).

1.2 Questions addressed

Once a data provider has decided to publish a bibliographical database as Linked Data, there are important components that should be considered, including:

a. What kinds of entities and relationships are involved in describing and accessing bibliographic resources?

LODE-BD believes that a conceptual model would help to establish an overall picture of involving entities and relationships in bibliographic descriptions. In a broader context, the use of a similar conceptual model among data providers should also help foster a common understanding of the involving data models. Thus, LODE-BD uses a simple conceptual model based on three entities: resource, agent and thema. Major relations can be identified between a resource instance (e.g. an article or a report) and the agent(s) (e.g. a personal author or a research team) that are responsible for the creation of the content and the dissemination of the resource, as well as the thema(s) (e.g. the subjects or topics of an article). The model, entitled Meaningful Bibliographic Metadata (M2B), provides sufficient capabilities for data providers to present their content (such as in document repositories and library catalogues) for sharing in the traditional environment (before the Semantic Web was created), or transferring to the Linked Data environment (for further explanation of M2B, see Section 2.1).

b. What properties should be considered for publishing meaningful/useful LOD-ready bibliographic data?

In the Linked Data context, any data provider can expose anything contained in its local database. However, in the case of bibliographical data, standardised types of information should be considered in order to maximise the impact of sharing and connecting the data. LODE-BD has built its recommendations on nine groups of common properties for describing bibliographic resources (see Section 2.2). These include specific best practice recommendations for about two-dozen properties used for describing a bibliographic resource, as well as an additional two sets of properties for describing relations between bibliographic resources or between agents.

c. What metadata terms are appropriate in any given property when producing LOD-ready bibliographic data from a local database?

LODE-BD has selected a number of well-accepted and widely used metadata/vocabularies and used their metadata terms in the recommendations. All metadata terms used in the recommendations are included in a crosswalk table (see Section 2.3, Table 4). Flowcharts are used to present individualised decision trees, which provide adjustable decision processes to data providers for their situation, when selecting metadata terms (see Chapter 3).

The comprehension of all of the components detailed in LODE-BD should enable a data provider to carry out a selection process of the metadata terms that fit their bibliographic data requirements. Most of the terms also apply to the metadata describing datasets. If there are specific characteristics in describing dataset resources, additional metadata terms will be provided.

1.3 The LODE-BD roadmap

The LODE-BD guide is presented as a whole package, encompassing the important components that a data provider may encounter when deciding to produce sharable LOD-ready structured data describing bibliographic resources (such as articles, monographs, theses, conference papers, presentation material, research reports, and learning objects, in print or electronic format) from a local database. It also considers the integration of metadata that describes research datasets within a bibliographic service. In the future the recommendations may continue to extend to accommodate other kinds of information resources.

The following recommendations are included in Chapters 2 and 3:

- **Chapter 2.** General recommendations present nine groups of common properties identified by LODE-BD and the selected metadata terms to be used for describing bibliographic resources.
- **Chapter 3.** Decision trees demonstrate how to select recommended properties according to local needs.

Table 1
The roadmap of the LODE-BD practical guide

Part	Focus	Document explanation
1	**Introduction**	Purpose of this guide
1.1	Questions addressed	Three major questions addressed in the guide
2	**General recommendations**	
2.1	M2B – a conceptual model	Recommendations of a set of metadata properties and encoding vocabularies
2.2	Metadata terms overview	All the metadata terms used in LODE-BD
2.3	Metadata terms used in the LODE-BD recommendations	LODE-BD metadata terms presented in a crosswalk table
3	**Decision trees**	A set of recommended decision-making trees for common properties used in describing a bibliographic resource instance. Each decision tree is delivered in a flowchart with various acting points. At the end of a decision tree, a set of matching encoding suggestions is provided.
4	**Further reading**	Recommended references for decision-makers, and links to general procedures of publishing Linked Data and useful syntax guidelines.
5	**Standards used**	Details of the metadata vocabularies used in LODE-BD.
6	**References**	List of references used in LODE-BD
7	**Appendix**	Crosswalk to Schema.org

1.4 Explanation of terminology

Certain terminology has been applied throughout this guide. Short explanations are provided below.

a. Metadata terms and properties
"[metadata] elements", "[metadata] fields", and "attributes [of an entity]" have been widely used by professionals who are involved in creating, designing, and implementing metadata standards and who have prepared their data structure standards using extended markup language (XML) schema as the primary medium. Nevertheless, as represented by DCMI Metadata Terms (DCTERMS), the RDF terminology is now gaining momentum. DC™ elements and element refinements are all considered as "properties" where a property is "a specific aspect, characteristic, attribute, or relation used to describe resources" (Johnston, 2005). The term "properties" of resources are used in this LODE-BD guide, for the characteristics that a resource may have, such as a title, publisher, subject, rights, etc. For example, "rights" is considered as a property of a resource:

Property: **rights**

Yet, there could be various levels of granularity and several corresponding ways that this property can be defined in different metadata element set specifications, e.g. for Property: rights

in dc /elements/1.1/ namespace: **dc:rights**

in dc /term/ namespace: **dcterms:rights**

Therefore, LODE-BD uses "metadata term" to refer to a specific element formally defined by a metadata namespace, no matter whether it is referred as an "element", "RDF property", or "term". For this example, property "rights" can be presented as metadata terms from different namespaces. The prefix of a term represents its namespace and this metadata term inherits the definition, constraints, and usages defined in that specification.

Metadata term: **dc:rights**

Metadata term: **dcterms:rights**

b. String and URI as values
In the LODE-BD guide, the words "string" and "URI" are used for the most commonly seen values in bibliographic data. They correspond to the terminology of RDF in the form of "literal" (typically a string of characters) and "non-literal".

Literal: "The most primitive value type represented in RDF, typically a string of characters. The content of a literal is not interpreted by RDF itself and may contain additional XML markup. Literals are distinguished from resources in that the RDF model does not permit literals to be the subject of a statement" (W3C, n.d.(a)).

Non-literal value: "A value which is a physical, digital or conceptual entity" (Powell *et al.*, 2007).

For example, "rice" is a concept included in the AGROVOC Thesaurus, with a preferred label (in English), "rice." When the thesaurus is published as Linked Data, the concept is considered as a resource and is given a unique URI, http://aims.fao.org/aos/agrovoc/c_6599. This means that a URI reference is used to identify this concept as a resource.

In this situation for the property: subject, the metadata terms for encoding this property include dc:subject and dcterms:subject. Because dcterms:subject is intended to be used with non-literal values (Dublin Core, 2020) as defined in the DCMI Abstract Model (Dublin Core, 2007), the value to be used associated with this metadata term should be the URI http://aims.fao.org/aos/agrovoc/c_6599 which represents the concept as a resource instead of "rice" or other language labels of the concept.

Based on the definition of these metadata terms, the following examples are provided:

Table 2
Examples of value types: string and URI

Metadata term	Value type	Example
dc:subject	string	rice
dcterms:subject	URI	http://aims.fao.org/aos/agrovoc/c_6599

3. [Bibliographic] resource
The term "resource" is used in the conceptual model to denote a general entity, the bibliographic resource. An instance of the bibliographic resource can be an article, monograph, thesis, conference paper, research report, presentation material, learning object, etc., regardless if it is in print or electronic format. In the flowcharts provided by the LODE-BD guide, the "resource" at the beginning oval box is an instance of the bibliographic resource (see Section 3.1.1).

2 General recommendations

2.1 M2B: a conceptual model

In order to have an overall picture and common understanding of involving entities and relationships in bibliographic descriptions, we have established a general conceptual model, Meaningful Bibliographic Metadata (M2B), following the International Federation of Library Associations and Institutions (IFLA) Functional Requirements for Bibliographic Records (FRBR)-family conceptual models which now have been consolidated into a unified IFLA Library Reference Model (IFLA LRM) of the bibliographic universe. While still using the Entity-Relationship modelling method that LRM used and some of the entities it defined, our M2B made enormous extensions and reconsiderations for the current recommendations.

M2B provides a high level of abstraction focusing on bibliographic resource entities (Figure 1). Three core entities are presented in the model: resource, agent, and thema. Major relations can be identified between a resource instance (e.g. an article or a report) and the agent(s) (e.g. a personal author or a research team) that are responsible for the creation of the content and the dissemination of the resource, as well as the thema(s) (i.e. the subjects or topics of an article). These led to the extended model presented in Figure 2, based on the implication of the general concept model with examples of possible relationships between and among the instances in different entities.

The models convey the following meanings (entity names are presented in italics):

- **Basic entities and their relationships.** The *resource* entity is the centre of every description here. The model does not exemplify the types of sub-entities, e.g. the sub-entities of resource would be various resource types. Relationships are established between the *resource* entity and two other major entitles: *agent* and *thema*.
- **Relationships between instances within the same entity.** Relationships between instances of an entity also exist. For example, a *resource* may be related to another *resource*. An *agent* may be related to another *agent*. Such relationships are demonstrated in the model.
- **Relationships between instances of different entities.** Relationships between any pair of instances vary and can be found at different levels. The sample relationships illustrated in Figure 2 are demonstrative and may apply at different levels of the bibliographic *resource* entity. For example, an *agent* may provide the funding for the creation of an original work, for the translation of a work, or the production of a new format of a translation.
- **Control of values.** Authority control is considered an important element of the model. The *agents*, regardless of their roles in relation to a *resource*, should be managed through name authority files. Concepts, topics, and geographic places as the *themas* of a *resource* should be controlled with value vocabularies. Although not emphasised in the model for the authority control of the titles of bibliographic resources given the context of this guide, it is also a logical step that resource uniform titles also be controlled.

More and more name authority files, controlled vocabularies, and resource datasets are becoming available as LOD. The model intentionally sets an extracted piece of the LOD cloud as the background for each entity, to remind the reader of reality.

The conceptual model holds the key for sharing common understanding of the important entities and relationships for bibliographic data. It can be used with different data models that have different implementation approaches.

Figure 1
The LODE-BD general concept model

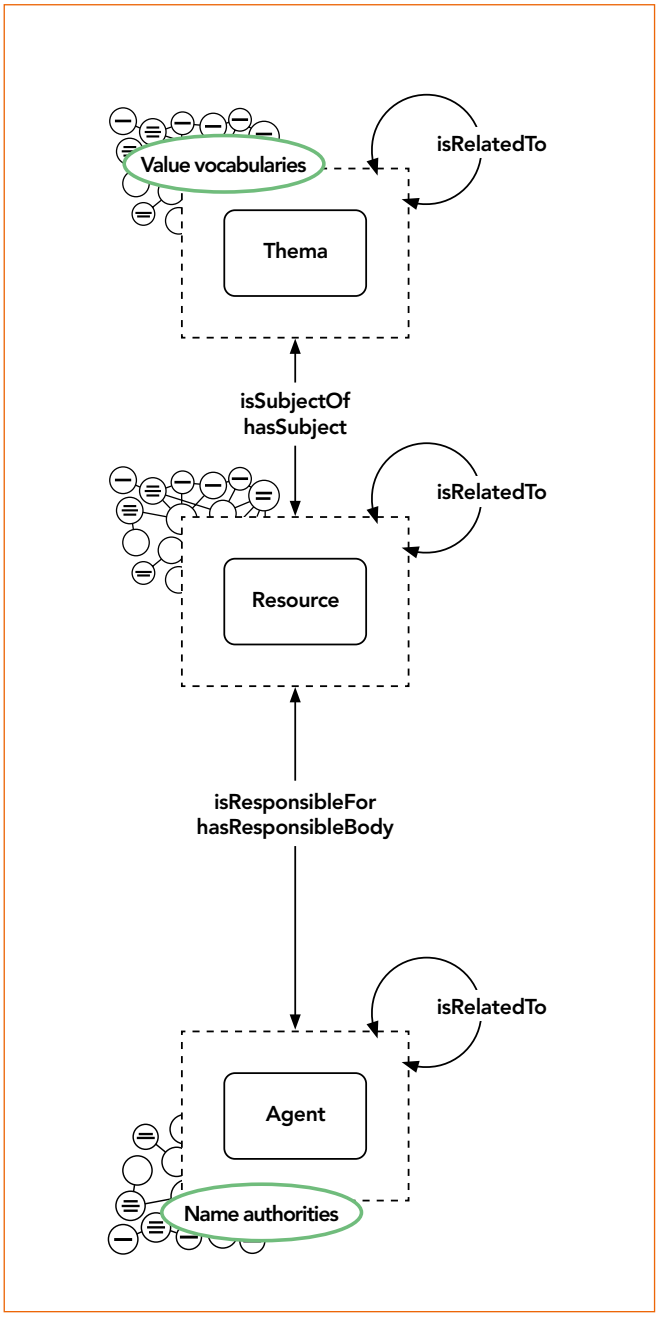
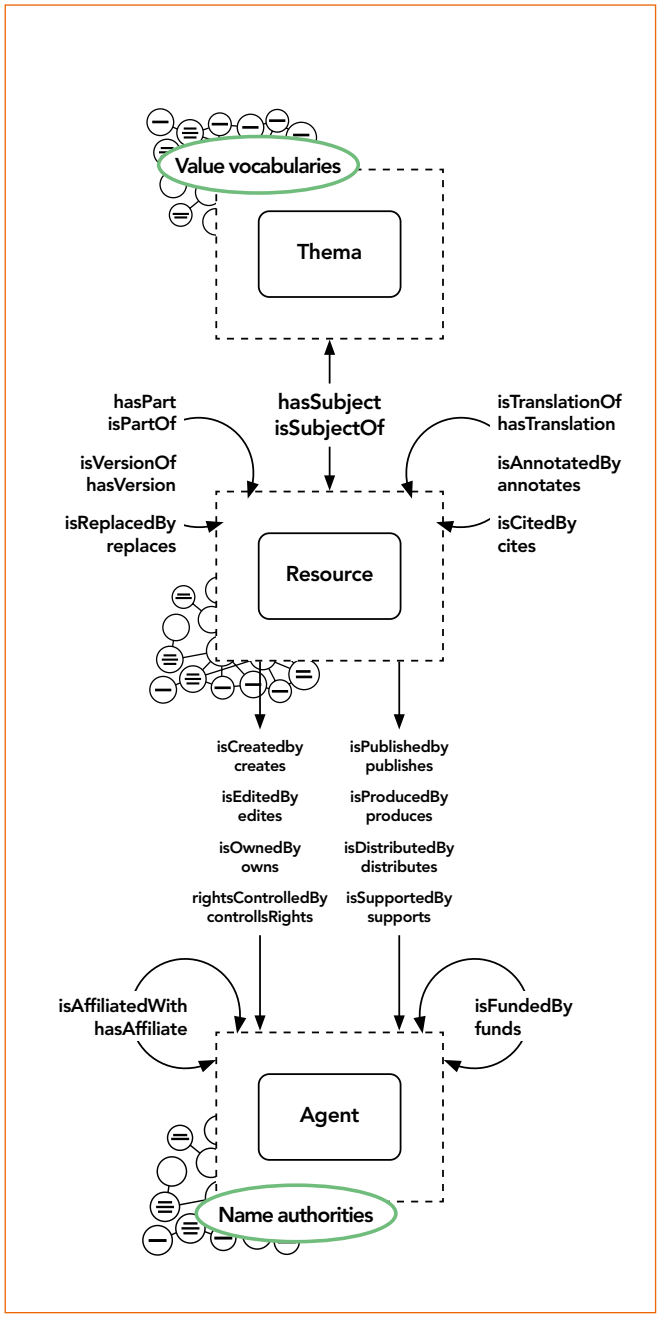


Figure 2
The implication of the general concept model in the LODE-BD v.1.1. case



2.2 Groups of common properties

Common properties for describing bibliographic resources are identified and grounded in nine groups based on our comprehensive studies of several open repositories. About two-dozen properties used for describing a bibliographic resource are included in groups one to eight. Two sets of properties for describing relations between bibliographic resources or between agents are included in group nine. In the guide, the word "resource" is used to represent bibliographic resources; a primary resource type to be described.

1. Title information
Title is one of the most important and relevant access points for any resource. The information is usually supplied through a number of properties including title, alternative title-(handling parallel title(s), translated title(s), transliterated title(s), etc.

2. Responsible body
This group contains the properties associated with any agent who is responsible for the creation and/or publication of the content of the resource, for example, the creator, contributor, and publisher or issuer of a resource.

3. Physical characteristics
Properties that describe the appearance and the characteristics of the physical form of a resource are placed into this group. They are: date, identifier, language, format, and edition/version.

4. Holding/location information
It is considered important for a resource to be located and obtained in the information exchange. Properties that record location and availability information are taken into account in this unique group.

5. Subject information
In contrast to the physical characteristics, the subject group embraces the properties that describe or otherwise help the discovery of what the resource is about or denotes, in the form of subject term, classification/category, freely assigned keyword and geographic term.

6. Description of content
Two major types of descriptions that focus on the content of the resource rather than the physical object are considered in this group: a) any representative description of the content, usually in the form of abstract, summary, note, and table of contents, and b) type or genre of the resource.

7. Intellectual property rights
Any property that deals with an aspect of intellectual property rights relating to access and use of a resource is included in this group, with special regard to rights, terms of use and access condition.

8. Usage
Properties that are related to the use of a resource, rather than the characteristics of the resource itself, are considered to belong to this group. Typical properties are: audience, literary indication, and education level.

9. Relation
This group has a different perspective for describing the resources from other groups that focus on describing the resource itself. Here, various relations between two resources or between two agents are the focus of description.

2.3

Metadata terms used in the LODE-BD guide: overview

These nine information groups are listed in Table 4, with the specific properties included in each group. Special attention should also be given to the additional recommendations on cardinality and value control (see Table 4, column 3 and 4). Table 4 comprises the following components in corresponding columns:

1. Groups of properties

2. Properties included in each group. Two special styles are used to signify the importance of the properties: two plus signs "++" for the mandatory property; one plus sign "+" for the highly recommended property in the context of bibliographic information exchange. The rest are recommended or optional.

3. Requirements of properties in the context of both non-analytical and analytical bibliographic descriptions, specified with (M)andatory; (H)ighly-(R)ecommended; (R)ecommended; and (O)ptional marked for either process.

4. Recommendation on the control of values, indicating: (n)ot controlled; should use a name authority or a controlled vocabulary; or should follow a syntax encoding rule.

The normative prefixes and namespaces included in this recommendation are detailed in Table 3.

Following the nine groups of recommended properties specified by M2B, the recommended metadata terms have been organised and presented as a crosswalk in Table 4. The semantics of the metadata terms (e.g. definition, usage, and relation with another property) defined by these specifications are inherited when a recommendation is made in a decision tree (see Chapter 3).

Table 3
A list of the prefix and namespaces used in LODE-BD (Refer to Chapter 5 Standards used).

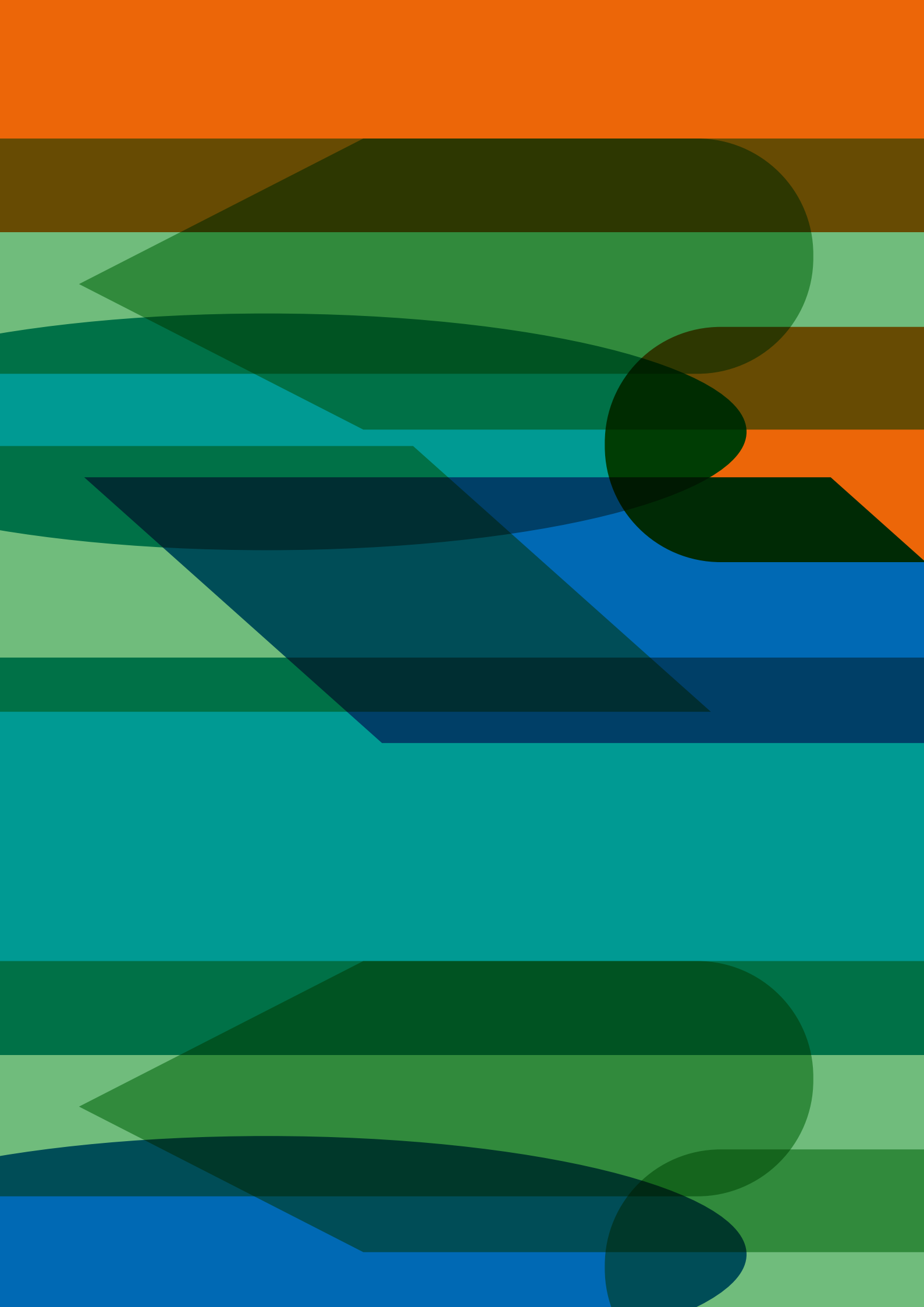
Prefix	Namespace	Standard
dc	http://purl.org/dc/elements/1.1/	Dublin Core Metadata Element Set
dcterms	http://purl.org/dc/terms/	DCMI Metadata Terms
bibo	http://purl.org/ontology/bibo/	Bibliographic Ontology
agls	http://www.agls.gov.au/agls/terms	AGLS Metadata Standard
eprint	http://purl.org/eprint/terms/	Eprints Terms
prov	http://www.w3.org/ns/prov#	PROV-O: The PROV Ontology
dcat	http://www.w3.org/ns/dcat#	Data Catalog Vocabulary

Table 4: Metadata terms used in the LODE-BD guide, grouped using the M2B model

LODE-BD group	M2B model recommendation				Metadata terms		
	Property	Requirement | M | HR | R | O |		Value control	General metadata terms		More specific metadata terms
		Non-analytical	Analytical		dc:-based	dcterms: -based	
1. Title information	title++	M	M	n	dc:title	dcterms: title	
	alternative title	O	O	n			dcterms:alternative
2. Responsible body	creator+	HR	HR	n or name authority (personal, corporate body, conference)	dc:creator	dcterms: creator	
	contributor	O	O	n or name authority	dc:contributor	dcterms: contributor	bibo:editor
	publisher/ issuer+	HR	R	n or name authority	dc:publisher	dcterms: publisher	bibo:issuer bibo:producer bibo:distributor bibo:owner
	[Other responsible body]	O	O	n or name authority	--	--	dcat:contactPoint dcat:qualifiedAttribution dcat:hadRole
3. Physical characteristics	date++	M	M	Syntax encoding rule	dc:date	dcterms:date	dcterms:created dcterms:dateAccepted dcterms:dateCopyrighted dcterms:dateSubmitted dcterms:modified dcterms:valid dcterms:available dcterms:issued
	identifier+	HR	HR	Syntax encoding rule	dc:identifier	dcterms: identifier	dcterms:bibliographicCitation bibo:asin bibo:coden bibo:doi bibo:eanucc13 bibo:eissn bibo:gtin14 bibo:handle bibo:isbn bibo:issn bibo:lccn bibo:oclcnum bibo:pmid bibo:sici bibo:upc bibo:uri
	language++	M	M	Controlled list	dc:language	dcterms: language	
	format/ medium+	HR	HR	Controlled list	dc:format	dcterms:format	dcterms:medium dcterms:extent
	edition / version	R	R	n			bibo:edition bibo:status
	source+	HR	R	n	dc:source	dcterms:source	bibo:pages bibo:pageStart bibo:pageEnd bibo:section bibo:volume bibo:issue bibo:chapter
	[Other physical characteristics]	O	O	n	--	--	dcat:landingPage dcat:accessURL

LODE-BD group	M2B model recommendation				Metadata terms		
4. Holding/ location information	location++	M	M	n or rule [Holding unit names may be managed through a controlled list]	dc:description	dcterms: description	agls:availability bibo:locator
	[Other locations]	O	O	n			[properties in] dcat:Distribution
5. Subject information	subject term+	HR	HR	Controlled vocabulary	dc:subject	dcterms: subject	
	classification	O	O	Controlled vocabulary, classification system			dcat:theme
	[freely assigned] keyword	R	R	n			dcat:keyword
	geographic term	O	O	Controlled vocabulary	dc:coverage	dcterms: coverage	dcterms:spatial dcterms:temporal
6. Description of content	description/abstract (or note/ summary/ table of contents)	HR	HR	n	dc:description	dcterms: description	dcterms:abstract dcterms:tableOfContent
	type/form/ genre	R	R	Controlled vocabulary	dc:type	dcterms: type	
7. Intellectual property rights	rights+ term of use/ access condition	R	R	n [Rights holders may be managed through name authorities]	dc:rights	dcterms: rights	dcterms:rightsHolder dcterms:accessRights dcterms:license
8. Usage	audience	O	O	Controlled list		dcterms: audience	dcterms:mediator
	instructional method	O	O	n or controlled list		dcterms:instructionalMethod	
	literary indication	O	O	Controlled list	dc:description	dcterms: description	
	education level	O	O	Controlled list		dcterms:educationLevel	
9. Relation	[relation between resources]+	O	HR	Controlled resource IDs	dc:relation	dcterms: relation	dcterms:isVersionOf dcterms:hasVersion dcterms:isReplacedBy dcterms:replaces dcterms:isRequiredBy dcterms:requires dcterms:isPartOf dcterms:hasPart dcterms:isReferencedBy dcterms:references dcterms:isFormatOf dcterms:hasFormat dcterms:conformsTo bibo:translationOf bibo:annotates bibo:citedBy bibo:cites dcat:qualifiedRelation
	[relation between agents]	O	O	n or name authority			eprint:affiliatedInstitution eprint:grantNumber dcat:hadRole [properties in OpenAIRE Version 4, for Funding Reference])

Legend used:
++ mandatory property
+ highly recommended property
Requirement: (M)andatory; (H)ighly-(R)ecommended; (R)ecommended; and (O)ptional
Value control: (n)o control



3 Decision trees: recommendations for individual properties

3

Decision trees

To assist in the metadata term selection, this chapter provides decision trees for the properties included in each of the nine groups presented in the crosswalk table (see Section 2.3 Table 4). Starting from the property that describes a resource instance, each flowchart presents decision points and gives a step-by-step solution to a given problem of metadata encoding. At the end of each flowchart, there are alternative sets of metadata terms for selection. A data provider can highlight the decision path and mark the metadata terms to be used at the end.

The types of values associated with a metadata term may be two types*:

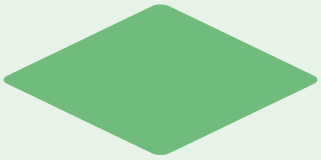
- literals (typically a string of characters; indicated by "string" in the flowcharts); or
- non-literals (a value which is a physical, digital or conceptual entity (Powell *et al.*, 2007); indicated by "URI" in the flowcharts), depending on the requirements expressed in the namespace.

***Note** that these are recommendations, not mandates. When a non-literal value is not possible or feasible, a literal value may be provided if it is relevant.

Text-based explanations corresponding to each of the flowcharts, with notes, steps, and examples, are also provided.

Figure 3: Flowchart symbols and meanings

A flowchart is a diagrammatic representation that uses standardised symbols to portray steps and processes involved in decision-making, with orders connected by flow lines with arrows. The basic shapes used in the figures follow flowchart conventions:

Name	Symbol	Description
Narrow oval		Beginning of a decision tree
Flow-line		Direction of logic flow
Dashed flow-line		Suggested direction of logic flow
Diamond		A decision to be made
Rectangle		A Process to be carried out
Parallelogram		An input or available information sources
Oval		End of a decision

3.1 Title information

Title is one of the most important and relevant access points for any resource. The information is usually supplied through a number of properties including title and alternative title (handling subtitle(s), parallel title(s), translated title(s), transliterated title(s)). Title information is essential in the description of a resource; therefore Figure 4 foresees title as a mandatory metadata property.

3.1.1 Title/alternative title

Relation with a resource being described: Resource has title.

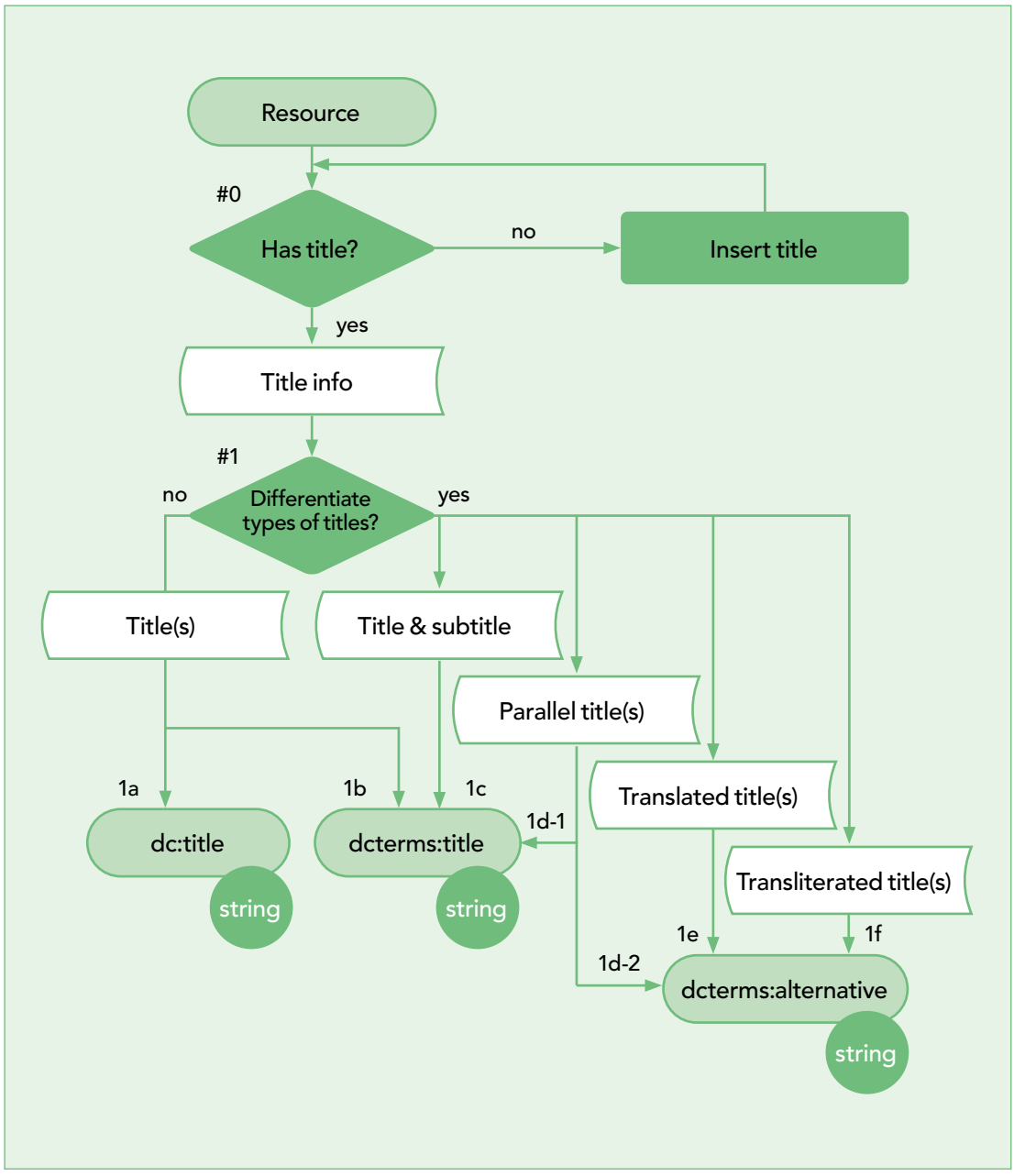


Figure 4: Title/alternative title decision tree

Note

- Values for this property are always text strings.
- Although not emphasised in this guide for the authority control of the titles of bibliographic resources given the context of this guide, it is a logical step that resource titles, especially uniform titles, are also controlled.

Table 5: Decision process with practical examples for title/alternative title

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	VALUE
#0	Has title?	No	Insert title and go back to #0			
		Yes	Continue to #1			
#1	Differentiate types of titles?	No	1a	String	dc:title	Solar radiation energy and its utilisation by Lucerne (Medicagosativa L.)
			1b	String	dcterms:title	On the state of man [world agricultural situation]
		Yes: title(s) and subtitle(s)	1c	String	dcterms:title	FAO yearbook of forest products, 1996-2000
		Yes: parallel title(s)	1d-1	String	dcterms:title	Annuaire des produits forestiers de la FAO, 1996-2000
			1d-2	String	dcterms: alternative	
		Yes: translated title(s)	1e	String	dcterms: alternative	Anuario de productos forestales de la FAO, 1996-2000 Working together for an International Alliance Against Hunger
		Yes: transliterated title(s)	1f	String	dcterms: alternative	Posly dobroj voli Prodovol'stvennoj i Sel'skokhozyajstvennoj Organizatsii Ob'edinennykh Natsij

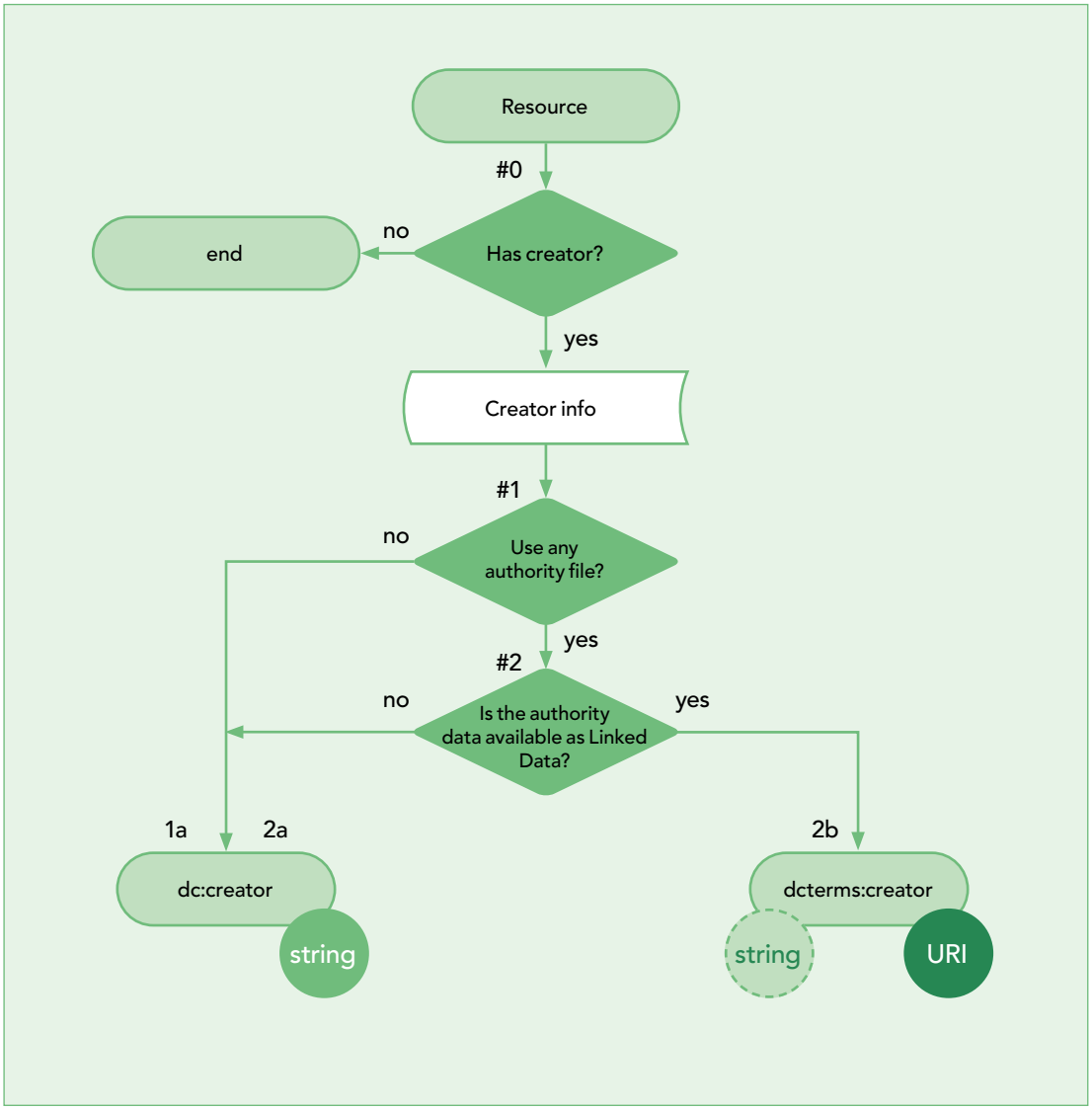
3.2 Responsible body

This group contains the properties associated with any agent who is responsible for the creation and/or publication of the content of the resource, for example, the creator, contributor, and publisher or issuer of a resource. Beyond these common properties, others are provided as additional responsible body information (see Section 3.2.4.)

3.2.1 Creator

Relation with a resource being described: Resource has creator.

Figure 5: Creator decision tree



Note It is always recommended that an authority file be used for the responsible body that has created the resource.

Table 6: Decision process with practical examples for creator

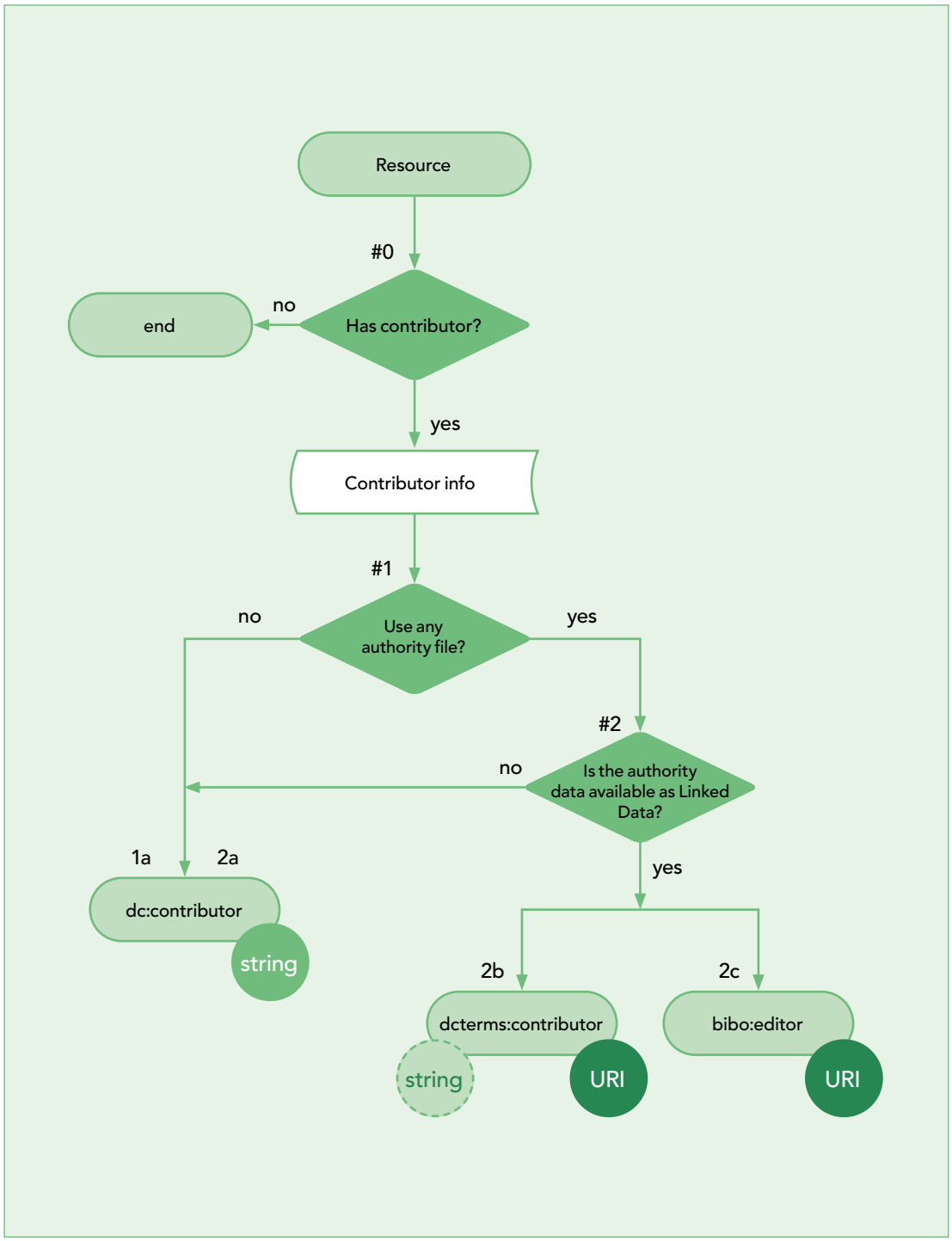
Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has creator?	No	End			
		Yes	Continue to #1			
#1	Use any authority file?	No	1a	String	dc:creator	[Unauthorised form]: • Tim Berners-Lee • Tim B-L • Timothy John Berners-Lee • FAO of the UN • FAO Council (78th Session) Nov. 24, 1980, Rome, Italy
		Yes	Go to #2			
#2	Is the authority data available as Linked Data?	No	2a	String	dc:creator	[Authorised form]: • Berners-Lee, Tim • Food and Agriculture Organization of the United Nations • FAO Council (Sess. 78 : 24 Nov 1980 : Rome, Italy) [values from an authority list]
		Yes	2b	URI or String (when URI is not available)	dcterms:creator	http://aims.fao.org/aos/corporate/c_1297* http://aims.fao.org/aos/conference/c_1842** http://viaf.org/viaf/85312226/#Berners-Lee,_Tim*** http://www.w3.org/People/Berners-Lee/card**** [URI of a responsible body] [A literal value that identifies the responsible body, if a URI is not possible or feasible.]

* A corporate body's URI, from the FAO Authority Description Concept Scheme
** A conference' URI, from the FAO Authority Description Concept Scheme
*** Tim Berners-Lee's URI, from the VIAF (Virtual International Authority File)
**** Tim Berners-Lee's URI http://www.w3.org/People/Berners-Lee/card#i , from Tim Berners-Lee's FOAF file .

3.2.2 Contributor

Relation with a resource being described: Resource has contributor.

Figure 6: Contributor decision tree



Note It is always recommended that an authority file be used for a responsible body that has contributed to the resource.

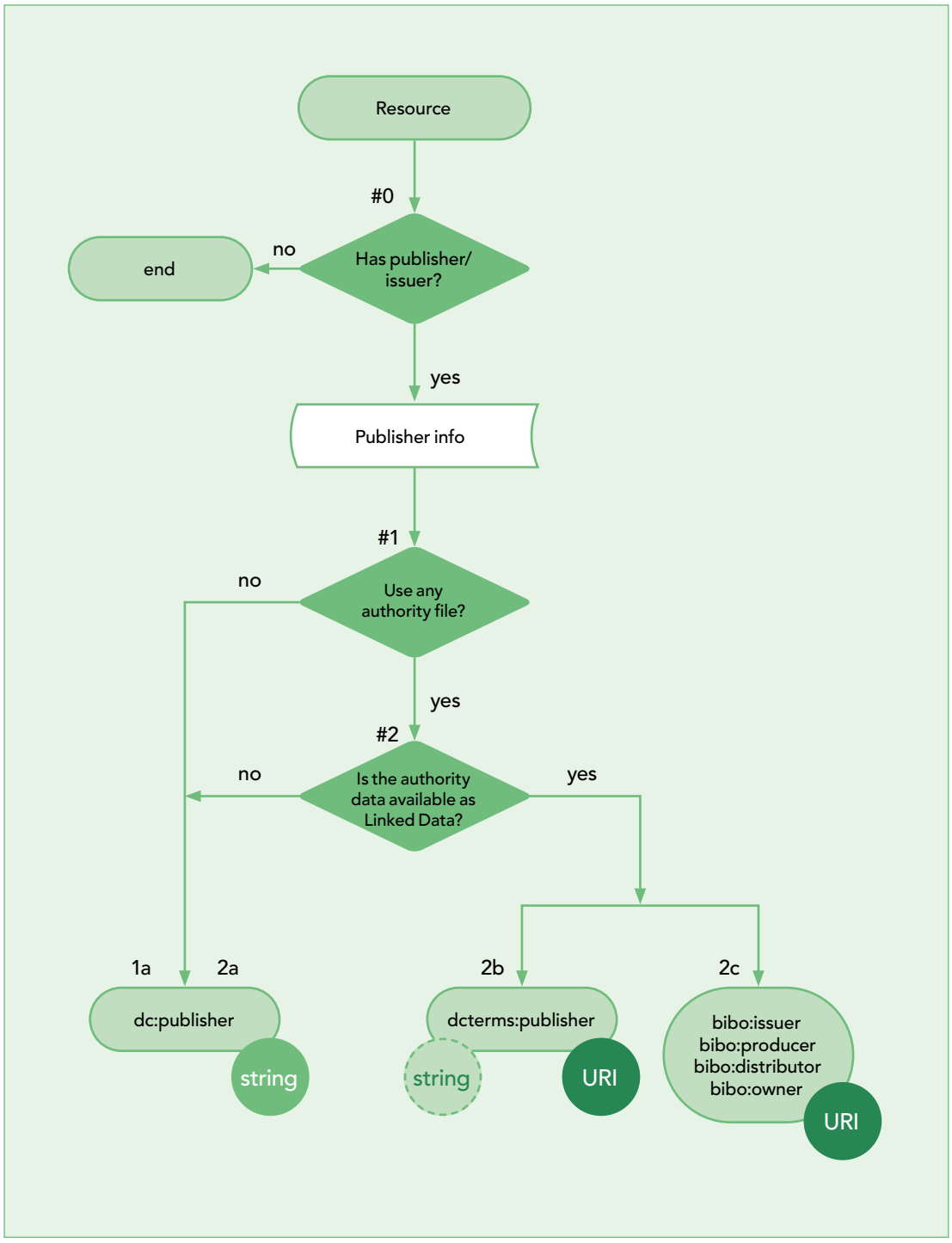
Table 7: Decision process with practical examples for contributor

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has contributor?	No	End			
		Yes	Continue to #1			
#1	Use any authority file?	No	1a	String	dc:contributor	[Unauthorised form]: • Tim Berners-Lee • Tim B-L • Timothy John Berners-Lee • FAO of the UN • FAO Council (78th Session) Nov. 24, 1980, Rome, Italy
		Yes	Go to #2			
#2	Is the authority data available as Linked Data?	No	2a	String	dc:contributor	[Authorised form]: • Berners-Lee, Tim • Food and Agriculture Organization of the United Nations • [values from a name authority]
		Yes	2b	URI or String (when URI is not available)	dcterms: contributor	[URI of a responsible body] [A literal value that identifies the responsible body, if a URI is not possible or feasible.]
			2c	URI	bibo:editor	

3.2.3 Publisher

Relation with a resource being described: Resource has publisher.

Figure 7: Publisher decision tree



Note
It is always recommended that an authority file be used for a responsible body that is responsible for publishing or producing the resource.

Table 8: Decision process with practical examples for Publisher

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has publisher/ issuer?	No	End			
		Yes	Continue to #1			
#1	Use any authority file?	No	1a	String	dc:publisher	[Un-authorised form]: • FAO Rome (Italy) • FAO • F.A.O. • FAO of the UN • FAO, Rome • Food and Agriculture Organization • F.A.O. of the U.N. [Authorised form]: • Food and Agriculture Organization of the United Nations
		Yes	Go to #2			
#2	Is the authority data available as Linked Data?	No	2a (See#1a) [Use authorised form from an authority file]			
		Yesw	2b	URI or String (when URI is not available)	dcterms: publisher	[URI of a responsible body] [A literal value that identifies the responsible body, if a URI is not possible or feasible.]
			2c	URI	bibo: issuer bibo:producer bibo:distributor bibo:owner	[URI of a responsible body]

3.2.4 Additional responsible body

These properties are useful for providing information to any responsible body that is known but does not match one of the above metadata terms (e.g. creator, publisher, contributor).

Metadata terms:

- dcat:contactPoint
  - Relevant contact information for the cataloged resource.
- prov:qualifiedAttribution
  - Link to an agent having some form of responsibility for the resource.
- dcat:hadRole
  - The function of an entity or agent with respect to another entity or resource.

Note

- It is always recommended that an authority file be used for the responsible body that has created the resource. Use of vCard is recommended (W3C, 2014).
- dcat:hadRole could be used on the prov:Attribution to capture the responsibility of the agent with respect to the resource.

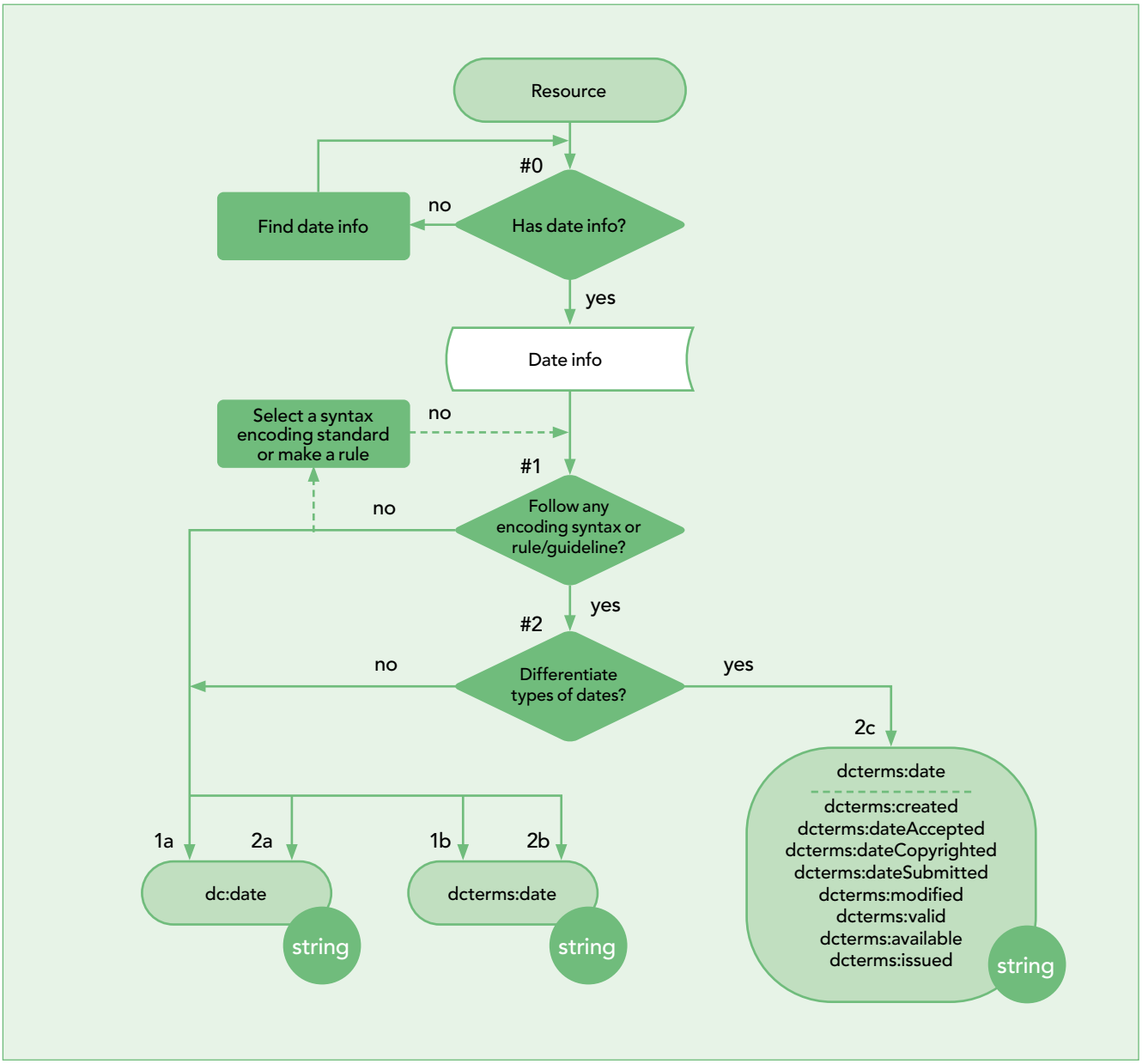
3.3 Physical characteristics

Properties that describe the appearance and the characteristics of the physical form of a resource are placed into this group. They are: date, identifier, language, format/medium, edition/version, and source.

3.3.1 Date

Relation with a resource being described: Resource has date.
Date is considered essential information in the description of a resource; therefore Figure 8 foresees date as a mandatory property.

Figure 8: Date decision tree



Note Recommended best practice is to follow an encoding syntax, such as that defined by the W3C date and time format profile of ISO 8601 (1988, 2019) or a published profile of the ISO standard, such as the W3C Note on Date and Time Formats (1997) or the Extended Date/Time Format Specification (LC, 2019).

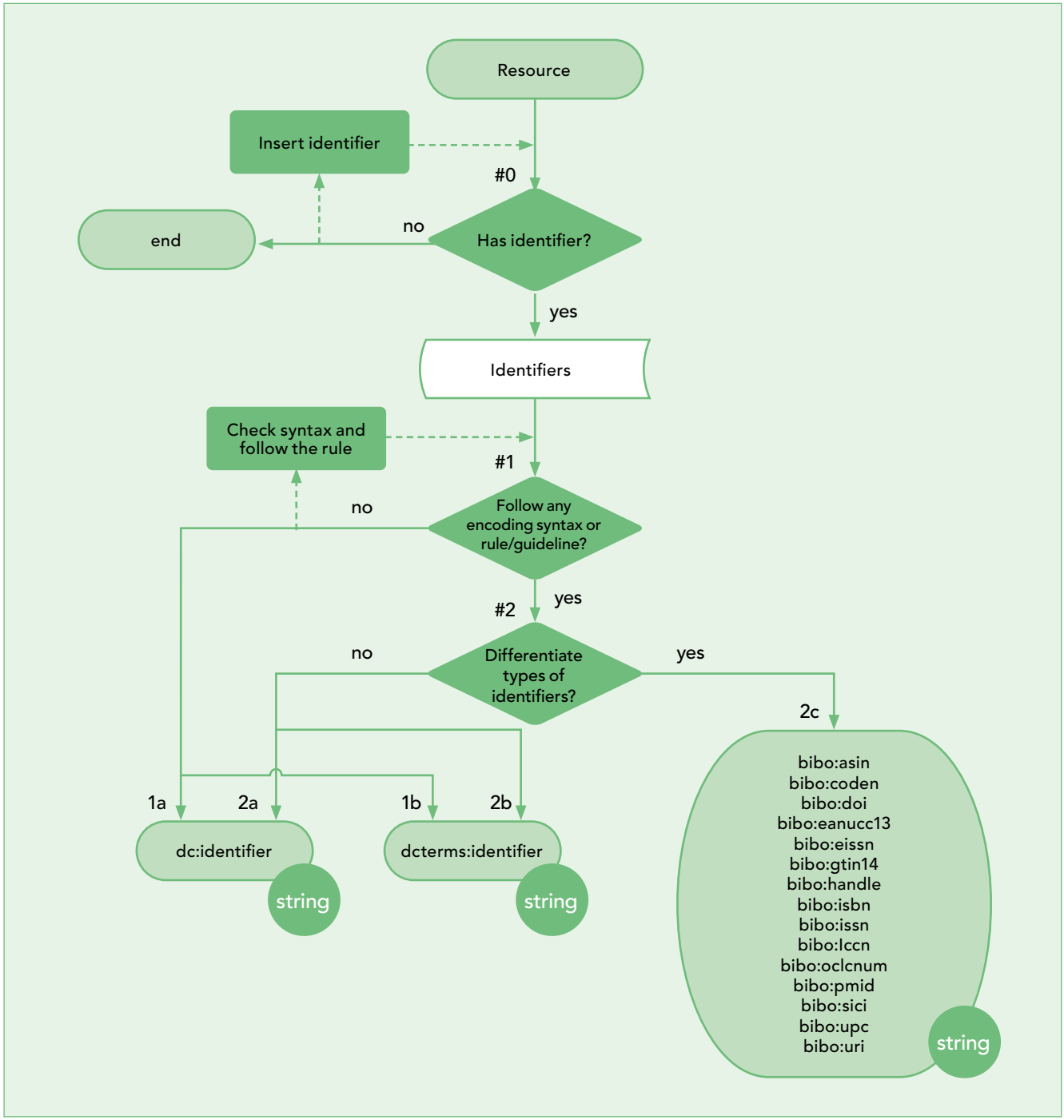
Table 9: Decision process with practical examples for date

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has date info?	No	Find date info and go back to #0			
		Yes	Continue to #1			
#1	Follow any encoding syntax or rule/ guideline?	Yes	Continue to #2			
		No	1a	String	dc:date	[198?] [1996] [1997?] 1968-2006 7 Jul 1989 7 July 1989 7-July-1989 Jul 1989 1989 Jul Jan-Feb 1997 1-5 Feb 1997 Spr 1997 20 Mar - 15 Apr 1995
			1b	String	dcterms:date	[see all examples above]
#2	Differentiate type of dates?	No	2a	String	dc:date	1997 1997-07 1997-07-16 1997-07-16T19:20+01:00 1997-07-16T19:20:30+01:00 1997-07-16T19:20:30.45+01:00* [values following an encoding syntax]
			2b	String	dcterms:date	
		Yes	2c	String	dcterms:date	
					dcterms:created	
					dcterms:dateAccepted	
					dcterms:dateCopyrighted	
					dcterms:dateSubmitted	
					dcterms:modified	
					dcterms:valid	
					dcterms:available	
					dcterms:issued	
* W3C date and time format (W3C, 1997).						

3.3.2 Identifier

Relation with a resource being described: Resource has identifier.

Figure 9: Identifier decision tree



Note

- It is always recommended that a resource has an identifier or identifiers.
- Established codes for identifiers (universal or local) should be used for any kind of identifiers. It is always recommended to check the syntax, and follow or create a rule/guideline when handling identifiers.

  In the bibliographic descriptions, a resource is always represented by a unique ID. This ID may be: a) locally assigned (or be local temporarily), b) be the same as its global recognisable identifiers, such as a URI, or c) contain the string that is from a universal identifier such as an International Standard Serial Number (ISSN) or a Digital Object Identifier (DOI) (for examples of the IDs, see Box 1). In this guide, such a unique ID is assumed to each resource being described, at the beginning of a decision tree.
- In addition to this unique ID, there are identifiers that are assigned to the original resource within the domains of various systems such as ISBN, DOI, ISSN, etc. The decision tree presented here is about those identifiers, even though one of the identifiers is the same as the unique ID of the resource being described (See Box 2).
- DCterms also provides a sub-property of identifiers as dcterms:bibliographicCitation which allows using literal.

Box 1: Examples of IDs

1. <info:doi/10.1134/S0003683806040089> a bibo:Article

2. <http://www.nytimes.com/2007/07/09/us/09cnd-penn.html> a bibo:Article;
<http://www.amazon.com/dp/026256212X> a bibo:Document
<urn:isbn:23983498> a bibo:Book
<urn:issn:23346587> a bibo:Journal
<http://www.w3.org/2007/Talks/0619-Nancy-IH/> a cc:Work, bibo:Slideshow

3. <doi:10.1045/july2015-bide>a bibo:Article

Examples 1 and 2 are from Bibliographic Ontology (BIBO)

Table 10: Decision process with practical examples for identifier

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has identifier?	No	End but recommended to insert an identifier			
		Yes	Continue to #1			
#1	Follow any encoding syntax, /rule /guideline?	No	1a	String	dc:identifier	http://www.ukoln.ac.uk/
						urn:ietf:rfc:1766
			1b	String	dcterms:identifier	http://www.ukoln.ac.uk/
						urn:ietf:rfc:1766
		Yes	Continue to #2			
#2	Differentiate types of identifiers?	No	2a	String	dc:identifier	http://www.ukoln.ac.uk/
						urn:ietf:rfc:1766
			2b	String	dcterms:identifier	http://www.ukoln.ac.uk/
						urn:ietf:rfc:1766
		Yes	2c	String	bibo:asin	020530902X
					bibo:coden	66HYAL
					bibo:doi	doi:10.1109/ISSTA.2002.1048560
					bibo:eanucc13	0123456789012
					bibo:eissn	0378-5955
					bibo:gtin14	00012345600012
					bibo:handle	https://hdl.handle.net/20.1000/100
					bibo:isbn	9-788175-257665
						9788175257665
					bibo:issn	0317-8471
					bibo:lccn	79051955
					bibo:oclcnum	ocn123456789
					bibo:pmid	20346624
					bibo:sici	0095-4403(199502/03)21:3<12:WATIIB>2.0.TX;2-J
					bibo:upc	5778400002
					bibo:uri	http://example.org/absolute/URI/with/absolute/path/to/resource.txt
						ftp://example.org/resource.txt

Box 2.
Identifiers assigned to original resources within the domains of various systems

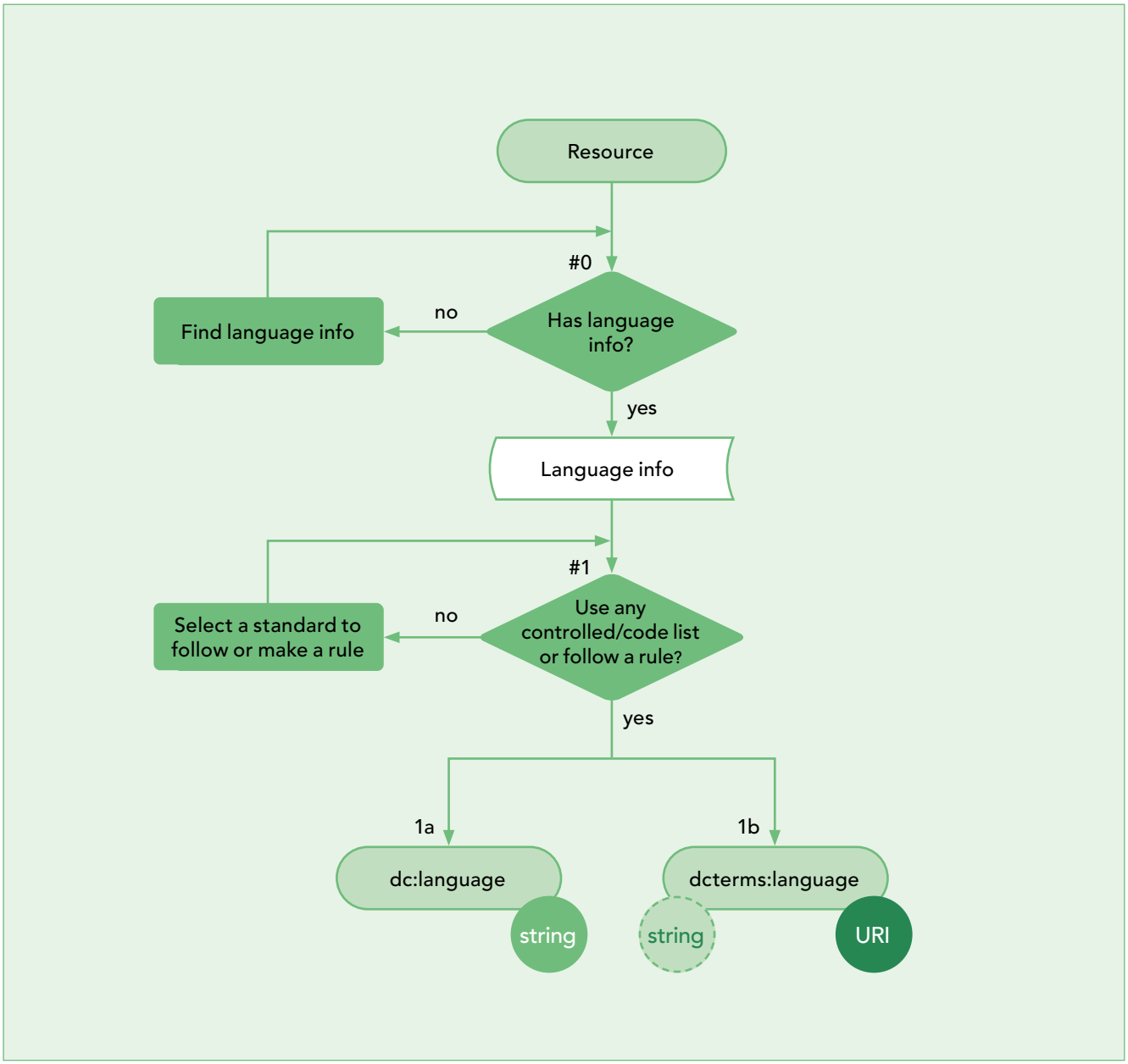
Code used	Related information	Reference
asin	Amazon Standard Identification Number (ASIN)	(Wikipedia, 2020a)
coden	CODEN, a six-character, alphanumeric bibliographic code	(Wikipedia, 2020b)
doi	Digital Object Identifier (DOI)	(DOI, 2020)
Ean/ucc13	Global Trade Item Number (GTIN) -13 (EAN/UCC-13) [a 13-digit number used predominately outside of North America]	(GTIN, 2020)
eissn	Electronic ISSN (International Standard Serial Number)	(Wikipedia, 2020d)
gtin14	Global Trade Item Number (GTIN)-14 [a 14-digit number used to identify trade items at various packaging levels]	(GTIN, 2020)
handle	Handle.Net Registry (HNR), a proprietary registry assigning persistent identifiers, or handles, to information resources	(Wikipedia, 2020h)
isbn	International Standard Book Number	(Wikipedia, 2020c)
issn	International Standard Serial Number	(Wikipedia, 2020d)
lccn	Library of Congress Control Number	(LC, 2020; OCLC, 2020)
oclcnum	OCLC control number	(OCLC. 2020)
pmid	The unique identifier number used in PubMed.	(NLM, 2020)
sici	Serial Item and Contribution Identifier	(Wikipedia, 2020e)
upc	Universal Product Code	(Wikipedia, 2020g)
uri	Uniform Resource Identifier	(Wikipedia, 2020f)

3.3.3 Language

Relation with a resource being described: Resource has language information.

Language is considered essential information in the description of a resource; therefore Figure 10 foresees language as a mandatory property.

Figure 10: Language decision tree



Note
Recommended best practice is to use an encoding scheme, such as the three-letter code (ISO639-2) or the two-letter code (ISO639-1) (LC, 2017) used in the Table 11. Literal value consisting of an BCP47 - Tags for Identifying Languages (2009) language tag may be used when needed.

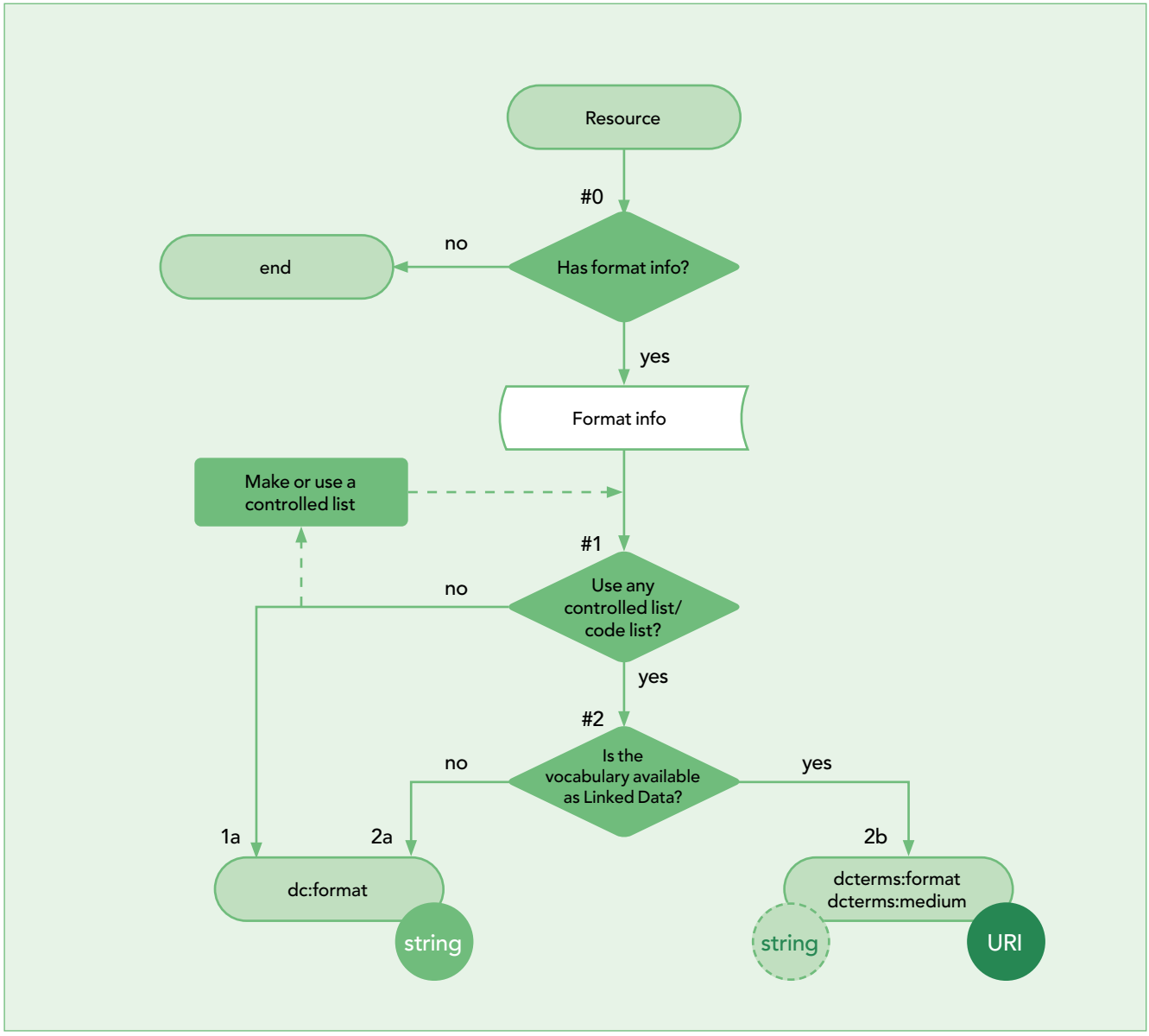
Table 11: Decision process with practical examples for language

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has language info?	No	Find language info and go back to #0			
		Yes	Continue to #1			
#1	Use any controlled list /code list or follow a rule?	No	Go back to #1			
		Yes	1a	String	dc:language	cat
						ca
			1b	URI or String (when URI is not available)	dcterms:language	cat ca [an identifier from ISO639 representing a language] [a literal value consisting of an IETF language tag]

3.3.4 Format/medium

Relation with a resource being described: Resource has format.

Figure 11: Format/medium decision tree



Note

- It is always recommended that a controlled vocabulary be used for your collection when describing "format" [e.g., the list of Media Types (previously known as MIME type)].
- It is also recommended that a controlled vocabulary be used for your collection when using dcterms:medium.
- Because dcterms:medium has the definition of material or physical carrier of the resource and may be used only for the description of physical resources (e.g. paper, canvas, or DVD), the Media Types should NOT be used for these values. Also, for dcterms:medium, the property is recommended to be used with non-literal values (Rühle *et al.*, n.d.).
- A dcterms:extent allows describing the size or duration of the resource. It is recommended to specify the file size in megabytes and duration in ISO 8601 format (Dublin Core, 2020).

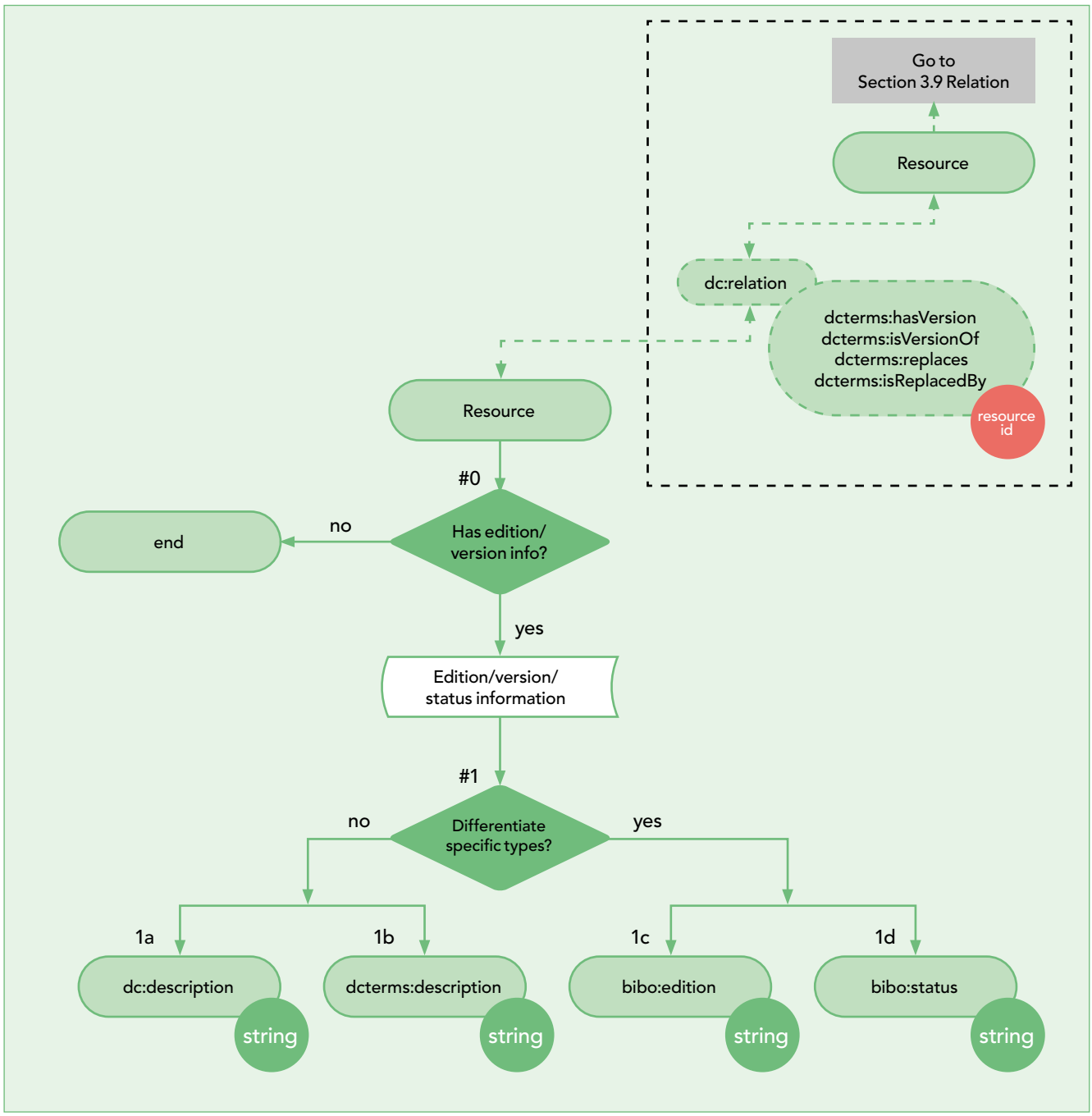
Table 12: Decision process with practical examples for format/medium

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has format info?	No	End			
		Yes	Continue to #1			
#1	Use any controlled list or code list?	No	1a	String	dc:format	html
		Yes	Continue to #2			
#2	Is the controlled vocabulary available as Linked Data?	No	2a	String	dc:format	text/html
		Yes	2b	URI or String (when URI is not available)	dcterms:format	image/jpeg [values from a controlled vocabulary]
					dcterms:medium	_:oilOnWood

3.3.5 Edition/version

Relation with a resource being described: Resource has edition/version/status.

Figure 12: Edition/version decision tree



Note

- When an edition or version of a resource is to be described, the relation between a resource and its related version(s) should also be described. In Figure 12, a dash-lined box signifies such relation(s) and points to Section 3.9, in this guide.
- The main body of Figure 12 only focuses on the description of edition or version as a part of the physical characteristics of a resource. For describing relations between different versions of resources, see Section 3.9.1.

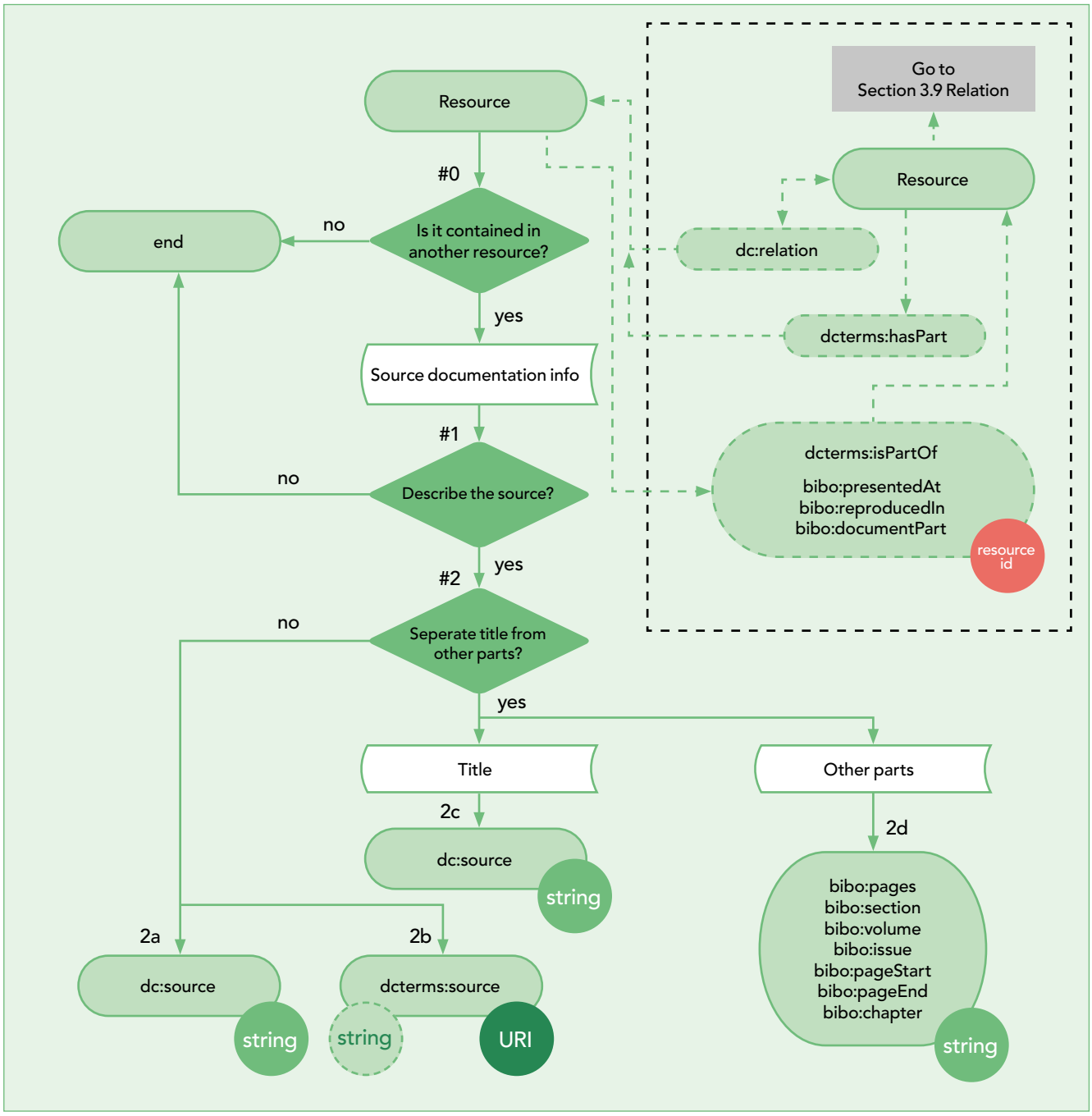
Table 13 Decision process with practical examples for edition/version

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has edition version info?	No	End			
		Yes	Continue to #1			
#1	Differentiate specific types?	No	1a	String	dc:description	2nd ed.
			1b	String	dcterms:description	2nd ed.
		Yes	1c	String	bibo:edition	2nd ed.
			1d	String	bibo:status	Final

3.3.6 Source

Relation with a resource being described: Resource has source.

Figure 13: Source decision tree



Note

- When a resource to be described is contained in another resource, the relations between the resources may be described according to the convention of a data provider. In Figure 13, a dash-lined box signifies such relation(s) and points to Section 3.9.
- The main body of Figure 13 is only focused on the description of the source of a resource. For describing relations between the resources involved, see Section 3.9.1.
- It is recommended that if the resource titles are controlled through an authority file, use the controlled title or identifier.

Table 14: Decision process with practical examples for source

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Is it contained in another resource?	No	End			
		Yes	Continue to #1			
#1	Describe the resource?	No	End			
		Yes	Continue to #2			
#2	Separate title from other parts?	No	2a	String	dc:source	Proceedings of the National Academy of Sciences of the United States of America, v. 95(10) p. 5632-5636
						http://www.pnas.org/content/by/year/2010
			2b	URI or String (when URI is not available)	dcterms:source	http://www.pnas.org/content/by/year/2010 [URI of a source] [a string conforming to a formal identification system]
		Yes	2c	String (Title)	dc:source	Proceedings of the National Academy of Sciences of the United States of America
			2d	String (Other parts)	bibo:pages	542
					bibo:section	2
					bibo:volume	95
					bibo:issue	10
					bibo:pageStart	5632
					bibo:pageEnd	5636
					bibo:chapter	II

3.3.7 Additional source information

Relation with a resource being described: Resource has source.

The properties of "landingPage" and "accessURL" are useful for providing information about the sources of datasets. The two terms are related. If the distribution(s) of datasets are accessible only through a landing page (i.e. direct download URLs are not known), then the landing page URL associated with the dcat:Dataset SHOULD be duplicated as access URL on a distribution. It is recommended to check DCAT 5.7: Dataset available only behind some web page (W3C, 2020).

Metadata terms:

- dcat:landingPage
  A web page that can be navigated to in a web browser to gain access to the catalog, a dataset, its distributions and/or additional information.
- dcat:accessURL
  A URL of the resource that gives access to a distribution of the dataset, e.g. landing page, feed, SPARQL endpoint.

3.4 Holding/location information

It is important for a resource to be located and obtained in the information exchange. Properties that record the location and availability information are taken into account in this unique group.

For datasets, DCAT includes a whole class of distribution (dcat:Distribution), as a dataset might be available in multiple serialisations that may differ in various ways, including natural language, MIME or format, schematic organisation, temporal and spatial resolution, level of detail, or profiles (which might specify any or all of the above). See the properties listed in Table 15.

3.4.1 Location/availability

Relation with a resource being described: Resource has holding or location information.

Location is considered essential information in the description of a resource in a digital repository; therefore Figure 14 foresees holding or location information as a mandatory property.

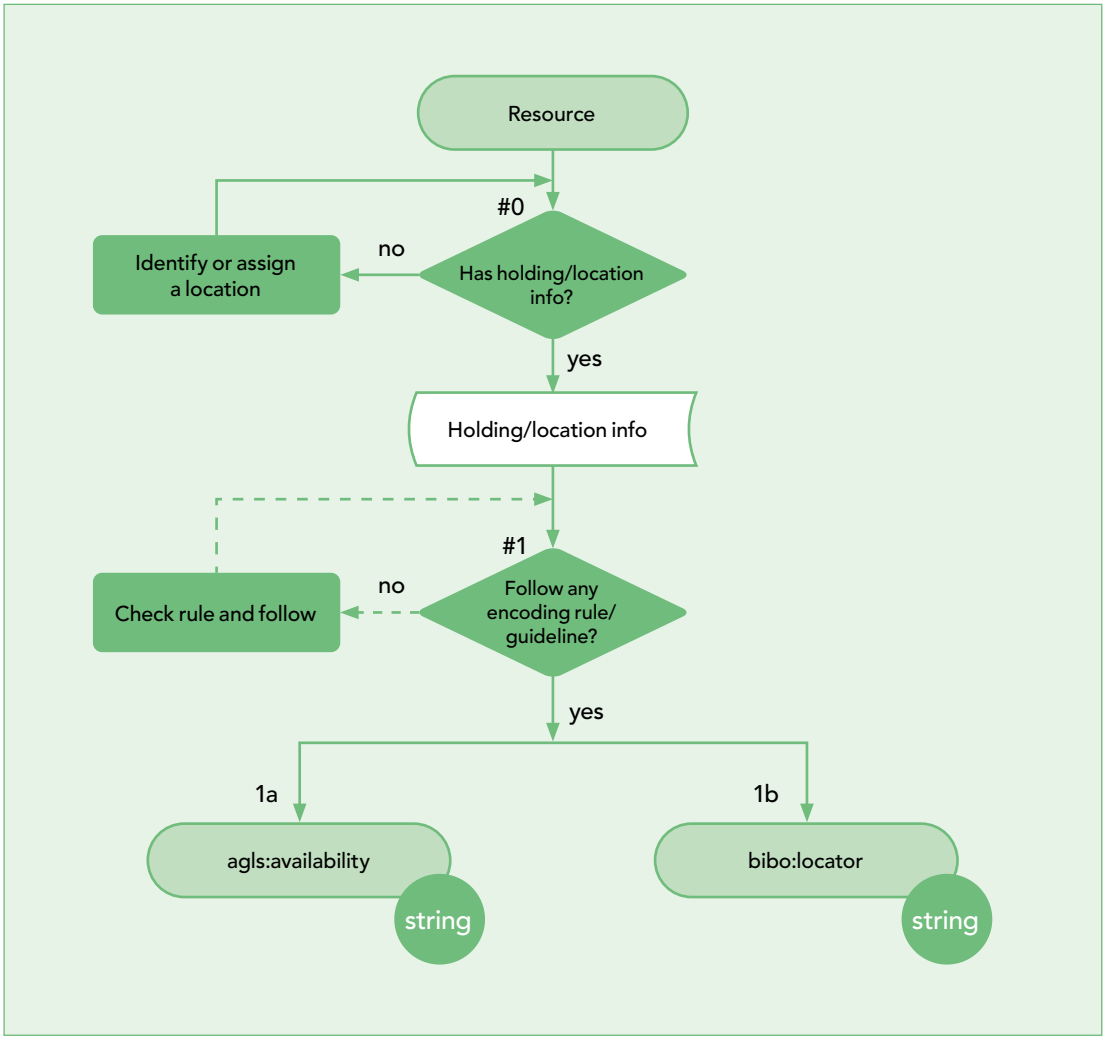


Figure 14: Location/availability decision tree

Note

- It is always recommended that location information be provided consistently by following an encoding rule or guideline.
- dc:description and dcterms: description can also be considered as metadata terms when needed.

Table 15: Decision process with practical examples for location/availability

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has holding/ location info?	No	Identify or assign a location and Go back to #0			
		Yes	Continue to #1			
#1	Follow any encoding rule or guideline?	No	Go back to #1			
		Yes	1a	String	agls:availability	http://www.example.org/services/id5678/ Contact the Publications Section on 1300 999 999 University of Vienna, Peter Jordanstr. 52, A-1190 Vienna, Austria
			1b	String	bibo:locator	Box 12, Folder 3

3.4.2 Additional properties for distribution

- **Class: dcat:Distribution**
  This represents a general availability of a dataset. It implies no information about the actual access method of the data, i.e. whether by direct download, application programming interface (API), or through a web page. The use of dcat:downloadURL property indicates directly downloadable distributions. Refer to DCAT 6.7: Distribution (W3C, 2020).

dcat:Distribution

dcat:accessService
dcat:accessURL
dcat:byteSize
dcat:compressFormat
dcat:downloadURL
dcat:mediaType
dcat:packageFormat
dcat:spatialResolutionInMeters
dcat:temporalResolution
dct:accessRights
dct:conformsTo
dct:description
dct:format
dct:issued
dct:license
dct:modified
dct:rights
dct:title
odrl:hasPolicy

Figure 15: Property of DCAT Distribution class

3.5 Subject information

In contrast to the physical characteristics, the subject group embraces the properties that describe or otherwise help the identification of what the resource is about or denotes, in the form of subject term, classification/category, freely assigned keyword, geographic term, and so on.

3.5.1 Subject

Relation with a resource being described: Resource has subject/topic.

Figure 16: Subject decision tree

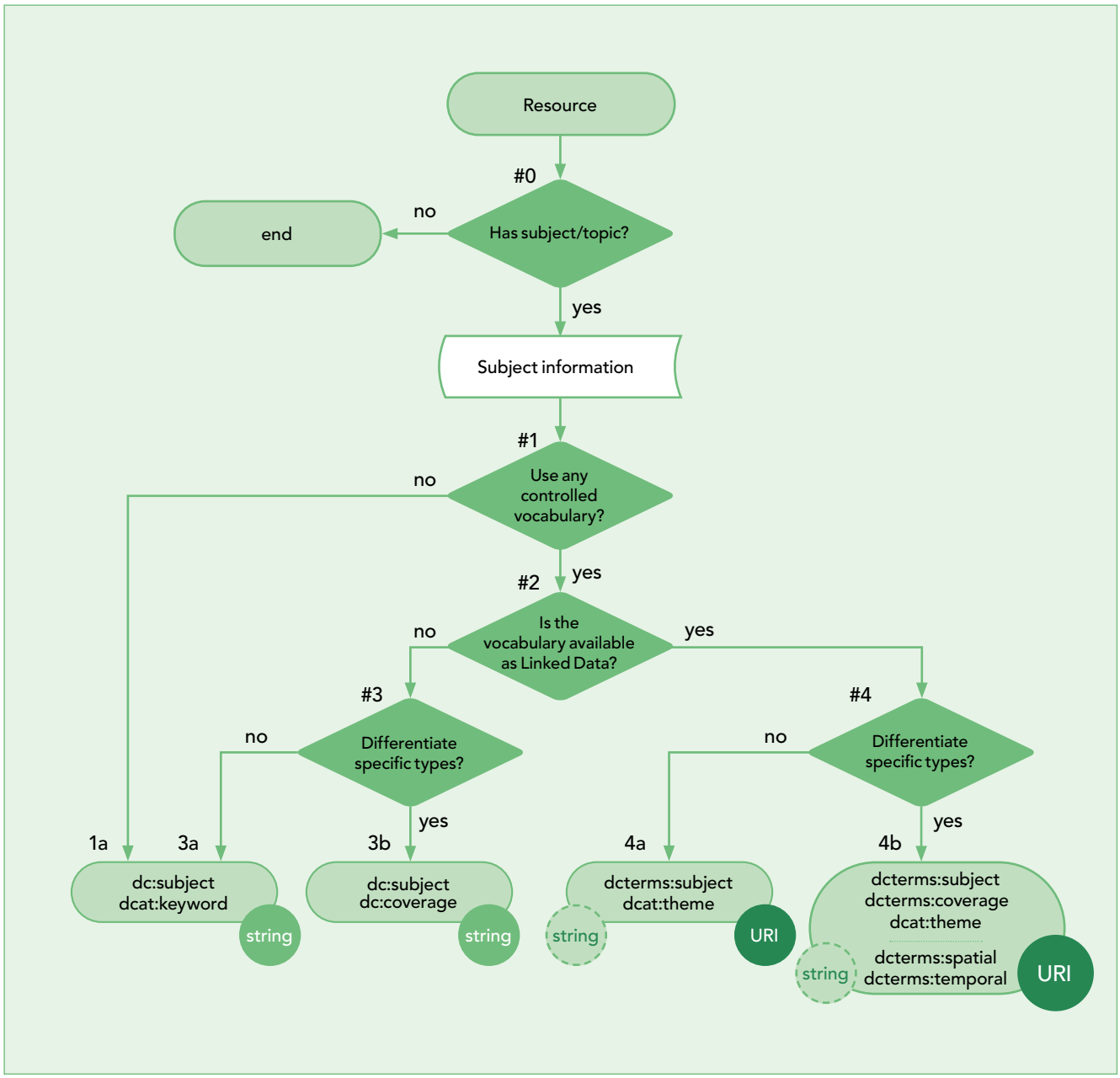


Table 16: Decision process with practical examples for Subject

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has subject/ topics?	No	End			
		Yes	Continue to #1			
#1	Use any controlled vocabulary?	No	1a	String	dc:subject dcat:keyword	paddy
						Pacific Islands & Oceania
						19th century
		Yes	Continue to #2			
#2	Is the vocabulary available as Linked Data?	No	Continue to #3			
		Yes	Continue to #4			
#3	Differentiate types of subjects?	No	3a	String	dc:subject	Rice
						Pacific Islands
						Nineteenth century
						[values from a controlled vocabulary]
		Yes	3b	String	dc:subject	Rice
						[values from a controlled vocabulary]
					dc:coverage	Pacific Islands
						Nineteenth century
						[values from a controlled vocabulary]
#4	Differentiate types of subjects?	No	4a	URI or String (when URI is not available)	dcterms:subject dcat:theme	http://aims.fao.org/aos/agrovoc/c_6599*
						http://aims.fao.org/aos/agrovoc/c_5487**
						[values from a controlled vocabulary]
		Yes	4b	URI or String (when URI is not available)	dcterms:subject dcat:theme	http://aims.fao.org/aos/agrovoc/c_6599*
						[values from a controlled vocabulary]
					dcterms:coverage	http://aims.fao.org/aos/agrovoc/c_5487**
						http://id.loc.gov/authorities/sh85091984***
						[values from a controlled vocabulary]
					dcterms:spatial	http://aims.fao.org/aos/agrovoc/c_5487**
						[values from a controlled vocabulary]
						[a literal value for spatial characteristics]
					dcterms:temporal	http://id.loc.gov/authorities/sh85091984***
						[values from a controlled vocabulary]
						[a literal value for temporal characteristics]

* http://aims.fao.org/aos/agrovoc/c_6599 is the URI of a concept in AGROVOC. Its preferred English label is "Rice".
** http://aims.fao.org/aos/agrovoc/c_5487 is the URI of a concept in AGROVOC. Its preferred English label is "Pacific Islands".
*** http://id.loc.gov/authorities/sh85091984 is the URI of a concept in LCSH. Its preferred English label is "Nineteenth century".

Note

- It is always recommended to index the concept/topic/subject/category of a resource. Examples of values include: concepts represented by terms from a controlled vocabulary; keywords; classes or categories represented by notations or labels from a classification system.
- More and more controlled vocabularies are published as Linked Data where concepts are represented by non-literal values (i.e. an identifier and/or a http URI). For example, each AGROVOC concept has its unique http URI. LODE-BD recommends using these URIs instead of the literal forms (i.e. the labels) as values when considering moving towards publishing your data as Linked Data.

Examples of values:

- A concept URI of your own controlled vocabulary.
- A URI of a concept from a published thesaurus (e.g. EuroVoc) or classification (e.g. Dewey Decimal Classification).
- A URI of an agent when the agent is the subject/focus of a resource (e.g. URI of a conference defined in a FOAF file, a URI of a person from Virtual International Authority File (VIAF).
- If none of these is not possible or feasible, a literal value that identifies the subject may be provided.
- Usually a value encoding scheme's title (e.g. AGROVOC or Library of Congress Subject Headings) should be indicated along with the value. Also when using literal forms than URIs, the language of the words should be indicated (consult References if needed).
- For coverage which is broadly defined, it is preferable to use the more specific sub-properties temporal and spatial.
- For datasets, the values describe the categories and their relations in the catalog.

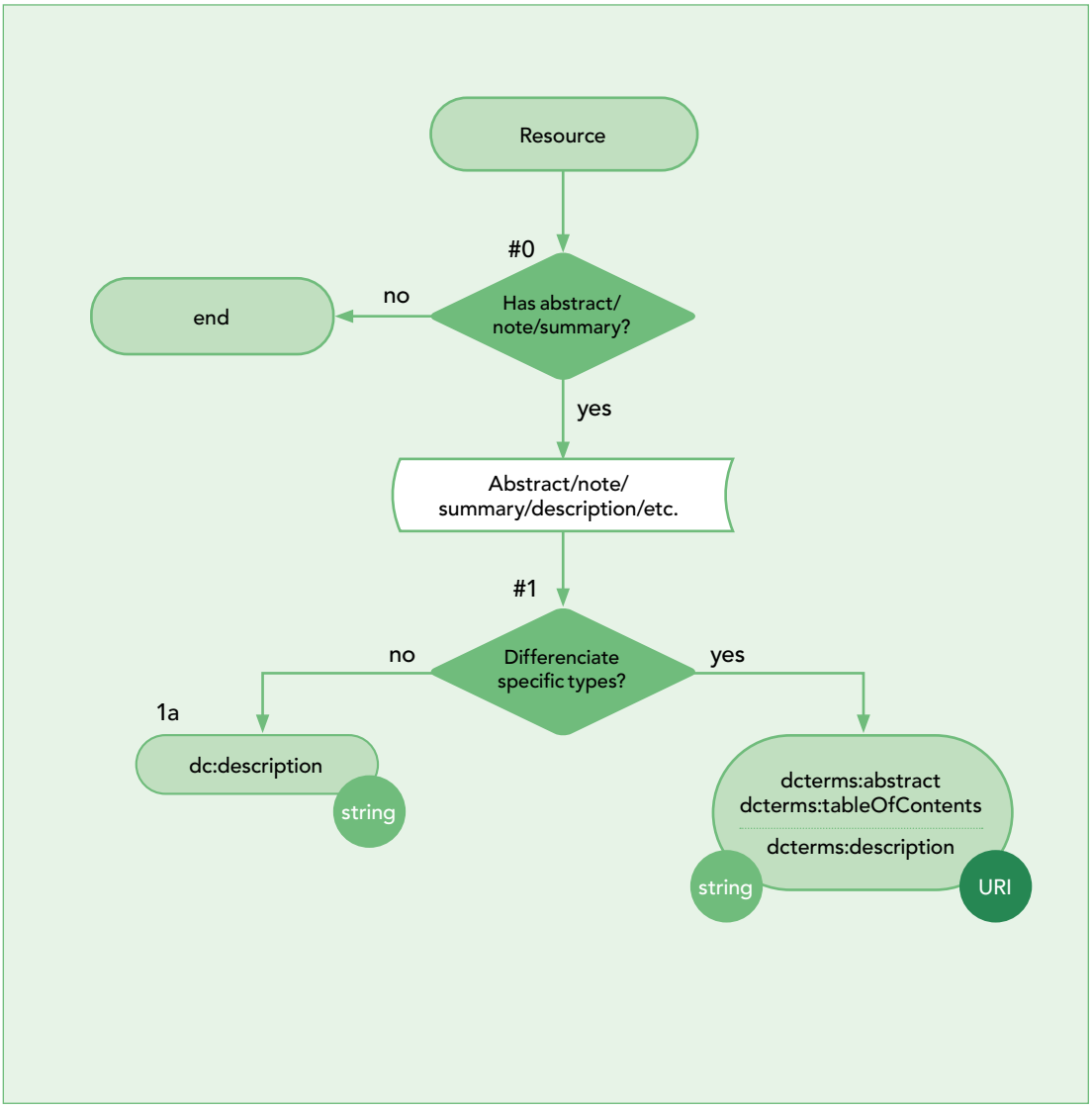
3.6 Description of content

Two major types of descriptions that focus on the content of the resource rather than the physical object are considered in this group: a) any representative description of the content, usually in the form of abstract, summary, note, and table of contents; and b) type or genre of the resource.

3.6.1 Description/abstract/table of contents

Relation with a resource being described: Resource has description, abstract or table of contents.

Figure 17
Description/abstract/table of contents decision tree



Note

- In describing the content, different words might have been used, such as "abstract" vs. "note", or "description" vs. "summary". A table of contents may also be presented in a description.

Table 17: Decision process with practical examples for description/abstract/table of contents

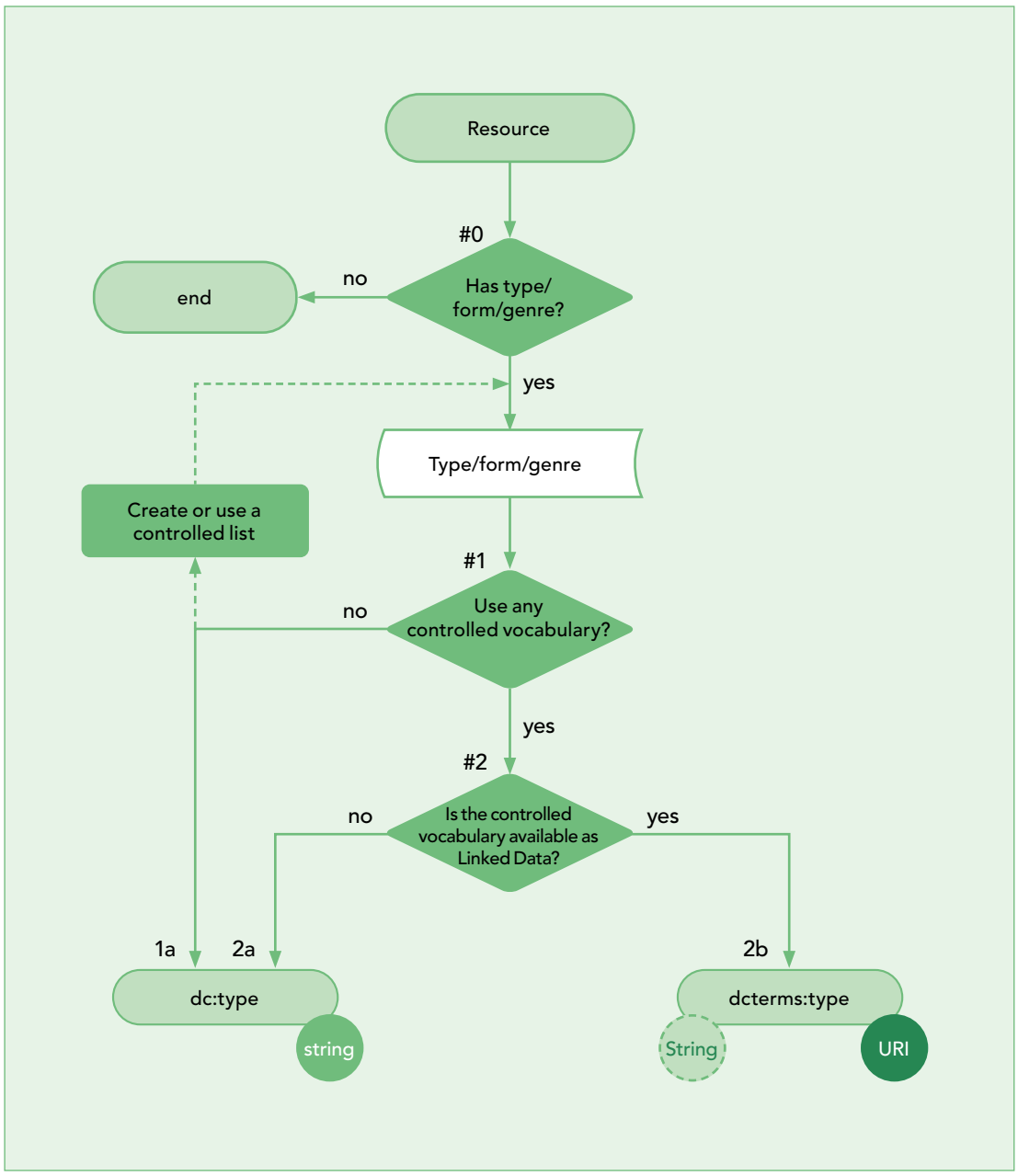
Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has abstract/ note/ summary?	No	End			
		Yes	Continue to #1			
#1	Differentiate types of content descriptions?	No	1a	String	dc:description	One of the least understood aspects of population biology is …
		Yes	1b	String or URI	dcterms:abstract	One of the least understood aspects of population biology is …
						http://jeclap.oxfordjournals.org/content/2/4/391.abstract*
					dcterms: tableOfContents	Introduction -- Formal theory -- Coevolution --
						http://preservationtutorial.library.cornell.edu/toc.html**
					dcterms:description	Contains a series of articles which are intended to …
						VocBench is a web-based, multilingual, vocabulary editing and workflow tool developed by FAO. It … ***
						http://aims.fao.org/tools/vocbench-2***

* The URL is the abstract of Moss (2001).
** The URL is the Table of Contents page of Cornell University Library (2000-2003).
*** Both the text and URL are from VocBench (n.d.).

3.6.2 Type/form/genre

Relation with a resource being described: Resource has type/form/genre.

Figure 18: Type/form/genre decision tree



Note

- It is always recommended that a controlled vocabulary (e.g., DCMI Type Vocabulary) be used or created for your collection when describing a resource type.
- To describe the file format, physical medium, or dimensions of the resource, use the property format.

Table 18: Decision process with practical examples for type/form/genre

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has type/form/genre?	No	End			
		Yes	Continue to #1			
#1	Use any controlled vocabulary?	No	1a	String	dc:type	Lecture; Poster, ...
		Yes	Continue to #2			
#2	Is the controlled vocabulary available as Linked Data?	No	2a	String	dc:type	Interactive Resource [values from a controlled vocabulary]
		Yes	2b	URI or String (when URI is not available)	dcterms:type	http://purl.org/dc/dcmitype/InteractiveResource* [values from a controlled vocabulary]

* http://purl.org/dc/dcmitype/InteractiveResource is the URI of the concept "Interactive Resource", from DCMI Type Vocabulary https://www.dublincore.org/specifications/dublin-core/dcmi-terms/#section-7.

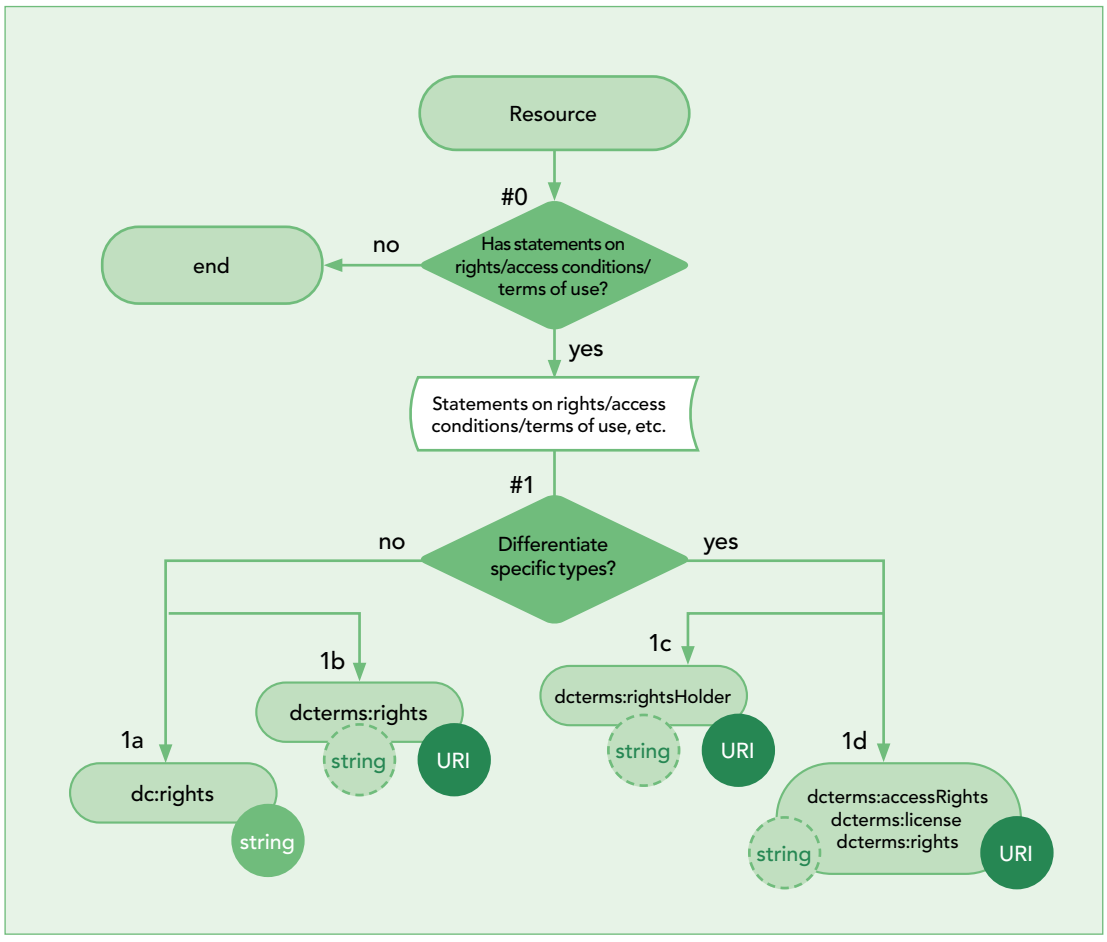
3.7 Intellectual property rights

Any property that deals with an aspect of intellectual property rights relating to access and use of a resource is included in this group, with special regard to rights, terms of use, and access condition.

3.7.1 Right statements

Relation with a resource being described: Resource has intellectual property rights statements.

Figure 19: Right statements decision tree



Note

- The property may be named as "rights" or "rights statement". More detailed types of statements may include access rights, terms of use, access condition/access rights, and license.
- Examples of the values (strings or URIs) are from FAO webpages and other resources.
- Value vocabularies are based on DCAT Section 8: License and rights statements recommendations (W3C, 2020).
- If an URI is not possible or feasible, a literal value (name, label, or short text) may be provided.

Table 19: Decision process with practical examples for right statements

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has statements on rights/ access conditions/ terms of use?	No	End			
		Yes	Continue to #1			
#1	Differentiate specific types or parts in the rights statement?	No	1a	string	dc:rights	Copyright 1996-2007 XYZ Productions. All rights reserved.
						http://www.fao.org/corp/copyright/en/
			1b	URI or String (when URI is not available)	dcterms:rights	http://www.fao.org/corp/copyright/en/ [URI of the rights holder] [a literal value that identifies the rights holder if a URI is not possible or feasible]
		Yes	1c	URI or String (when URI is not available)	dcterms:rightsHolder	http://www.fao.org/ [URI of the rights holder] [a literal value that identifies the rights holder if a URI is not possible or feasible]
			1d	URI or String (when URI is not available)	dcterms:accessRights	http://www.fao.org/corp/copyright/en/
						[values from code lists/taxonomies].*
					dcterms:license	http://creativecommons.org/licenses/by/3.0/
						[URIs defined by Creative Commons.]** [a literal value that identifies the license if a URI is not possible or feasible]
					dcterms:rights	[URL of any other type of statements] [a literal value that identifies the rights, if a URI is not possible or feasible]

* Examples of the code list: (1) Access rights name authority list of EU: https://op.europa.eu/en/web/eu-vocabularies/at-dataset/-/resource/dataset/access-right; (2) Eprints Access Rights Vocabulary Encoding Scheme http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_AccessRights_Vocabulary_Encoding_Scheme

** https://creativecommons.org/about/cclicenses/

3.7.2 [Additional] intellectual property rights information

Information about rights expressed as an The Open Digital Rights Language (ODRL) policy using the ODRL vocabulary may be provided using odrl:hasPolicy property. ODRL enables the statements about the usage (i.e. permissions, prohibitions, and obligations) of content and services of the particular resource (W3C, 2018).

Metadata term

- odrl:hasPolicy - It is used for a policy expressing the rights associated with the resource.

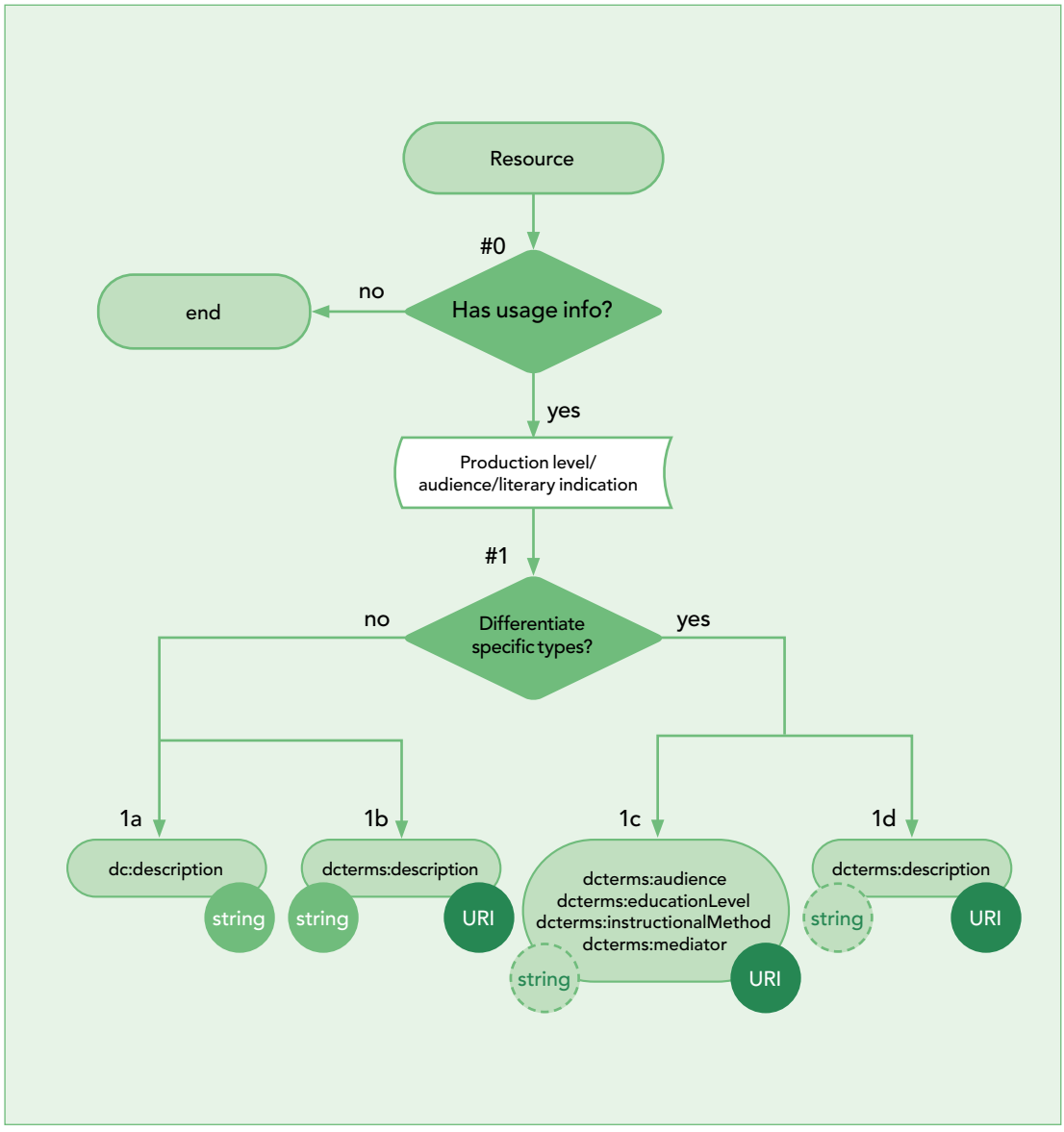
3.8 Usage

Properties that are related to the use of a resource, rather than the characteristics of the resource itself, are considered to belong to this group. Typical properties are: audience, literary indication, and education level.

3.8.1 Audience/literary indication/education level

Relation with a resource being described: Resource has usage information.

Figure 20: Audience/literary indication/education level decision tree



Note

- In presenting the usage-related information, different words might be used in your situation, for example, "Production Level", "Audience", "Literary Indication", etc.

Table 20: Decision process with practical examples for audience/literary indication/education level

Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Has usage info?	No	End			
		Yes	Continue to #1			
#1	Differentiate specific types of usage data? (e.g. Production level/ audience/ literary indication, etc.)	No	1a	String	dc:description	audience: Public*
			1b	String or URI	dcterms:description	audience: Public*
						[URI]
		Yes	1c	URI or String (when URI is not available)	dcterms:audience	rdfs:label "Public"* [values from a controlled vocabulary]
					dcterms:educationLevel	rdfs:label "UK Educational Level 1"** [values from a controlled vocabulary]
					dcterms:instructionalMethod	rdfs:label "Direct Teaching"***
					dcterms:mediator	rdfs:label "Reading specialist" ****
			1d	URI or String	dcterms:description	[String or URI for any other usage data]

* Example taken from ProdINRA sample record.

** Example taken from UK Educational Levels (UKEL) list: http://www.ukoln.ac.uk/metadata/education/ukel/

*** Example taken from ADPRIMA Instructional Methods Information list of Instructional Methods: http://www.adprima.com/teachmeth.htm

**** Example taken from Hillman, 2005.

3.9 Relation

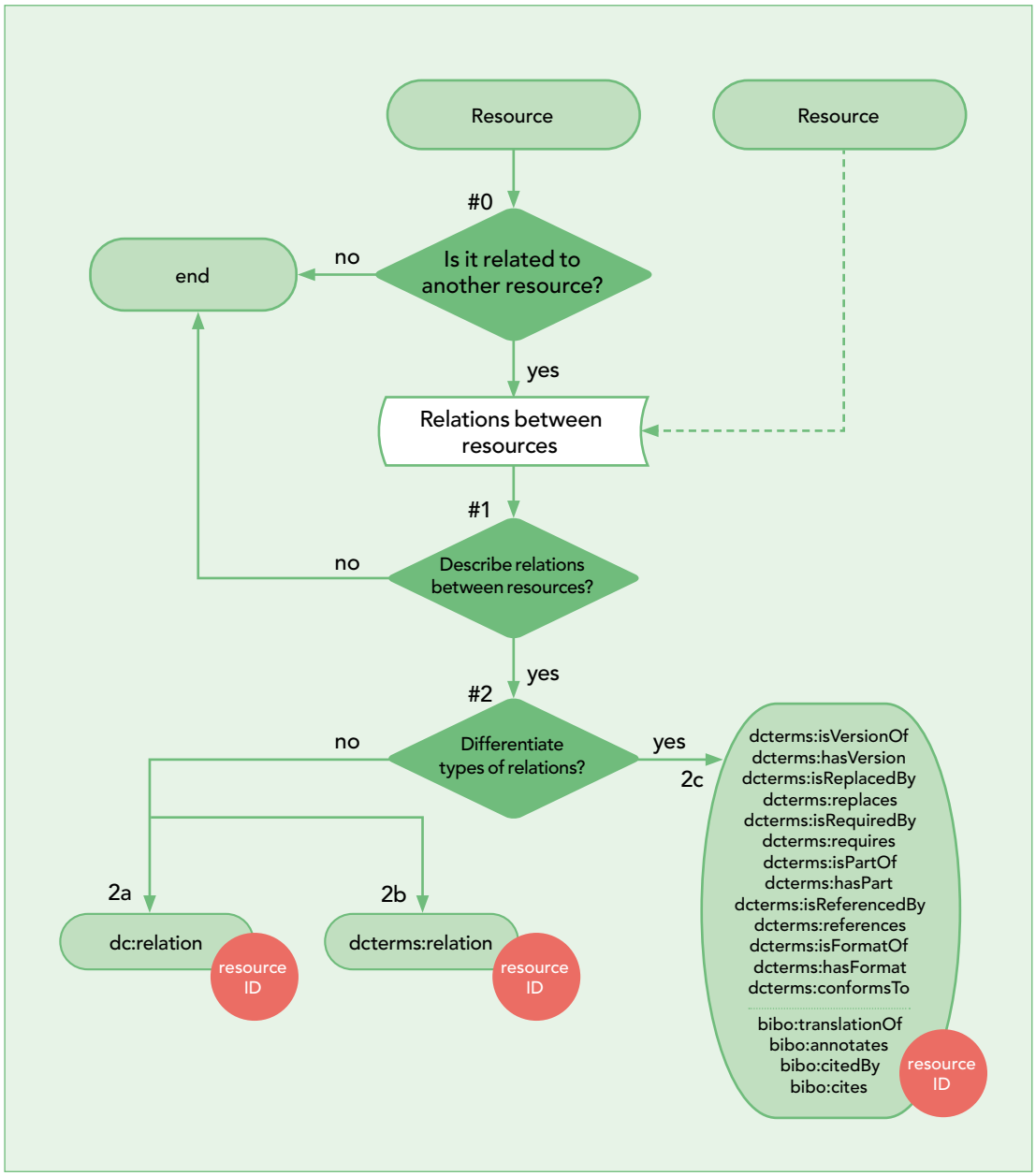
This group has a different perspective for describing the resources from other groups that focus on describing the resource itself.

Here, various relations between two resources or between two agents are the focus of the description.

3.9.1 Relation between resources

Relation being described: The resource is related to another resource.

Figure 21: Relation between resources decision tree



Note

- When a resource is related to another resource, a decision should be made regarding whether the relations between the two resources need to be described.
- In describing the relations, a great number of relation types can be used. The available metadata terms listed below do not form an exhaustive list. Other types may exist.
- The involved resources should always be represented by their identifiers. Values for this property are always the identifiers.

Table 21: Decision process with practical examples for relation between resources

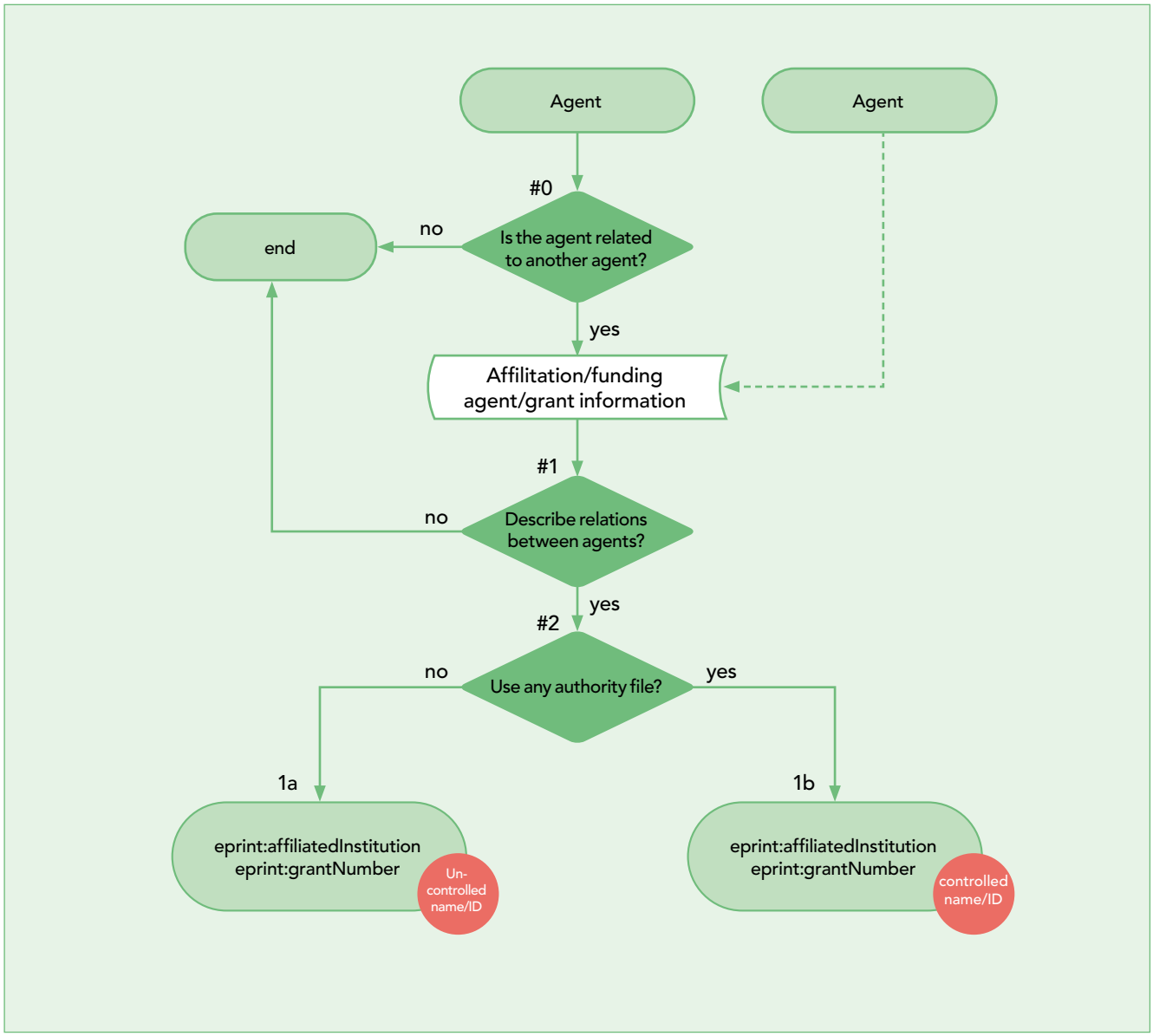
Decision	Question	Answer	Action	Value type	Examples	
					Metadata term	Value
#0	Is it related to another resource?	No	End			
		Yes	Continue to #1			
#1	Describe relations between resources?	No	End			
		Yes	Continue to #2			
#2	Differentiate types of relations?	No	2a	ID	dc:relation	[a resource]
			2b	ID	dcterms:relation dcat:qualifiedRelation*	[a resource]
		Yes	2c	ID	dcterms:isVersionOf dcterms:hasVersion dcterms:isReplacedBy dcterms:replaces dcterms:isRequiredBy dcterms:requires dcterms:isPartOf dcterms:hasPart dcterms:isReferencedBy dcterms:references dcterms:isFormatOf dcterms:hasFormat dcterms:conformsTo bibo:translationOf bibo:annotates bibo:citedBy bibo:cites	[a resource]

* dcat:qualifiedRelation is for linking to another resource where the nature of the relationship is known but does not match one of the DCTERMS properties listed in this table.

3.9.2 Relation between agents

Relation being described: The agent is related to another agent, specifically affiliation or funding relation.

Figure 22
Decision tree for relation between agents



Note

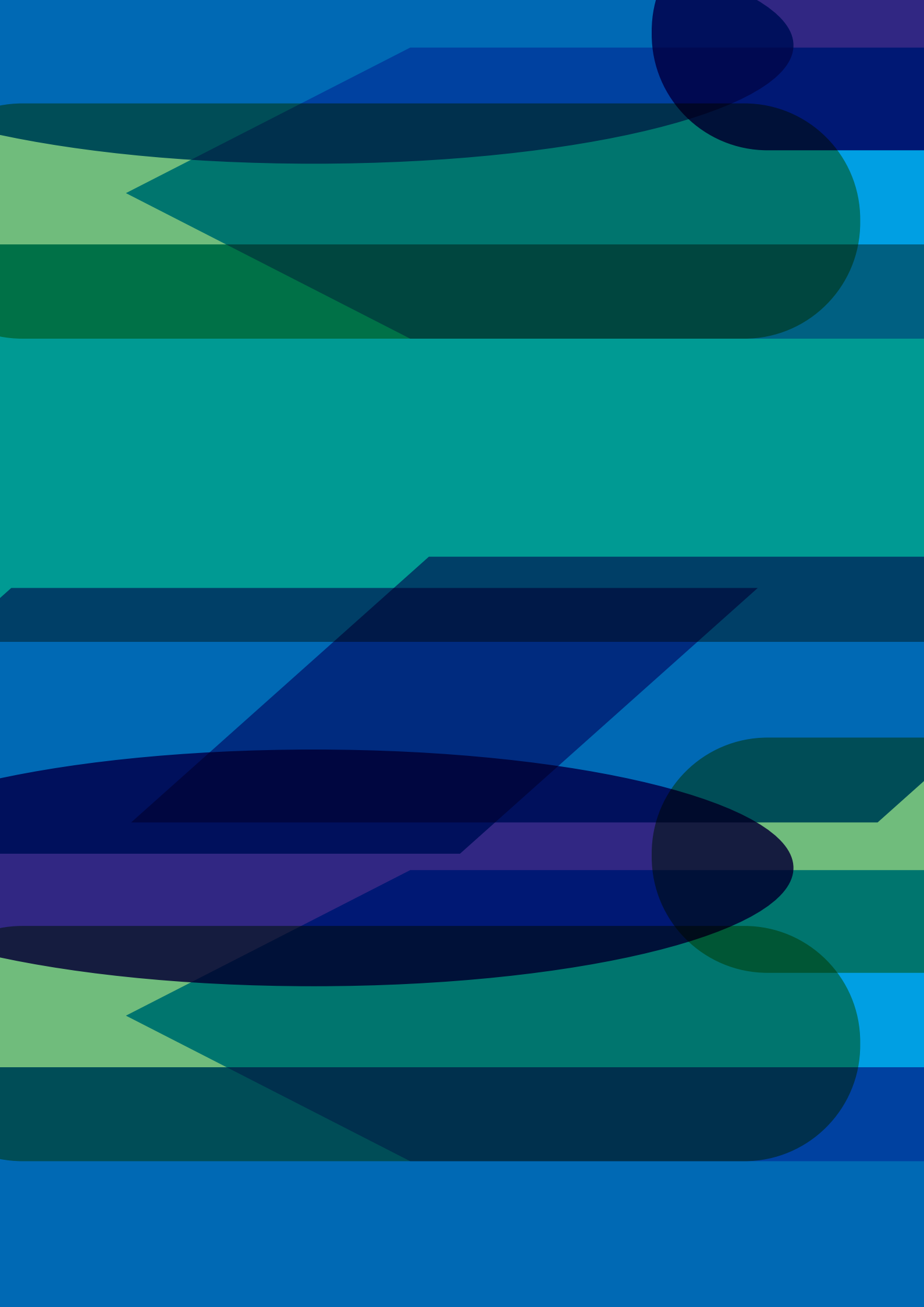
- When an agent is related to another agent, a decision needs to be made regarding whether the relations between the two agents should be described.
- There could be various types of relations between agents. The available metadata terms listed below focus on the affiliation and funding information and do not form an exhaustive list. Consult MARC List for Relators (LC, n.d.) for more types of relators.
- It is highly recommended that agents always be represented by their identifiers or controlled names.
- For research funding agents, dcat:hadRole can be used if applicable. Properties in the incoming OpenAIRE Version 4, for Funding Reference may give appropriate metadata elements.

Table 22: Decision process with practical examples for relation between agents

Decision	Question	Answer	Action	Value type	Examples	
					Metadata form	Value
#0	Is the agent related to another agent?	No	End			
		Yes	Continue to #1			
#1	Describe relations between agents?	No	End			
		Yes	Continue to #1			
#2	Do you use any authority file for the names of the agents?	No	1a	Un-controlled name/ID	eprint:affiliatedInstitution*	University of Bristol
					eprint:affiliatedInstitution*	A456X
		Yes	1b	Controlled name/ID	eprint:affiliatedInstitution*	University of Bristol
					eprint:grantNumber**	A456X

* The eprint:affiliatedInstitution originally has a constraint of domain "ScholarlyWork", which would not make use of it for relating agents. EPrint defined "Affiliated Institution" as "An organisation to which a creator of the eprint is affiliated"; LODE-BD considers it appropriate and uses this metadata term without the domain constraint. If a better namespace and metadata term can be identified in the future, LODE-BD will consider a replacement.

** The eprint:grantNumber originally has a constraint of domain "ScholarlyWork", which would not make the use of it for relating agents. Based on its definition, a Grant Number is "An alpha-numeric string identifying the funding grant under which the eprint was written". LODE-BD considers it appropriate and uses this metadata term without the domain constrain. If a better namespace and metadata term can be identified in the future, LODE-BD will consider a replacement.



4 The step forward (with further reading)

4.1 Implementation options

For anyone who is contributing to an open bibliographic data repository and considering preparing LODE-BD, this guide has provided recommendations on issues related to processes and strategies. LODE-BD addressed three questions:

a. What kinds of entities and relationships are involved in describing and accessing bibliographic resources?

b. What properties should be considered for publishing meaningful/useful LOD-ready bibliographic data?

c. What metadata terms are appropriate in any given property when producing LOD-ready bibliographic data from a local database?

In Chapter 2, LODE-BD presents nine groups of common properties and the selected metadata terms to be used for describing bibliographic resources. In Chapter 3, this guide demonstrates how to make decisions on selecting recommended properties according to local needs.

After metadata terms are selected based on the decision trees, a data provider should come up with a list of the metadata terms that are appropriate for its existing bibliographic data. To implement these metadata terms, LODE-BD shares two options based on best practices.

• **Option #1**, "Design-time" strategy: The data provider would need to change its current ad-hoc model, replacing it with the LODE-BD recommended M2B model and those selected metadata terms. This means some changes to a database and the services that access it.

• **Option #2**, "Run-time" strategy: The data provider would keep the original ad-hoc model and database structure. A data provider would convert bibliographic data on the fly to a M2B model upon request. This means adding a conversion service and leaving the ad-hoc model unchanged.

Turning bibliographic data from an ad-hoc modeled database in a silo, to data in a standardised metadata repository, is a giant leap and enables unified data records from various data providers to be maximised when users are searching and browsing through the repository. Furthermore, this could also be a step forward to the LOD Cloud. One outcome is that individual data providers can directly produce RDF triples using LODE-BD recommended metadata terms. Moving another step towards the LOD Cloud can also be accomplished through the metadata repository, which would publish its bibliographic data as Linked Data, as illustrated in Figure 23. For the creation of RDF triples or a metadata repository, preparing LOD-ready metadata by data providers is essential.

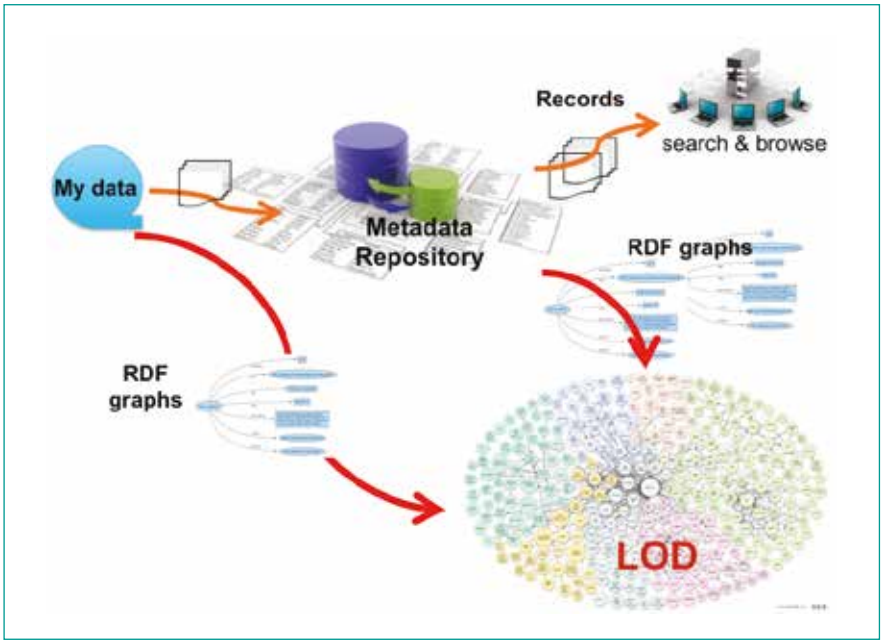


Figure 23
Output of LOD-ready metadata. Source: Marcia Zeng, 2011. CC0

4.2

How to create and consume Linked Data

Linked Data: Evolving the Web into a Global Data Space (1st edition), Tom Heath and Christian Bizer (2011). Synthesis Lectures on the Semantic Web: Theory and Technology. http://linkeddatabook.com/editions/1.0/

"This book gives an overview of the principles of Linked Data as well as the Web of Data that has emerged through the application of these principles. The book discusses patterns for publishing Linked Data, describes deployed Linked Data applications and examines their architecture."

Linked Data Patterns, Leigh Dodds and Ian Davis (2011). http://patterns.dataincubator.org/book

"A pattern catalogue for modelling, publishing, and consuming Linked Data."

Linked Open Data Star Scheme by Example, Posted by woddiscovery (2010) https://webofdata.wordpress.com/2010/11/12/linked-open-data-star-scheme-by-example

"Tim Berners-Lee suggested a 5-star deployment scheme for Linked Open Data and Ed Summers provided a nice rendering of it."

Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish your Metadata, Seth van Hooland and Ruben Verborgh (2014).

Linked Data – Design Issues, Tim Berners-Lee (2006). www.w3.org/DesignIssues/LinkedData.html

One of the first discussions of the topic, mentioning the "four rules of Linked Data".

The FAIR Guiding Principles for Scientific Data Management and Stewardship, Mark Wilkinson, Michael Dumontier, IJusbrand Jan Aalbersberg *et al.* (2016).

Cool URIs for the Semantic Web. Leo Sauermann and Richard Cyganiak (2008). www.w3.org/TR/cooluris

Linked Data, W3C (n.d.) www.w3.org/standards/semanticweb/data

Explains: what is linked data • what is linked data used for • examples • learn more •current status of specifications and groups. Provides links to a collection of semantic web case studies and use cases.

Large-scale Semantic Integration of Linked Data: A Survey. Mountantonakis Michalis, and Yannis Tzitzikas (2019). https://dl.acm.org/doi/fullHtml/10.1145/3345551

Surveys the work that has been done in the area of Linked Data integration; identifies the main actors and use cases; analyses and factorises the integration process according to various dimensions, and discusses the methods that are used in each step.

Tutorials and Courses, W3C (n.d.). www.w3.org/2002/03/tutorials#semanticweb_data

Provides selected tutorials and online courses to help people learn about W3C technologies www.w3.org/standards.

Data on the Web Best Practices, Bernadette Lóscio, Caroline Burle and Newton Calegari (2017). www.w3.org/TR/dwbp

Provides best practices related to the publication and usage of data on the web designed to help support a self-sustaining ecosystem. Following these best practices will facilitate interaction between publishers and consumers.

Linked Data Glossary, W3C (2013-) www.w3.org/TR/ld-glossary

Provides a glossary of terms defined and used to describe Linked Data, and its associated vocabularies and Best Practices.

4.3

Where to find vocabularies (metadata vocabularies and value vocabularies)

BARTOC (Basel Register of Thesauri, Ontologies & Classifications)
URL: https://bartoc.org

BioPortal
URL: www.bioportal.bioontology.org

Linked Open Data Cloud
URL: http://lod-cloud.net/

The diagram is based on metadata collected and curated by contributors to the CKAN's Datahub directory https://old.datahub.io. See datasets in the Linking Open Data (LOD) Cloud's group https://lod-cloud.net of the https://old.datahub.io.

Linked Open Vocabularies (LOV)
URL: https://lov.linkeddata.es/dataset/lov

A portal of Resource Description Framework Schema (RDFS) vocabularies or Web Ontology Language (OWL) ontologies defined for and used by LD datasets.

Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets, Antoine, Isaac, William Waites, Jeff Young, and Marcia Zeng (2011).

URL: www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025

A side delivery of the W3C Linked Library Data XG which lists relevant metadata element sets, value vocabularies that are reported in the Linked Library Data use cases and case studies. Each entry contains link URL, namespace, and short description.

4.4

How to express metadata with different syntaxes: text, html. xml, rdf, and rdfa

DC-TEXT [DCMI Recommendation]. "Expressing Dublin Core metadata using the DC-Text format", Pete Johnston (2007)

URL: www.dublincore.org/specifications/ublin-core/dc-text

Its primary use is in presenting metadata constructs for human consumption.

DC-HTML [DCMI Recommendation]. "Expressing Dublin Core metadata using HTML/XHTML meta and link elements", Pete Johnston and Andy Powell (2008)

URL: www.dublincore.org/specifications/dublin-core/dc-html

It describes how a Dublin Core metadata description set can be encoded using the HTML/XHTML <meta> and <link> elements. This specification is also an HTML "meta data profile" as defined by the HTML specification.

DC-DS-XML [DCMI Proposed Recommendation]. "Expressing Dublin Core Description Sets using XML (DC-DS-XML)", Pete Johnston and Andy Powell (2008)

URL: www.dublincore.org/specifications/dublin-core/dc-ds-xml

It specifies an XML format for representing a Dublin Core metadata description set.

DC-RDF [DCMI Recommendation]. "Expressing Dublin Core metadata using the Resource Description Framework (RDF)", Mikael Nilsson, Andy Powell, Pete Johnston, and Ambjorn Naeve (2008)

URL: www.dublincore.org/specifications/dublin-core/dc-rdf

It describes how constructs of the DCMI Abstract Model may be expressed in RDF graphs.

User Guide/Publishing Metadata, Stefanie Rühle, Tom Baker and Pete Johnston (n.d.).

URL: www.dublincore.org/resources/userguide/publishing_metadata

"How to use DCMI Metadata as Linked Data."

Linked Data Tutorial NG – Publishing and Consuming Linked Data Embedded in HTML, Michael Hausenblas and Richard Cyganiak (2001).

URL: www.w3.org/2001/sw/interest/ldh

"This document provides guidelines for how to create and consume Linked Data embedded in HTML."

Validating and converting tools

- RDF Validator: www.w3.org/RDF/Validator
- XML Validator: www.w3schools.com/xml/xml_validator.asp
- RDF Converter (including RDF/XML and other formats): www.easyrdf.org/converter

4.5 Why publish bibliographic data as Linked Data?

Library Linked Data Incubator Group Final Report, Thomas Baker, Emmanuelle Bermès, Karen Coyle *et al.*, (2011).

URL: www.w3.org/2005/Incubator/lld/XGR-lld-20111025

- Draft Benefits of the Linked Data Approach, W3C (2011) http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/#Benefits_of_the_Linked_Data_Approach
- Recommendations www.w3.org/2005/Incubator/lld/XGR-lld-20111025/#Recommendations

Linked Data: from Library Entities to the Web of Data, Richard Wallis (2014).

URL: www.slideshare.net/rjw/linked-data-from-library-entities-to-the-web-of-data

5 Standards used

Metadata terms used in the LODE-BD guide are rooted in various well-known metadata specifications that are presented in Table 3 and Table 4. Usually metadata terms from the Dublin Core namespaces are the fundamentals, while metadata terms from other namespaces are supplemented when additional needs are to be satisfied. They are:

@prefix **dc:** <http://purl.org/dc/elements/1.1/>
(Dublin Core Metadata Element Set namespace) (Dublin Core, 2012)

Dublin Core Metadata Element Set (DCMES or DC)
Dublin Core Metadata Initiative (DCMI)
Namespace: http://purl.org/dc/elements/1.1
Page: http://dublincore.org/documents/dces

Note: The Dublin Core Metadata Element Set (ISO 15836) is a vocabulary of fifteen properties for use in resource description.

@prefix **dcterms:** <http://purl.org/dc/terms/>
(DCMI Metadata Terms namespace) (Dublin Core, 2020)

DCMI Metadata Terms
Dublin Core Metadata Initiative (DCMI)
Namespace: http://purl.org/dc/terms
Page: http://dublincore.org/documents/dcmi-terms

Note: The DCMI Metadata Terms is an authoritative specification of all metadata terms maintained by DCMI. As a full set of DCMI vocabularies it also includes sets of resource classes (including the DCMI Type Vocabulary), vocabulary encoding schemes, and syntax encoding schemes.

@prefix **bibo:** <http://purl.org/ontology/bibo/>
(Bibliographic Ontology namespace) (Bibliographic Ontology, 2009)

Bibliographic Ontology
Bibliographic Ontology Specification Group
Namespace: http://purl.org/ontology/bibo
Page: http://bibliontology.com/specification

Note: The Bibliographic Ontology is designed for use in describing bibliographic things on the Semantic Web in RDF.

@prefix **agls:** <http://www.agls.gov.au/agls/terms/>
(AGLS Metadata Standard namespace) (AGLS, 2008)

AGLS Metadata Standard
Australian Government Locator Service
Namespace: www.agls.gov.au/agls/terms
Page: www.agls.gov.au/documents/aglsterms

Note: The AGLS Metadata Standard (Australian Standard AS 5044-2010) is developed to promote consistency of discovery of government resources. It provides a set of metadata properties and associated usage guidelines to improve the visibility, manageability and interoperability of online information and services.

@prefix **eprint:** <http://purl.org/eprint/terms/>
(Eprints namespace) (UKOLN, 2008)

Eprints Terms
UKOLN, JISC
Namespace: http://purl.org/eprint/terms/
Page: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Terms

Note: The Eprints Terms include eprints-specific metadata properties and encoding schemes that have been created as part of the Dublin Core-based Scholarly Works Application Profile.

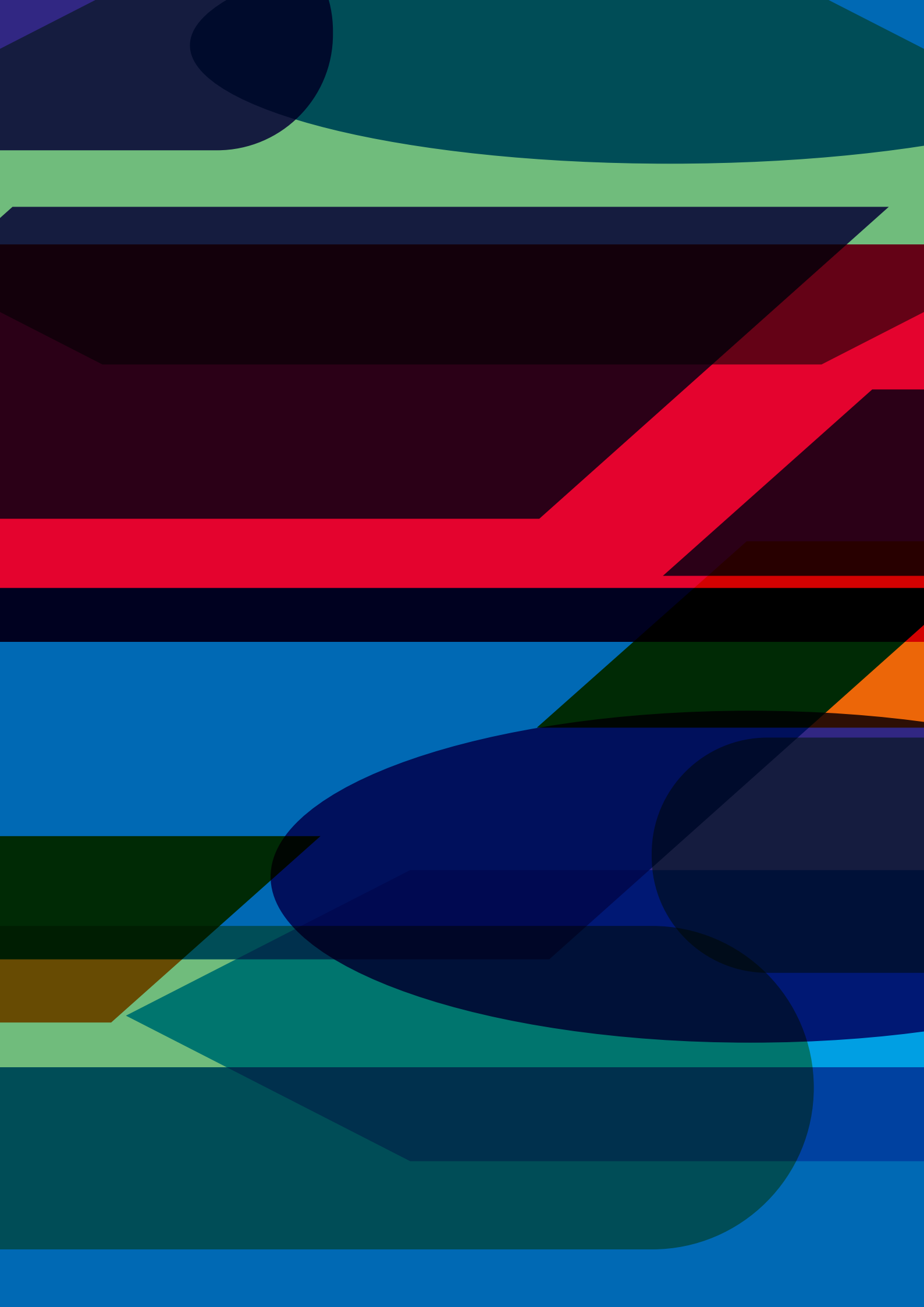
@prefix **prov:** < http://www.w3.org/ns/prov#>
(PROV-O: The PROV Ontology namespace) (W3C, 2013)

PROV-O: The PROV Ontology
W3C Recommendation 30 April 2013
Namespace: www.w3.org/ns/prov#
Page: www.w3.org/TR/prov-o
PROV-O provides a set of classes, properties, and restrictions that can be used to represent and interchange provenance information generated in different systems and under different contexts.

@prefix **dcat:** <http://www.w3.org/ns/dcat#> (Data Catalog Vocabulary – Version 2 namespace) (W3C, 2020)

Data Catalog Vocabulary – Version 2
W3C Recommendation 04 February 2020.
Namespace: www.w3.org/ns/dcat#
Prefix for the DCAT namespace: dcat
Page: www.w3.org/TR/vocab-dcat-2

Note: DCAT is an RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web. DCAT 2 supersedes DCAT [VOCAB-DCAT-20140116], while maintaining the DCAT namespace.



6 References

AGLS. 2008. AGLS Metadata Standard – Reference [online]. [Cited July 2020]. www.agls.gov.au/documents/aglsterms

AGLS. 2020. AGLS Metadata Standard [online]. [Cited July 2020]. www.agls.gov.au

Baker, T., Bermès, E., Coyle, K., Dunsire, G., Isaac, A., Murray, P., Panzer, M., Schneider, J., Singer, R., Summers, E., Waites, W., Young, J. & Zeng, M. 2011. Library Linked Data Incubator Group Final Report [online]. [Cited July 2020]. www.w3.org/2005/Incubator/lld/XGR-lld-20111025

BCP47 - Tags for Identifying Languages. 2009. Ed. Phillips, A., & Davis, M. IETF Trust. [online]. [Cited July 2020]. https://tools.ietf.org/html/bcp47

Berners-Lee, T. 2006. Linked Data. World Wide Web Consortium [online]. [Cited July 2020]. www.w3.org/DesignIssues/LinkedData

Bibliographic Ontology. 2009. Bibliographic Ontology Specification [online]. [Cited July 2020]. http://bibliontology.com

Cornell University Library. 2000-2003. Moving Theory into Practice [online]. [Cited July 2020]. http://preservationtutorial.library.cornell.edu

Dodds, L. & Davis, I. 2012. Linked Data Patterns: A Pattern Catalogue for Modelling, Publishing, and Consuming Linked Data [online]. [Cited July 2020]. https://patterns.dataincubator.org/book

DOI. 2020. The DOI® System [online]. [Cited July 2020]. www.doi.org

Dublin Core. 2009. DCMI Abstract Model. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/abstract-model

Dublin Core. 2012. Dublin CoreTM Metadata Element Set, Version 1.1: Reference Description [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/dcmi-terms

Dublin Core. 2020. DCMI Metadata Terms [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/dcmi-terms

E-prints in Library & Information Service (e-LIS). 2020. Welcome to e-LIS Repository [online]. [Cited July 2020]. http://eprints.rclis.org

GTIN. 2020. GTIN Definition: Information [online]. [Cited July 2020]. www.gtin.info

Hausenblas, M. & Cyganiak. R. 2001. Publishing and Consuming Linked Data Embedded in HTML [online]. [Cited July 2020]. www.w3.org/2001/sw/interest/ldh

Heath, T. & Bizer, C. 2011. Linked Data: Evolving the Web into a Global Data Space. Williston, Morgan & Claypool. 136 pp. (also available at http://linkeddatabook.com/editions/1.0)

Hillmann, D. 2001. Using Dublin Core. Simple HTML Examples. [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/usageguide/2001-04-12/simple-html

Hillmann, D. 2005. Using Dublin Core [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/usageguide

Isaac, A., Waites, W., Young, J. & Zeng, M. 2011. Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets [online]. [Cited July 2020]. www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025

ISO 8601. 1988. [Reviewed 2019]. Data elements and interchange formats – Information interchange - Representation of dates and times. [Cited July 2020]

Internet Assigned Numbers Authority. 20xx. Media Types. [online]. [Cited July 2020]. www.iana.org/assignments/media-types/media-types.xhtml

Johnston, P. 2005. Element Refinement in Dublin CoreTM Metadata [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/dc-elem-refine

Johnston, P. 2007. Expressing Dublin Core™ Metadata Using the DC-Text Format [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/dc-text

Johnston, P. & Powell, A. 2008a. Expressing Dublin Core™ Description Sets Using XML (DC-DS-XML) [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/dc-ds-xml

Johnston, P. & Powell, A. 2008b. Expressing Dublin Core™ Metadata Using HTML/XHTML Meta and Link Elements [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/dc-html

Library of Congress (LC). N.d. MARC Code List for Relators Scheme [online]. [Cited July 2020]. https://id.loc.gov/vocabulary/relators.html

Library of Congress (LC). 2017. ISO 639.2: Codes for the Representation of Names of Languages [online]. [Cited July 2020]. www.loc.gov/standards/iso639-2/php/code_list.php

Library of Congress (LC). 2019. Extended Date/Time Format (EDTF) [online]. [Cited July 2020]. www.loc.gov/standards/datetime

Library of Congress (LC). 2020. Catalogue [online]. [Cited July 2020]. https://catalog.loc.gov

Lóscio, B., Burle, C,. & Calegari, N. 2017. Data on the Web Best Practices. [online]. [Cited July 2020]. www.w3.org/TR/dwbp

Michalis, M., & Tzitzikas, Y. (2019). Large-scale Semantic Integration of Linked Data: A Survey. ACM Computing Surveys (CSUR) (52)5: 1–40.

Moss, D.L. 2011. Transatlantic Airline Alliances: The Joint EU-US Report. Journal of European Competition Law & Practice, 2 (4): 391–395.

National Library of Medicine (NLM). 2020. PubMed® [online]. [Cited July 2020]. https://pubmed.ncbi.nlm.nih.gov

Nilsson, M., Powell, A., Johnston, P. & Naeve, A. 2008. Expressing Dublin Core™ Metadata Using the Resource Description Framework (RDF) [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/dc-rdf

OCLC. 2020. 035 Field and OCLC Control Numbers [online]. [Cited July 2020]. https://help.oclc.org/Metadata_Services/WorldShare_Collection_Manager/Choose_your_Collection_Manager_workflow/Data_sync_collections/Prepare_your_data/30035_field_and_OCLC_control_numbers

Powell, A., Nilsson, M., Naeve, A., Johnston, P. & Baker, T. 2007. DCMI Abstract Model [online]. [Cited July 2020]. www.dublincore.org/specifications/dublin-core/abstract-model

Rühle, S., Baker, T., & Johnston, P. n.d. Publishing Metadata [online]. [Cited July 2020]. www.dublincore.org/resources/userguide/publishing_metadata

Sauermann, S. & Cyganiak, R. (W3C). 2008. Cool URIs for the Semantic Web [online]. [Cited July 2020]. www.w3.org/TR/cooluris

UKOLN. 2008. Eprints Terms [online]. [Cited July 2020]. www.ukoln.ac.uk/repositories/digirep/index/Eprints_Terms

van Hooland. S. & Verborgh. R. 2014. Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish your Metadata. Chicago, Neal-Schuman. 282 pp.

VocBench. N.d. VocBench [online]. [Cited July 2020]. http://vocbench.uniroma2.it

Wallis, R. 2014. Linked Data: From Library Entities to the Web of Data. Presentation to the Association for Library Collections and Technical Services session, "International Developments in Library Linked Data: Think Globally" at the American Library Association Conference, June 2014 [online]. [Cited July 2020]. www.slideshare.net/rjw/linked-data-from-library-entities-to-the-web-of-data

Wikipedia. 2020a. Amazon Standard Identification Number [online]. [Cited July 2020]. https://en.wikipedia.org/wiki/Amazon_Standard_Identification_Number

Wikipedia. 2020b. CODEN [online]. [Cited July 2020]. https://en.wikipedia.org/wiki/CODEN

Wikipedia. 2020c. International Standard Book Number [online]. [Cited July 2020]. https://en.wikipedia.org/wiki/International_Standard_Book_Number

Wikipedia. 2020d. International Standard Serial Number [online]. [Cited July 2020]. https://en.wikipedia.org/wiki/International_Standard_Serial_Number

Wikipedia. 2020e. Serial Item and Contribution Identifier [online]. [Cited July 2020]. https://en.wikipedia.org/wiki/Serial_Item_and_Contribution_Identifier

Wikipedia. 2020f. Uniform Resource Identifier [online]. [Cited July 2020]. https://en.wikipedia.org/wiki/Uniform_Resource_Identifier

Wikipedia. 2020g. Universal Product Code [online]. [Cited July 2020]. https://en.wikipedia.org/wiki/Universal_Product_Code

Wilkinson, M.D., Dumontier, M., Aalbersberg, I., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., A. Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. 2016. The FAIR Guiding Principles for Scientific Data Management and Stewardship. Scientific Data 3, 160018 (2016) DOI:10.1038/sdata.2016.18

Woddiscovery. 2010. Linked Open Data Star Scheme by Example [online]. [Cited July 2020]. https://webofdata.wordpress.com/2010/11/12/linked-open-data-star-scheme-by-example

World Wide Web Consortium (W3C). n.d(a). Glossary [online]. [Cited July 2020]. www.w3.org/2003/glossary/keyword/All/literal.html?keywords=literal

World Wide Web Consortium (W3C). n.d(b). Linked Data [online]. [Cited July 2020]. https://www.w3.org/standards/semanticweb/data

World Wide Web Consortium (W3C). n.d(c). Tutorials and Courses [online]. [Cited July 2020]. www.w3.org/2002/03/tutorials#semanticweb_data

World Wide Web Consortium (W3C). 1997. Date and Time Formats [online]. [Cited July 2020]. www.w3.org/TR/NOTE-datetime

World Wide Web Consortium (W3C). 2011. Draft Benefits [online]. [Cited July 2020]. www.w3.org/2005/Incubator/lld/wiki/Draft_Benefits

World Wide Web Consortium (W3C). 2013. PROV-O: The PROV Ontology [online]. [Cited July 2020]. www.w3.org/TR/prov-o

World Wide Web Consortium (W3C). 2014. vCard Ontology – For Describing People and Organizations [online]. [Cited July 2020]. www.w3.org/TR/vcard-rdf

World Wide Web Consortium (W3C). ODRL Vocabulary & Expression 2.2. [online]. [Cited July 2020]. www.w3.org/TR/odrl-vocab

World Wide Web Consortium (W3C). 2020. Data Catalog Vocabulary (DCAT) – Version 2 [online]. [Cited July 2020]. www.w3.org/TR/vocab-dcat-2/#example-landing-page

7 Appendix

Crosswalk of metadata terms used
in LODE BD and Schema.org terms

Appendix

Crosswalk of metadata terms used in LODE-BD and Schema.org terms

Note: This crosswalk table is based on Table 4, Metadata terms used in the LODE-BD guide. The last two columns are added to map those terms further to the Schema.org terms. The prefix is "schema".

For the constraints of the values associated with the terms, please consult Schema.org specification at: schema.org/docs/full.html.

As Schema.org might be updated frequently, when using this crosswalk, please also verify the newest release of Schema.org documentation at schema.org/docs/documents.html.

Updated 2020-09-05

LODE-BD Group	Metadata Terms			schema.org Terms	schema.org Types
	General Metadata Terms		More Specific Terms dcterms: + bibo: + dcat:	Legend	
	dc:-based	dcterms:-based		> narrowMatch < broadMatch ~ relatedMatch	[Dots indicate the level of a sub-type]
1. Title Information	dc:title	dcterms:title		name	Thing
				>headline	. CreativeWork
			dcterms:alternative	alternateName	Thing
				>alternativeHeadline	. CreativeWork
2. Responsible Body	dc:creator	dcterms:creator		creator	. CreativeWork
				>author	. CreativeWork
	dc:contributor	dcterms:contributor		contributor	. CreativeWork
				>editor	. CreativeWork
				>translator	. CreativeWork
				>illustrator	. . Book
			bibo:editor	editor	. CreativeWork
	dc:publisher	dcterms:publisher		publisher	. CreativeWork
			bibo:issuer	<producer	. CreativeWork
			bibo:producer		
			bibo:distributor	<provider	. CreativeWork
			bibo:owner	~ acquiredFrom	. OwnershipInfo
			dcat:contacctPoint	contactPoint	. Person II . Organization
			dcat:qualifiedAttribution		
			dcat:hadRole	>roleName	..Role
3. Physical Characteristics	dc:date	dcterms:date	dcterms:created	dateCreated	. CreativeWork
			dcterms:dateAccepted		
			dcterms:dateCopyrighted	copyrightYear	. CreativeWork
			dcterms:dateSubmitted		
			dcterms:modified	dateModified	. CreativeWork
			dcterms:valid		
			dcterms:available		
			dcterms:issued	date:Published	. CreativeWork

Appendix - continued

LODE-BD Group	Metadata Terms				schema.org Terms	schema.org Types
	General Metadata Terms		More Specific Terms dcterms: + bibo: + dcat:		Legend	
	dc:-based	dcterms:-based			> narrowMatch < broadMatch ~ relatedMatch	[Dots indicate the level of a sub-type]
3. Physical Characteristics	dc:identifier	dcterms:identifier	dcterms:biblographicCitation		<citation	. CreativeWork
			bibo:asin		<identifier	Thing
			bibo:coden			
			bibo:doi		<identifier	Thing
			bibo:eanucc13		gtin13	. Product
			bibo:eissn		<issn	.CreativeWorkSeries
			bibo:gtin14		gtin14	. Product
			bibo:handle			
			bibo:isbn		isbn	. . Book
			bibo:issn		issn	.CreativeWorkSeries
			bibo:lccn			
			bibo:oclcnum			
			bibo:pmid			
			bibo:sici			
			bibo:upc			
			bibo:uri		>url ||< identifier	Thing
	dc:language	dcterms:language			inLanugauge	. CreativeWork
	dc:format	dcterms:format	dcterms:medium		<encodingFormat >associatedMedia	. CreativeWork
			dcterms:extent		> encoding	
			bibo:edition		> bookEdition	. . Book
					> version	. CreativeWork
					> schemaVersion	. CreativeWork
			bibo:status		creativeWorkStatus	. CreativeWork
			bibo:pages		~pagination	. . Article
					>numberOfPages	. . Book
				bibo:pageStart	pageStart	. . Article
				bibo:pageEnd	pageEnd	. . Article
			bibo:section		>articleSection	. . Article
			bibo:volume		volumeNumber	. . PublicationVolume
			bibo:issue		issueNumber	. . PublicationIssue
			bibo:chapter			
			dcat:landingPage		url	Thing
			dcat:accessURL		< url	Thing
4. Holding/ Location			agls: availability	bibo: locator	contentLocation >conditionsOfAccess	. CreativeWork

Appendix - continued

LODE-BD Group	Metadata Terms				schema.org Terms	schema.org Types
	General Metadata Terms		More Specific Terms dcterms: + bibo: + dcat:		Legend	
	dc:-based	dcterms:-based			> narrowMatch < broadMatch ~ relatedMatch	[Dots indicate the level of a sub-type]
5. Subject	dc:subject	dcterms:subject	dcat:theme	dcat: keyword	about	. CreativeWork
					keywords	. CreativeWork
					> mainEntity	. CreativeWork
					> mainEntityOfPage	Thing
	dc:coverage	dcterms:coverage			< mentions	. CreativeWork
			dcterms:spatial		< keywords spatialCoverage temporalCoverage	. CreativeWork
			dcterms:temporal			
6. Description of Content	dc:description	dcterms:description			description	Thing
			dcterms:abstract		abstract	. CreativeWork
			dcterms:tableOfContent			
	dc:type	dcterms:type			< genre	. CreativeWork
					>learningResourceType	. CreativeWork
					~additionalType	Thing
7. Intellectual Property	dc:rights	dcterms:rights	dcterms:rightsHolder		copyrightHolder	. CreativeWork
					>accountablePerson	. CreativeWork
			dcterms:accessRights			
			dcterms:license		license	. CreativeWork
8. Usage	dc:description	dcterms:description			description	Thing
		dcterms:audience			audience	. CreativeWork
					~typicalAgeRange	. CreativeWork
			dcterms:educationLevel		educationalLevel ~educationalUse	. CreativeWork
					~educational Alignment	. CreativeWork
			dcterms:mediator			
		dcterms:instructionalMethod			> interactivityType	. CreativeWork
					~ timeRequired	. CreativeWork

Appendix - continued

LODE-BD Group	Metadata Terms			schema.org Terms	schema.org Types
	General Metadata Terms		More Specific Terms dcterms: + bibo: + dcat:	Legend	
	dc:-based	dcterms:-based		> narrowMatch < broadMatch ~ relatedMatch	[Dots indicate the level of a sub-type]
9. Relation -a [between resources]	dc:relation	dcterms:relation		~ relatedTo	.Product
			dcterms:isVersionOf	~ sameAs	Thing
			dcterms:hasVersion		
			dcterms:isReplacedBy		
			dcterms:replaces		
			dcterms:isRequiredBy		
			dcterms:requires		
			dcterms:isPartOf	isPartOf	. CreativeWork
				~ exampleOfWork	. CreativeWork
			dcterms:hasPart	hasPart	. CreativeWork
				~ workExample	. CreativeWork
			dcterms:isReferencedBy	~ review	. CreativeWork
				> mentions	. CreativeWork
			dcterms:isFormatOf		
			dcterms:hasFormat		
			dcterms:conformsTo		
			dcat:qualifiedRelation		
			dcterms:references	>isBasedOn	. CreativeWork
			bibo:translationOf	translationOfWork	. CreativeWork
			bibo:annotates	> mentions	. CreativeWork
			bibo:citedBy | bibo:cites	citation	. CreativeWork
9. Relation -b [between agents]			eprint:affiliatedInstitution	affiliation	. Person
				>sourceOrganization	. CreativeWork
				>worksFor	. Person
			eprint:grantNumber		
			dcat:hadRole	> sponsor	. Person
					.Organization

The Linked Open Data Enabled Bibliographical Data (LODE-BD) 3.0 guide is designed to provide an overview of how to select appropriate metadata properties to create, manage, and exchange bibliographic information. The strength of LODE-DB are the decision trees designed to facilitate the selection of the appropriate strategies adjustable to data providers according to their local needs.

LODE-BD's objectives include ensuring a set of common metadata properties; encouraging the use of authority data, controlled vocabularies and syntax coding standards; and providing a reference support that is open for suggestions of new properties and metadata terms.

This guide provides recommendations applicable for structured data, describing bibliographic resources such as articles, monographs, theses, conference papers, presentation materials, research reports, learning objects, in print or electronic format.

Food and Agriculture Organization of the United Nations

Rome, Italy

ISBN 978-92-5-133655-7

9 789251 336557

CB2209EN/1/11.20