DataCite - International Data Citation

DataCite Metadata Schema for the Publication and Citation of Research Data

Version 2.2

March 2011

doi:10.5438/0003

Members of the Metadata Working Group

Joan Starr, California Digital Library (head of working group)

Jan Ashton, British Library

Jan Brase, TIB / DataCite

Paul Bracke, Purdue University

Angela Gastl, ETH Zürich

Jacqueline Gillet, Inist

Alfred Heller, DTU Library

Birthe Krog, DTU Library

Lynne McAvoy, CISTI

Karen Morgenroth, CISTI

Elizabeth Newbold, British Library

Madeleine de Smaele, TU Delft

Anja Wilde, GESIS

Scott Yeadon, ANDS

Wolfgang Zenk-Möltgen, GESIS

Frauke Ziedorn, TIB (Metadata Supervisor)

DataCite

Table of Contents

1 Introduction

1.1 The DataCite Consortium

Scholarly research is producing ever increasing amounts of digital research data, and it depends on data to verify research findings, create new research, and share findings. In this context, what has been missing until recently, is a *persistent* approach to access, identification, sharing, and re-use of datasets. To address this need, the DataCite[1] international consortium was founded in late 2009 with these three fundamental goals:

- establish easier access to scientific research data on the Internet,
- increase acceptance of research data as legitimate, citable contributions to the scientific record, and
- support data archiving that will permit results to be verified and re-purposed for future study.

The benefits provided by DataCite are domain agnostic and serve scholars in a range of disciplines, from the sciences, social sciences and humanities.

Key to the DataCite service is the concept of a *persistent identifier.* A persistent identifier is an association between a character string and an object. Objects can be files, parts of files, persons, organizations, abstractions, etc. DataCite uses *Digital Object Identifiers* (DOIs)[2], at the present time and is investigating the use of other identifier schemes in the future. For this reason, the Metadata Kernel has been designed with flexibility and extensibility in mind.

1.2 The Metadata Schema

The DataCite Metadata Schema is a list of core metadata properties chosen for the accurate and consistent identification of data for citation and retrieval purposes, along with recommended use instructions. At a minimum, the mandatory metadata schema properties must be provided at the time of identifier registration. Data centers and other submitters may also choose to use the optional properties to identify their data more clearly. This metadata schema can fulfill several key functions in support of the larger goals of DataCite. Primarily these are:

- recommending a standard citation format for datasets, based on a small number of properties required for identifier registration;
- providing the basis for interoperability with other data management schemas;
- promoting dataset discovery with optional properties allowing for flexible description of the resource, including its relationship to other resources;
- and, laying the groundwork for future services (e.g., discovery) through the use of controlled terms from both a DataCite vocabulary and external vocabularies as applicable. The DataCite vocabularies will be administered by the DataCite Metadata Supervisor who will establish and publicize procedures for submitting changes.

[1] http://www.datacite.org

[2] DOIs are administered by the International DOI Foundation, http://www.doi.org/

This version of the metadata schema has benefited greatly from the comments provided during the community feedback period in the summer of 2010. Many changes too numerous to list here have been made in response to the suggestions. A few of the major adaptations include a reduction in the number of controlled lists and a simplification of those that remain, a radical alteration of the way that a resource's topic or subject can be described, and requiring only those properties needed to support resource citation.

In addition, the Metadata Working Group has two provisions for openness and future extensibility of the schema. The first is the release of a second version of the schema, this one being in a Dublin Core application profile format. This will be implemented by the Metadata Supervisor in a timeframe to be determined. The second is the maintenance of the DataCite controlled lists in separate files from the main schema. This approach not only promotes ease of use but also facilitates sharing the lists with the larger community. The schema maintenance methodology has already been implemented.

1.3 A Note about DataCite DOI registration

The DataCite DOI registration process treats information about the resource's current location (for example, its URL) separate from all other metadata. This is so that updates to the current location are streamlined for the best possible response time and service.

Additionally, one of the main purposes of assigning DOI names (or any persistent identifier) is to separate the location information from any other metadata about a resource. Changeable location information is not considered part of the resource description. Once a resource has been registered with a persistent identifier, the only location information relevant for this resource from now on is that identifier, e.g., http://dx.doi.org/10.xx.

Any subsequent URL-updates would affect only the identifier, not the metadata record.

1.4 Final Thoughts

This metadata schema will be successful if it can fulfill the functions listed above in the simplest manner. It is hoped that the recent changes are steps in this direction. In the fullness of time, DataCite-compliant metadata submissions may originate from widely disparate locations. Therefore, it is highly desirable that the schema be easy to implement, be open to understanding, and be able to adapt to the full extent of academic and research use cases. The schema will be judged in the short-term according to some of these standards, but time will introduce the real tests.

In closing, we'd like to acknowledge the contributions to this work of many colleagues in our institutions who provided us with assistance of all kinds. Their help has been greatly appreciated.

1.5 Version 2.1 Update

Version 2.1 of the DataCite Metadata Schema introduces several changes, as noted below:

- Addition of a namespace (http://datacite.org/schema/namespace) to the schema in order to support OAI PMH compatibility.
- Enforcement of content for mandatory properties.
- New type for the Date property to conform with the specification that it handles both YYYY and YYYY-MM-DD values.

Version 2.1 of the DataCite Metadata Schema documentation includes these changes:

- Addition of a column to the Mandatory and Optional Properties tables (pp. 7-13) providing an indicator of whether the property being described is an attribute or a child of the corresponding property that has preceded it.

- Revision of the allowed values description for the attribute 12.2 relationType. These have been reviewed and rewritten for increased clarity. In several cases, corrections to the definitions have occurred. The numerous changes are found on page 12 of this document.

1.6 Version 2.2 Update

Version 2.2 of the DataCite Metadata Schema introduces several changes, as noted below:

- Addition of "URL" to list of allowed values for relatedIdentifierType.
- Addition of the following values to list of allowed values for contributorType: Producer, Distributor, RelatedPerson, Supervisor, Sponsor, Funder, RightsHolder.
- Addition of "SeriesInformation" to list of allowed values for descriptionType.
- Addition of "Model" to list of allowed values for resourceTypeGeneral.

Version 2.2 of the DataCite Metadata Schema documentation includes these changes:

- Provision of more examples of xml for different types of objects.
- Explanation of the PublicationYear property in consideration of the requirements of citation.
- A change to the definition of the Publication property, which now reads, "The name of the entity that holds, archives, publishes, prints, distributes, releases, issues, or produces the resource. This property will be used to formulate the citation, so consider the prominence of the role."

2 DataCite Metadata Properties

2.1 Overview

The tables below display in a simple manner the mandatory and optional properties. Also shown are two administrative metadata properties which are not submitted, but instead are supplied at the time of registration with the managing agent for DataCite[3]. More detail about these properties is provided in following sections of this document. Please note that in this document, the resource that is being described can be of any kind, but it is typically a dataset. We use the term 'dataset' in its broadest sense. We mean by it to include not only numerical data, but any other research data outputs.

Please note that DataCite reserves the right to share metadata with information indexes and other entities.

The properties listed in Table 1 *must be* supplied when submitting DataCite metadata. The optional properties listed in Table 2 *may be* supplied when submitting DataCite metadata.

Table 1: DataCite Mandatory Properties

ID	*Property*
1	Identifier (with type attribute)

[3]At the present time the Technische Informationsbibliothek (TIB) is the managing agent for DataCite.

2	Creator (with name identifier attributes)
3	Title (with optional type attribute)
4	Publisher
5	PublicationYear

Table 2: DataCite Optional Properties

ID	*Property*
6	Subject (with schema attribute)
7	Contributor (with type and name identifier attributes)
8	Date (with type attribute)
9	Language
10	ResourceType (with description attribute)
11	AlternateIdentifier (with type attribute)
12	RelatedIdentifier (with type and relation type attributes)
13	Size
14	Format
15	Version
16	Rights
17	Description (with type attribute)

In addition to the metadata that submitters supply with registrations and updates, there are two administrative metadata properties that the managing agency will assign to each DataCite metadata record, shown in Table 3. These properties convey the date on which the metadata description was stored by DataCite (LastMetadataUpdate) and a sequence number assigned to the metadata description by DataCite (MetadataVersionNumber).

Table 3: DataCite Administrative Metadata

ID	*Property*
0.1	LastMetadataUpdate
0.2	MetadataVersionNumber

2.2 Citation

Because many users of this schema are members of a variety of academic disciplines, DataCite remains discipline-agnostic concerning matters pertaining to academic style sheet requirements. Therefore, DataCite recommends rather than requires a particular citation format. In keeping with

this approach, the following is the *recommended* format for rendering a DataCite citation for human readers using the first five properties of the schema:

Creator (PublicationYear): Title. Publisher. Identifier

It may also be desirable to include information from two optional properties, Version and ResourceType (as appropriate). If so, the recommended form is as follows:

Creator (PublicationYear): Title. Version. Publisher. ResourceType. Identifier

For citation purposes, the Identifier may optionally appear both in its original format and in a linkable, http format, as it is practiced by the Organisation for Economic Co-operation and Development (OECD), as shown below.

Regarding the PublicationYear, DataCite recommends, for resources that do not have a standard publication year value, to submitt the date that would be preferred from a citation perspective.

Here are several examples:

- Irino, T; Tada, R (2009): Chemical and mineral compositions of sediments from ODP Site 127-797. Geological Institute, University of Tokyo.doi:10.1594/PANGAEA.726855. http://dx.doi.org/10.1594/PANGAEA.726855
- Geofon operator (2009): GEFON event gfz2009kciu (NW Balkan Region). GeoForschungsZentrum Potsdam (GFZ). doi:10.1594/GFG.GEOFON.gfz2009kciu. http://dx.doi.org/10.1594/GFZ.GEOFON.gfz2009kciu
- Denhard, Michael (2009): dphase_mpeps: MicroPEPS LAF-Ensemble run by DWD for the MAP D-PHASE project. World Data Center for Climate. doi: 10.1594/WDCC/dphase mpeps. http://dx.doi.org/10.1594/WDCC/dphase_mpeps

2.3 DataCite Mandatory Properties

The table below provides a detailed description of the mandatory properties which must be supplied with any initial metadata submission to the managing agent for DataCite. For an example of how to make a submission in XML format, please see the XML Example provided at the end of this document.

A naming convention has been used for all properties and attributes as follows: properties begin with a capital letter, whereas attributes and child properties begin with a lower case letter. If the name of the element is a compound of more than one word, subsequent words begin with capital letters.

The indicator A/C shows whether the property being described is an attribute (A) or a child (C) of the corresponding property that has preceded it.

The attribute **Occ**urrence explains if a property can have multiple instances, which is indicated by the notation: Occ.1-n, meaning that a property must occur once (1), and may occur multiple times (n). A notation of "Req" indicates that an attribute is required if the corresponding property is applied; "Opt" indicates that it is optional.

Where the Allowed values specify a "Controlled List," the members of the list are governed by the DataCite Metadata Supervisor. The Metadata Supervisor will establish and publicize procedures for submitting changes.

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
1	Identifier	The Identifier is a unique string that identifies a resource.		1	DOI (Digital Object Identifier) registered by a DataCite member. Format should be "10.1234/foo"
1.1	identifierType	The type of the Identifier.	A	Req	*Controlled List* Allowed values: DOI
2	Creator	The main researchers involved in producing the data, or the authors of the publication, in priority order.		1-n	May be a corporate/institutional or personal name. The personal name format should be: family, given. Non-roman names should be transliterated according to the ALA-LC schemes found at http://www.loc.gov/catdir/cpso/roman.html
2.1	creatorName	The name of the creator.	C	1	Examples: Toru, Nozawa; Utor, Awazon

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
2.2	nameIdentifier	Uniquely identifies an individual or legal entity, according to various schemes.	C	0-1	The format is dependent upon scheme.
2.2.1	nameIdentifierScheme	The name and/or the URL of the name identifier scheme.	A	Req	Examples are Open Researcher and Contributer ID[4] (ORCID),International Standard Name Identifier[5] (ISNI)
3	Title	A name or title by which a resource is known.		1-n	The format is open.
3.1	titleType	The type of Title.	A	Opt	*Controlled List* Allowed values: Alternative Title Subtitle TranslatedTitle
4	Publisher	The name of the entity that holds, archives, publishes, prints, distributes, releases, issues, or produces the resource. This property will be used to formulate the citation, so consider the prominence of the role.		1	Examples: World Data Center for Climate (WDCC); GeoForschungsZentrum Potsdam (GFZ); Geological Institute, University of Tokyo *** In the case of datasets, "publish" is understood to mean making the data available to the community of researchers.

[4] http://www.orcid.org/
[5] http://www.isni.org/

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
5	PublicationYear	The year when the data was or will be made publicly available.		1	YYYY *** If an embargo period has been in effect, use the date when the embargo period ends. If there is no standard publication year value, use the date that would be preferred from a citation perspective.

2.4 DataCite Optional Properties

The table below provides a detailed description of the optional properties which may be supplied with any metadata submission to the managing agent for DataCite. For an example of how to make a submission in XML format, please see the XML Example provided at the end of this document.

A naming convention has been used for all properties and attributes as follows: properties begin with a capital letter, whereas attributes and child properties begin with a lower case letter. If the name of the element is a compound of more than one word, subsequent words begin with capital letters.

The indicator A/C shows whether the property being described is an attribute (A) or a child (C) of the corresponding property that has preceded it.

The attribute **Occ**urrence explains if a property can have multiple instances, which is indicated by the notation: Occ. 0-n, meaning that a property can occur never (0), once, or multiple times (n). A notation of "Req" indicates that an attribute is required if the parent property is applied; "Opt" indicates that it is optional.

Where the Allowed values specify a "Controlled List", the members of the list are governed by the DataCite Metadata Supervisor. The DataCite Metadata Supervisor will establish and publicize procedures for submitting changes.

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
6	Subject	Subject, keyword, classification code, or key phrase describing the resource.		0-n	The format is open.
6.1	subjectScheme	The name and/or URL of the Subject scheme or classification code, if one is used.	A	Opt	The format is open.
7	Contributor	The institution or person responsible for collecting, creating, or otherwise contributing to the development of the dataset.		0-n	The personal name format should be: family, given. Non-roman names should be transliterated according to the ALA-LC schemes found at http://www.loc.gov/catdir/cpso/roman.html

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
7.1	contributorType	The type of contributor of the resource.	A	Req	*Controlled List* Allowed values: ContactPerson DataCollector DataManager Distributor Editor Funder HostingInstitution Producer ProjectLeader ProjectMember RegistrationAgency RegistrationAuthority RelatedPerson Researcher RightsHolder Sponsor Supervisor WorkPackageLeader
7.2	contributorName	The name of the contributor.	C	1	Examples: Patel, Emily; Doe, John
7.3	nameIdentifier	Uniquely identifies an individual or legal entity, according to various schemes.	C	0-1	The format is dependent upon scheme.
7.3.1	nameIdentifierScheme	The name and/or the URL of the name identifier scheme.	A	Req	Examples are ORCID, ISNI
8	Date	Different dates relevant to the work.		0-n	YYYY or YYYY-MM-DD or any other format described in W3CDTF.[6] *** May be repeated to indicate a date range.

[6] http://www.w3.org/TR/NOTE-datetime

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
8.1	dateType	The type of date.	A	Req	*Controlled List* Allowed values: Accepted (The date that the publisher accepted the resource into their system.) Available (The date the resource is made publicly available. May be a range.) Copyrighted (The specific, documented date at which the resource receives a copyrighted status, if applicable.) Created (The date the resource itself was put together; this could be a date range or a single date for a final component, e.g., the finalised file with all of the data.) EndDate (Use if any other date type covers a range) Issued (The date that the resource is published or distributed e.g. to a data center.) StartDate (Use if any other date type covers a range) Submitted (The date the creator submits the resource to the publisher. This could be different from Accepted if the publisher then applies a selection process.) Updated (The date of the last update to the resource, when the resource is being added to. May be a range.) Valid (The date or date range during which the dataset or resources are accurate. May be a range.) *** To indicate a date period, provide two dates, specifying the StartDate and the EndDate. To indicate the end of an embargo period, use Available. To indicate the start of an embargo period, use Submitted or Accepted, as appropriate.
9	Language	The primary language of the resource.		0-1	*Controlled Vocabulary* Allowed values from: ISO 639-2/B, ISO 639-3 Examples: eng, fre, ger
10	ResourceType	A description of the resource.		0-1	The format is open, but the preferred format is a single term of some detail so that a pair can be formed with the attribute. *** Example: Image/Animation, where 'Image' is resourceTypeGeneral value and 'Animation' is ResourceType value.

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
10.1	resourceTypeGeneral	The general type of a resource.	A	Req	*Controlled List* Allowed values: Collection Dataset Event Film Image InteractiveResource Model PhysicalObject Service Software Sound Text
11	AlternateIdentifier	An identifier or identifiers other than the primary Identifier applied to the resource being registered. This may be any alphanumeric string which is unique within its domain of issue.		0-n	The format is open.
11.1	alternateIdentifierType	The type of the AlternateIdentifier.	A	Req	The format is open.
12	RelatedIdentifier	Identifiers of related resources.		0-n	The format is open. *** Use this property to indicate subsets of properties, as appropriate.
12.1	relatedIdentifierType	The type of the RelatedIdentifier.	A	Req	*Controlled List* Allowed values: ARK DOI EAN13 EISSN Handle ISBN ISSN ISTC LISSN LSID PURL UPC URL URN

ID	DataCite-Property	Definition	A/C	Occ	Allowed values, examples, other constraints
12.2	relationType	Description of the relationship of the resource being registered (A) and the related resource (B).	A	Req	*Controlled List* Allowed values: IsCitedBy (indicates that B includes A in a citation) Cites (indicates that A includes B in a citation) IsSupplementTo (indicates that A is a supplement to B) IsSupplementedBy (indicates that B is a supplement to A) IsContinuedBy (indicates A is continued by the work B) Continues (indicates A is a continuation of the work B) IsNewVersionOf (indicates A is a new edition of B, where the new edition has been modified or updated) IsPreviousVersionOf (indicates A is a previous edition of B) IsPartOf (indicates A is a portion of B; may be used for elements of a series) HasPart (indicates A includes the part B) IsReferencedBy (indicates A is used as a source of information by B) References (indicates B is used as a source of information for A) IsDocumentedBy (indicates B is documentation about/explaining A) Documents (indicates A is documentation about/explaining B) isCompiledBy (indicates B is used to compile or create A) Compiles (indicates B is the result of a compile or creation event using A) IsVariantFormOf (indicates A is a variant or different form of B, e.g. calculated or calibrated form or different packaging) IsOriginalFormOf (indicates A is the original form of B)
13	Size	Unstructured size information about the resource.		0-n	The format is open. *** Examples: "15 pages", "6 MB"
14	Format	Technical format of the resource.		0-n	The format is open. *** Use file extension or MIME type where possible, e.g., PDF, XML, MPG or application/pdf, text/xml, video/mpeg.
15	Version	The version number of the resource.		0-1	If the primary resource has changed, the version number increases. Register a new DOI (or primary identifier) when the version of the resource changes to enable the citation of the exact version of a research dataset (or other resource). May be used in conjunction with properties 11 and 12 (AlternateIdentifier and RelatedIdentifier) to indicate various information updates.

ID	*DataCite-Property*	*Definition*	*A/C*	*Occ*	*Allowed values, examples, other constraints*
16	Rights	Any rights information for this resource.		0-1	The format is open. *** Provide a rights management statement for the resource or reference a service providing such information. Include embargo information if applicable.
17	Description	All additional information that does not fit in any of the other categories.		0-n	The format is open. *** It is a best practice to supply a description.
17.1	descriptionType	The type of the description.	A	Req	*Controlled List* Allowed values: Abstract SeriesInformation TableOfContents Other *** Use the type SeriesInformation when supplying the description of a resource that is part of a series.

3 XML Example

This XML example conforms to the XML schema. More examples for various object types can be found at http://schema.datacite.org/meta/kernel-2.2/index.html.

```
<resource xmlns="http://datacite.org/schema/kernel-2.2"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://datacite.org/schema/kernel-2.2 http://schema.datacite.org/meta/kernel-
2.2/metadata.xsd">
        <identifier identifierType="DOI">10.1594/WDCC/CCSRNIES_SRES_B2</identifier>
        <creators>
                <creator>
                        <creatorName>Miller, John</creatorName>
                </creator>
                <creator>
                        <creatorName>Smith, Jane</creatorName>
                        <nameIdentifier nameIdentifierScheme="ISNI">1422 4586 3573
0476</nameIdentifier>
                </creator>
        </creators>
        <titles>
                <title>National Institute for Environmental Studies and Center for Climate System
Research Japan</title>
                <title titleType="Subtitle">A survey</title>
        </titles>
        <publisher>World Data Center for Climate (WDCC)</publisher>
        <publicationYear>2004</publicationYear>
        <subjects>
                <subject>Earth sciences and geology</subject>
                <subject subjectScheme="DDC">551 Geology, hydrology, meteorology</subject>
        </subjects>
        <contributors>
                <contributor contributorType="DataManager">
                        <contributorName>PANGAEA</contributorName>
                </contributor>
                <contributor contributorType="ContactPerson">
                        <contributorName>Doe, John</contributorName>
                        <nameIdentifier nameIdentifierScheme="ORCID">xyz789</nameIdentifier>
                </contributor>
        </contributors>
        <dates>
                <date dateType="Valid">2005-04-05</date>
                <date dateType="Accepted">2005-01-01</date>
        </dates>
        <language>eng</language>
        <resourceType resourceTypeGeneral="Image">Animation</resourceType>
        <alternateIdentifiers>
                <alternateIdentifier alternateIdentifierType="ISBN">937-0-1234-56789-
X</alternateIdentifier>
        </alternateIdentifiers>
        <relatedIdentifiers>
                <relatedIdentifier relatedIdentifierType="DOI"
relationType="IsCitedBy">10.1234/testpub</relatedIdentifier>
                <relatedIdentifier relatedIdentifierType="URN"
relationType="Cites">http://testing.ts/testpub</relatedIdentifier>
        </relatedIdentifiers>
        <sizes>
```

```
                <size>285 kb</size>
                <size>100 pages</size>
        </sizes>
        <formats>
                <format>text/plain</format>
        </formats>
        <version>1.0</version>
        <rights>Open Database License [ODbL]</rights>
        <descriptions>
                <description descriptionType="Other">
                The current xml-example for a DataCite record is the official example from the documentation.
                <br/>Please look on datacite.org to find the newest versions of sample data and schemas.
                </description>
        </descriptions>
</resource>
```

4 XML Schema

The XML Schema is available here:

http://schema.datacite.org/meta/kernel-2.1/metadata.xsd

doi:10.5438/0004

Note that the schema and this documentation will always have the same version number.

Each subsequent version of the schema will be at this same location using an address composed in the same manner, that is: http://schema.datacite.org/meta/kernel-versionnumber/metadata.xsd.

Earlier versions will continue to be available at their previous locations for backward compatibility.

DataCite

Appendices

This appendix to the DataCite Metadata Schema contains for mappings to four metadata element sets. The DataCite Metadata Working Group offers these initial mappings as a service to the community, but will not be responsible for maintaining them on an ongoing basis.

Dublin Core Mapping

The table below provides a mapping of the DataCite properties to the Dublin Core Simple elements and Qualified terms.[7]

ID	DataCite-Property	Dublin Core Simple Mapping (elements namespace)	Dublin Core Qualified Mapping (terms namespace)
1	Identifier	dc:identifier	dcterms:identifier
1.1	identifierType	dc:identifier	dcterms:identifier
2	Creator	dc:creator	dcterms:creator
2.1	creatorType	dc:creator	dcterms:creator
2.2	nameIdentifier	Not present in Dublin Core	Not present in Dublin Core
2.2.1	nameIdentifierScheme	Not present in Dublin Core	Not present in Dublin Core
3	Title	dc:title	dcterms:title
3.1	titleType	dc:title	dcterms:alternative
4	Publisher	dc:publisher	dcterms:publisher
5	PublicationYear	dc:date	dcterms:issued
6	Subject	dc:subject	dcterms:subject
6.1	subjectScheme	Not present in Dublin Core	Not present in Dublin Core
7	Contributor	dc:contributor	dcterms:contributor
7.1	contributorType	dc:contributor	dcterms:contributor
7.2	contributorName	dc:contributor	dcterms:contributor
7.3	nameIdentifier	Not present in Dublin Core	Not present in Dublin Core
7.3.1	nameIdentifierScheme	Not present in Dublin Core	Not present in Dublin Core
8	Date	dc:date	dcterms:date

[7] http://www.dublincore.org/documents/dcmi-terms/

ID	DataCite-Property	Dublin Core Simple Mapping (elements namespace)	Dublin Core Qualified Mapping (terms namespace)
8.1	dateType	dc:date	dcterms:available dcterms:created dcterms:dateAccepted dcterms:dateCopyrighted dcterms:dateSubmitted dcterms:modified
8.1	dateType (for StartDate/EndDate)	dc:coverage	dcterms:temporal
9	Language	dc:language	dcterms:language
10	ResourceType	dc:type	dcterms:type
10.1	resourceTypeGeneral	dc:type	dcterms:type
11	AlternateIdentifier	dc:identifier	dcterms:identifier
11.1	alternateIdentifierType	dc:identifier	dcterms:identifier
12	RelatedIdentifiers	dc:relation + dc:identifier	dcterms:relation + dcterms:identifier
12.1	relatedIdentifierType	dc:relation + dc:identifier	dcterms:relation + dcterms:identifier
12.2	relationType	dc:relation	dcterms:relation dcterms:conformsTo: dcterms:isReferencedBy dcterms:references dcterms:isVersionOf dcterms:hasVersion dcterms:isFormatOf dcterms:hasFormat dcterms:isPartOf dcterms:hasPart dcterms:isReplacedBy dcterms:replaces dcterms:source

ID	DataCite-Property	Dublin Core Simple Mapping (elements namespace)	Dublin Core Qualified Mapping (terms namespace)
13	Size	dc:format	dcterms:extent
14	Format	dc:format	dcterms:format
15	Version	Not present in Dublin Core	Not present in Dublin Core
16	Rights	dc:rights	dcterms:rights
17	Description	dc:description	dcterms:description
17.1	descriptionType	dc:description	dcterms:abstract dcterms:tableOfContents

IDF Metadata Kernel Mapping

The table below provides a mapping of the DataCite properties to the International DOI Foundation (IDF) Metadata Kernel.[8]

ID	*DataCite-Property*	*IDF Metadata Kernel*
1	Identifier	DOI Name
2	Creator	principalAgent + agentName + agentRole (with value 'author')
2.1	creatorName	
2.2	nameIdentifier	+ agentIdentifier
2.2.1	nameIdentifierScheme	+ type
3	Title	resourceName
3.1	titleType	resourceNameType with value: principalTitle, title, alternativeTitle, subtitle, translatedTitle.
4	Publisher	principalAgent + agentName + agentRole (with value 'publisher')
5	PublicationDate	Not present in IDF
6	Subject	Not present in IDF
6.1	subjectScheme	Not present in IDF
7	Contributor	principalAgent + agentName
7.1	contributorType	+ agentRole (with relevant value from the controlled list e.g. hostingInstitution).
7.2	contributorName	principalAgent + agentName
7.3	nameIdentifier	+ agentIdentifier
7.3.1	nameIdentifierScheme	+ type
8	Date	Not present in IDF
8.1	dateType	Not present in IDF
9	Language	resourceName + primaryLanguage
10	ResourceType	resourceType
10.1	resourceTypeGeneral	Not present in IDF
11	AlternateIdentifier	resourceIdentifier
11.1	AlternateIdentifierType	+ type
12	RelatedIdentifiers	Not present in IDF
12.1	relatedIdentifierType	Not present in IDF

[8]The DOI® Handbook, Edition 4.4.1, 2006, http://dx.doi.org/10.1000/186, Chapter 4.3.1

ID	*DataCite-Property*	*IDF Metadata Kernel*
12.2	relationType	Not present in IDF
13	Size	Not present in IDF
14	Format	Not present in IDF
15	Version	Not present in IDF
16	Rights	Not present in IDF
17	Description	Not present in IDF
17.1	descriptionType	Not present in IDF
	Not present in DataCite Schema	structuralType
	Not present in DataCite Schema	mode
	Not present in DataCite Schema	RegistrationAgency

OECD Mapping

The table below provides a mapping of the DataCite Properties to the proposed OECD Dataset metadata elements. This included OECD elements mentioned in the Dataset Annex.[9]

ID	*DataCite Properties*	*OECD Dataset Metadata*
1	Identifier	Unique, persistent, global identifier, e.g. DOI and DOI number
2	Creator	Author(s)
2.1	creatorName	Author
2.2	nameIdentifier	Not present in OECD
2.2.1	nameIdentifierScheme	Not present in OECD
3	Title	Main Title Subtitle
3.1	titleType	Main Title Subtitle
4	Publisher	Is Copyrighted By
5	PublicationYear	Publication date
6	Subject	Classification, e.g. JEL Classification, Themes
6.1	subjectScheme	Classification, e.g. JEL Classification, Themes
7	Contributor	Is imprinted by, Is edited by
7.1	contributorType	Not present in OECD
7.2	contributorName	Is imprinted by, Is edited by
7.2	nameIdentifier	Not present in OECD
7.2.1	nameIdentifierScheme	Not present in OECD
8	Date	Next publication date (3 fields), Time range
8.1	dateType	Period covered: Start Year Period covered: End Year
9	Language	Languages
10	ResourceType	Not present in OECD (assumes dataset)
10.1	resourceTypeGeneral	Not present in OECD
11	AlternateIdentifier	ISSN, Dataset Code
11.1	alternateIdentifierType	Not present in OECD

[9] Green, Toby (2009): "We Need Publishing Standards for Datasets and Data Tables," OECD Publishing White Paper, OECD Publishing. http://dx.doi.org/10.1787/603233448430

ID	*DataCite Properties*	*OECD Dataset Metadata*
12	RelatedIdentifiers	External links
12.1	relatedIdentifierType	External links
12.2	relationType	Links: Belongs to Has main parent Is related to Has external links Has renditions Supersedes Is continued by
13	Size	Size
14	Format	Has physical form
15	Version	Not present in OECD
16	Rights	Not present in OECD
17	Description	Keyword(s)
17.1	descriptionType	Short abstract Long abstract
	Not present in DataCite Schema	Authors (order)
	Not present in DataCite Schema	Periodicity
	Not present in DataCite Schema	Countries covered
	Not present in DataCite Schema	Update method
	Not present in DataCite Schema	Related countries
	Not present in DataCite Schema	Variable index

DDI Mapping

The table below provides a mapping of the DataCite properties to the DDI 3.1 elements.[10] All DDI elements are included in the <ddi:DDIInstance> element.[11]

ID	*DataCite-Property*	*DDI 3.1 Mapping*	*Used in*
1	Identifier	<r:InternationalIdentifier type="DOI"> OR <r:UserID type="DOI">	<s:StudyUnit> <r:Citation> <pi:PhysicalInstance> <pi:DataFileIdentification>
2	Creator	<r:Creator> OR <a:DisplayName> (use together with <a:Role> <a:IndividualReference> <r:Description>Creator)	<s:StudyUnit> <r:Citation> <a:Archive> <a:OrganizationScheme> <a:Individual>
2.1	creatorName	<r:Creator> OR <a:DisplayName> (use together with <a:Role> <a:IndividualReference> <r:Description>Creator)	<s:StudyUnit> <r:Citation> <a:Archive> <a:OrganizationScheme> <a:Individual>
2.2	nameIdentifier	<a:ResearcherID> <a:Identifier>	<a:Archive> <a:OrganizationScheme> <a:Individual>

[10] http://www.ddialliance.org/specification/ddi3.1

[11] The following namespaces are used:
xmlns:ddi="ddi:instance:3_1"
xmlns:s="ddi:studyunit:3_1"
xmlns:pd="ddi:physicaldataproduct:3_1"
xmlns:pi="ddi:physicalinstance:3_1"
xmlns:c="ddi:conceptualcomponent:3_1"
xmlns:l="ddi:logicalproduct:3_1"
xmlns:r="ddi:reusable:3_1"
xmlns:dc="ddi:datacollection:3_1"
xmlns:a="ddi:archive:3_1"
xmlns:xhtml="http://www.w3.org/1999/xhtml"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:dcore="ddi:dcelements:3_1"
xmlns:dcore2="http://purl.org/dc/elements/1.1/"

ID	*DataCite-Property*	*DDI 3.1 Mapping*	*Used in*
2.2.1	nameIdentifierScheme	<a:ResearcherID> <a:Type>	<a:Archive> <a:OrganizationScheme> <a:Individual>
3	Title	<r:Title> OR (depending on titleType) <r:SubTitle> <r:AlternateTitle> <r:AlternateTitle translated=”true”>	<s:StudyUnit> <r:Citation>
3.1	titleType	(defines mapping of Title)	
4	Publisher	<r:Publisher>	<s:StudyUnit> <r:Citation>
5	PublicationYear	<r:PublicationDate> <r:SimpleDate>	<s:StudyUnit> <r:Citation>
6	Subject	<r:UserID> OR <r:Subject> OR <r:Keywords>	<s:StudyUnit> <r:Coverage> <r:TopicalCoverage>
6.1	subjectScheme	<r:UserID type=””> OR <r:Subject codeListID=““> OR <r:Keywords codeListID=““>	
7	Contributor	<r:Contributor>	<s:StudyUnit> <r:Citation>
7.1	contributorType	<r:Contributor role="">	
7.2	contributorName	<r:Contributor>	<s:StudyUnit> <r:Citation>
7.3	nameIdentifier	<a:ResearcherID> <a:Identifier>	<a:Archive> <a:OrganizationScheme> <a:Individual>
7.3.1	nameIdentifierScheme	<a:ResearcherID> <a:Type>	<a:Archive> <a:OrganizationScheme> <a:Individual>
8	Date	<r:Date> <r:SimpleDate> OR <r:StartDate> OR <r:EndDate>	<s:StudyUnit> <r:LifecycleInformation> <r:LifecycleEvent> OR <s:StudyUnit> <dc:DataCollection> <dc:CollectionEvent> <dc:DataCollectionDate>
8.1	dateType	<r:Description>	<s:StudyUnit> <r:LifecycleInformation> <r:LifecycleEvent>
9	Language	<r:Language>	<s:StudyUnit> <r:Citation>

ID	*DataCite-Property*	*DDI 3.1 Mapping*	*Used in*
10	ResourceType	<dcore:DCElements> <dcore2:type>	<s:StudyUnit> <r:Citation>
10.1	resourceTypeGeneral	<dcore:DCElements> <dcore2:description>	<s:StudyUnit> <r:Citation>
11	AlternateIdentifier	<r:UserID> OR <a:CallNumber>	<s:StudyUnit> <a:Archive> <a:ArchiveSpecific> <a:Item>
11.1	alternateIdentifierType	<r:UserID type=””>	
12	RelatedIdentifiers	<r:UserID>	<s:StudyUnit> <r:OtherMaterial>
12.1	relatedIdentifierType	<r:UserID type=””>	
12.2	relationType	<r:Relationship> <r:RelationshipDescription>	<s:StudyUnit> <r:OtherMaterial>
13	Size	<r:Content> OR (if unit “cases” is known) <pi:CaseQuantity> OR (if unit “datafile” is known) <a:DataFileQuantity>	<s:StudyUnit> <r:Note> <pi:PhysicalInstance> <pi:GrossFileStructure> < a:Archive> <a:ArchiveSpecific> <a:Item>
14	Format	<pd:Format> OR <a:Format>	<s:StudyUnit> <pd:PhysicalDataProduct> <pd:PhysicalStructureScheme> <pd:PhysicalStructure> < a:Archive> <a:ArchiveSpecific> <a:Item>
15	Version	IF Version syntax is like n.n.n: <pi:PhysicalInstance version=””> ELSE <pi:PhysicalInstance> <r:VersionRationale>	 (May be misleading, because VersionRationale should contain the versioning reason, not the version number)