

# DataCite Best Practice Guide

Sonja Kümmer, Universitätsbibliothek der LMU  
([sonja.kuemmet@ub.uni-muenchen.de](mailto:sonja.kuemmet@ub.uni-muenchen.de) <https://orcid.org/0000-0002-8954-0200>)  
Stephan Lücke, VerbaAlpina and IT-Gruppe Geisteswissenschaften (LMU)  
([luecke@lmu.de](mailto:luecke@lmu.de) <https://orcid.org/0000-0002-5853-1918>)  
Julian Schulz, IT-Gruppe Geisteswissenschaften (LMU)  
([julian.schulz@lmu.de](mailto:julian.schulz@lmu.de) <https://orcid.org/0000-0003-4374-2680>)  
Martin Spenger, Universitätsbibliothek der LMU  
([martin.spenger@ub.uni-muenchen.de](mailto:martin.spenger@ub.uni-muenchen.de) <https://orcid.org/0000-0002-8841-5985>)  
Tobias Weber, Leibniz Supercomputing Centre  
([mail@tgweber.de](mailto:mail@tgweber.de) <https://orcid.org/0000-0003-1815-7041>)

Version 1.0 (15.11.2019)

## Contents

<b>Overview</b>	<b>2</b>
<b>General Best Practice</b>	<b>2</b>
What do the metadata describe? . . . . .	2
What is the language of the metadata? . . . . .	3
How should I specify an institution? . . . . .	3
How should I specify a person? . . . . .	4
How should I handle different versions of the same research data? . . . . .	4
<b>Best practice for specific fields</b>	<b>5</b>
Identifier . . . . .	5
Title . . . . .	5
Creator . . . . .	5
Publisher . . . . .	6
PublicationYear . . . . .	6
ResourceType . . . . .	7
Subject . . . . .	7
Mandatory subject annotations . . . . .	8
Geotagging . . . . .	8
Additional subject annotations . . . . .	8
Contributor . . . . .	9
Date . . . . .	11
RelatedIdentifier . . . . .	11
Description . . . . .	15
GeoLocation . . . . .	15

Language . . . . .	17
AlternateIdentifier . . . . .	17
Size . . . . .	17
Format . . . . .	17
Version . . . . .	18
Rights . . . . .	18
FundingReference . . . . .	19
<b>Examples</b>	<b>19</b>
Bayerisches Musiker-Lexikon Online (BMLO) . . . . .	19
VerbaAlpina . . . . .	20
ClimEx . . . . .	20

This document is a guideline to use the [DataCite Metadata Schema \[external link\]](#), version 4.3.

It is meant for researchers, IT and library support staff in the context of the following institutions:

- Ludwig-Maximilians-Universität München
- Leibniz Supercomputing Centre

If you wish to publish your research data in one of these institutions, this guide helps to make your submission compliant to best practices in the context of research data management.

This guide is also designed to be reused by other institutions as well.

To create the DataCite-XML file for the project you want to describe, we recommend to you to use our [DataCite-Generator \[external link\]](#). This tool is kept in sync with this guideline, safe for transmission times inbetween versions. If you want to create metadata for research data on a scale that is too large for manual procedures, please contact one of the institutions named above.

## Overview

The first part, **General Best Practice**, is a selection of recommendations and obligations when using DataCite in general. It has been written in an FAQ-style (Frequently Asked Questions).

The second part, **Best Practice for specific fields**, gives more details for each of the 19 metadata fields of the DataCite metadata standard. Note that the following fields are optional/recommended in the DataCite standard, but mandatory according to this guide: **Subject**, **Description**, and **Rights**.

The third part consists of four complete examples.

## General Best Practice

### What do the metadata describe?

Unless otherwise specified all information in the metadata concerns the research data (also denoted as “resource”), neither the project in whose context the data have been created or collected nor the metadata themselves.

## What is the language of the metadata?

- The default language of the metadata is English. If another language is used, the same information must additionally be specified in English.
- Where language variations are possible (e.g. title, description, affiliations), the language should be specified by `xml:lang` attributes:

```
<title xml:lang="de">
    Bayerisches Musiker-Lexikon Online (BMLÖ)
</title>
<title xml:lang="en" titleType="TranslatedTitle">
    Digital Encyclopedia of Bavarian Musicians
</title>
```

- Proper nouns need not be translated.
- Use standardized data (e.g. controlled vocabularies) when possible. This might allow data aggregators to display the information in the language most suitable to the use case at hand.

## How should I specify an institution?

- Follow the policy of the institution (as local as possible).
- If no policy is findable or multiple names in multiple languages are given, use the English name in doubt.
- In any case specify the language in which the name is given as *xml:lang* tag:

```
<publisher xml:lang="en">Leibniz Supercomputing Centre</publisher>
```

- State the name of the institution as specific as possible (e.g. start with the chair/group, not with the university).
- Start with the more specific organisational units first and end with the most generic unit, separated by comma:

```
<affiliation xml:lang="de">
    Institut für Romanische Philologie, Ludwig-Maximilians-Universität München
</affiliation>
```

- Affiliations are specified as appropriate at the time of creation of the resource.
- When possible, specify a persistent identifier (PID) for the institution (Research Organization Registry (ROR [\[external link\]](#)), if there is no entry in ROR: International Standard Name Identifier (ISNI [\[external link\]](#))).

```
<fundingReference>
    <funderName>Deutsche Forschungsgemeinschaft (DFG)</funderName>
    <funderIdentifier
        funderIdentifierType="ROR">
        https://ror.org/018mejw64
    </funderIdentifier>
```

...  
</fundingReference>

## How should I specify a person?

- State the name as specified in the standard: family name, given name:

```
<creatorName nameType="Personal">Krefeld, Thomas</creatorName>
```

- Recommendation: Additionally, separate family name and given name in a specific field and generate the name field:

```
<givenName>Thomas</givenName>  
<familyName>Krefeld</familyName>
```

- Add a persistent identifier for persons (Gemeinsame Normdatei (GND) and Open Researcher and Contributor iD (ORCID)). This will make attributions robust to changes of names or affiliations:

```
<nameIdentifier  
  schemeURI="http://d-nb.info/gnd/"  
  nameIdentifierScheme="GND">  
  123778689  
</nameIdentifier>  
<nameIdentifier  
  schemeURI="http://orcid.org/"  
  nameIdentifierScheme="ORCID">  
  0000-0001-9657-6052  
</nameIdentifier>
```

- Recommendation for GND-entries: Only use PF/PS GND-entries. PF/PS entries are for a specific individual (e.g. *the* Thomas Müller); PN entries could include all people with a name (e.g. all persons called Thomas Müller).
- It is recommended to not use titles in names (as they can change).
- If a person has multiple affiliations:
  - It is recommended to state only one institution (the context of the resource determines the affiliation).
  - If unavoidable, multiple affiliations can be specified in the order of importance for the dataset published.
  - An affiliation is an institution, not a project (see [How should I specify an institution](#)).

## How should I handle different versions of the same research data?

Metadata can be updated without releasing a new version of the research data, but not vice versa; if the research data change, you need to update the metadata to reflect these changes.

If you want to publish several versions of the research data, but also want to have a point of reference for all of these publications together, we recommend to use a form of [DOI-versioning](#) [\[external link\]](#):

- Specify a set of metadata that is valid for all versions.
- Specify a set of metadata for each version.
- Update all these metadata with the according references (e.g. *isNewVersionOf*, see [RelatedIdentifier](#) for details).

## Best practice for specific fields

### Identifier

- This field can be omitted on submission: it is mandatory according to the DataCite standard, but it will be set by the data publisher.
- The assigned Digital Object Identifier (DOI) will be reported to you by the data publisher.

#### Example

```
<identifier identifierType="DOI">10.5282/ubm/data.158</identifier>
```

### Title

- This field is mandatory.
- Be as specific as you would be in the context of a journal publication.
- It is recommended to avoid filenames (e.g. “survey.csv”) or generic descriptions (e.g. “survey data”) as a title.
- The main title is specified without a *titleType*.
- Every title must be specified with its language (*xml:lang attribute*).
- If the main title is not specified in English, a title of type “TranslatedTitle” must be given in English.
- Title types “AlternativeTitle” and “Subtitle” are supported but not recommended. “Other” may not be used as a title type.

#### Example

```
<titles>
  <title
    xml:lang="de">
    Bayerisches Musiker-Lexikon Online (BML0)
  </title>
  <title
    xml:lang="en"
    titleType="TranslatedTitle">
    Digital Encyclopedia of Bavarian Musicians
  </title>
</titles>
```

### Creator

- This field is mandatory.

- Consult section on [how to specify a person](#).
- Always prefer natural persons over institutions. If no natural person can be identified, you must set the attribute *nameType* to “Organizational”.

### Example

```
<creators>
  <creator>
    <creatorName nameType="Personal">Krefeld, Thomas</creatorName>
    <givenName>Thomas</givenName>
    <familyName>Krefeld</familyName>
    <nameIdentifier
      schemeURI="http://d-nb.info/gnd/"
      nameIdentifierScheme="GND">
      123778689
    </nameIdentifier>
    <nameIdentifier
      schemeURI="http://orcid.org/"
      nameIdentifierScheme="ORCID">
      0000-0001-9657-6052
    </nameIdentifier>
    <affiliation
      xml:lang="de">
      Institut für Romanische Philologie, Ludwig-Maximilians-Universität München
    </affiliation>
  </creator>
</creators>
```

### Publisher

- This field can be omitted on submission: it is mandatory according to the DataCite standard, but it will be set by the data publisher, i.e. the institution that hosts the (meta)data.

### Example

```
<publisher
  xml:lang="de">
  Universitätsbibliothek der Ludwig-Maximilians-Universität München
</publisher>
```

### PublicationYear

- This field can be omitted on submission: it is mandatory according to the DataCite standard, but it will be set by the data publisher.

### Example

```
<publicationYear>2019</publicationYear>
```

## ResourceType

- This field is mandatory.
- Decision tree to pick the right *resourceTypeGeneral*:
  1. If the data submission contains heterogeneous data, consider publishing it in separate data publications or (less preferred) use “Collection”. If the data are homogeneous, proceed with 2.
  2. If the data are movies, images or sound files use “Audiovisual”, “Image” or “Sound”, respectively. If not, proceed with 3.
  3. If the data are a digital, interactive representations of some real-world phenomena (e.g. trained models in the context of machine learning) use “Model”. If not, proceed with 4.
  4. If the data are descriptions of a workflow (e.g. in the common workflow language), use “Workflow”. If not, proceed with 5.
  5. If the data are source code files (incl. configuration and built artefacts), use “Software”. If not, proceed with 6.
  6. If the data have a fixed structure (e.g. table-like), use “Dataset”. If not, proceed with 7.
  7. If the data are text files, use “Text”. If not, proceed with 8.
  8. Use “Other”.
- As a consequence, only use these values for *resourceTypeGeneral*:
  - Audiovisual
  - Collection
  - Dataset
  - Image
  - Model
  - Software
  - Sound
  - Text
  - Workflow
  - Other
- These *ResourceTypes* should **not** be used:
  - DataPaper
  - Event
  - InteractiveResource
  - PhysicalObject
  - Service
- Note: Only items with the *resourceTypeGeneral* “Dataset” will be included in the Google Dataset Search. All other types are currently not supported.

### Example

```
<resourceType resourceTypeGeneral="Dataset">
  Regional Climate Simulation
</resourceType>
```

## Subject

- This field is mandatory, in the DataCite standard it is only recommended.

## Mandatory subject annotations

- The following subject annotations are mandatory (must occur at least once):

Type of Subject	Standard	Type of standard	Usage hint
Discipline	DDC	Classification	Add 3 digits DDC as specific as possible followed by the English term for the discipline ( <a href="#">Canonical Source [external link]</a> )
Keywords	Wikidata Q-ID and GND	Keyword	Wikidata and GND terms are both mandatory, including redundancy (if an appropriate entry does not exist contact the ITG). Use <a href="#">Wikidata-Search [external link]</a> and <a href="#">GND-Search [external link]</a> to find the appropriate identifiers.

There should be no overlap between the discipline specifier(s) and the keywords.

## Geotagging

Specifying the location via subject is mandatory, if applicable to the resource:

- Canonical source for geonames is the [GeoNames Service \[external link\]](#) (a registration for API access is necessary).
- See [GeoLocation section](#) for a more detailed specification.

## Additional subject annotations

- Every other subject may be added.
- Specify the language of the subject.
- It is recommended to always qualify subjects by URL or scheme name. A good starting point to research existing schemes is [Bartog.org \[external link\]](#) - Basel Register of Thesauri, Ontologies & Classifications. Unqualified subjects (not controlled by a controlled vocabulary, ontology or any other standard for the subject terms) are often useless for research data aggregators due to ambiguities.

## Example

```
<subjects>
  <subject xml:lang="en" schemeURI="http://dewey.info/" subjectScheme="dewey">
    410 Linguistics
  </subject>
  <subject xml:lang="en" schemeURI="http://dewey.info/" subjectScheme="dewey">
    004 Data processing computer science
  </subject>
  <subject
    xml:lang="de"
```



```

    schemeURI="http://d-nb.info/gnd/"
    subjectScheme="GND"
    valueURI="http://d-nb.info/gnd/4740815-7">
        Chalet (Definition: Urspr. die franz. Bezeichnung für Alphütte)
    </subject>
    <subject
        xml:lang="de"
        schemeURI="https://www.wikidata.org/wiki/"
        subjectScheme="wikidata"
        valueURI="https://www.wikidata.org/wiki/Q136689">
            Sennhütte (Gebäude, einfach, auf der Alm zur Verarbeitung von Milch)
    </subject>
    <subject
        xml:lang="en"
        schemeURI="https://glottolog.org/resource/languoid/id/"
        subjectScheme="Glottocode"
        valueURI="https://glottolog.org/resource/languoid/id/high1286">
            High German
    </subject>
    <subject
        xml:lang="de"
        schemeURI="http://www.geonames.org/"
        subjectScheme="geonames"
        valueURI="http://www.geonames.org/2764958">
            Hall in Tirol
    </subject>
</subjects>

```

## Contributor

- This field is recommended if the data are published with a free license.
- If the license specified via the **Rights** field restricts the usage in a way, that possibly necessitates interaction with the rightsholder, contributor of type “RightsHolder” must be specified. Examples of free licenses are CC-0, CC-BY, or CC-SA; non-free licenses are for example CC-NC or CC-ND.
- Consult the section on **how to specify a person**.
- If contributors change over versions, the version metadata should only include the actual contributors of the updated version. A metadata set representing all versions of the dataset (including links to the versions) can include all contributors with the dates of participation, see **how to handle different versions of the research data**.
- Duplicate mentions between Creator and Contributor are unproblematic.
- If a person has multiple roles, it is recommended to identify the most important role of that person and select only one.
- Be as specific as possible (a “ProjectLeader” is also considered to be a “ProjectMember”, but “ProjectLeader” carries more information). Use generic role descriptions only when nothing else fits.
- The following roles are recommended:

Option	Description from DataCite standard ( <i>italics</i> ) and usage hints
ContactPerson	<i>Person with knowledge of how to access, troubleshoot, or otherwise field issues related to the resource.</i>
DataCollector	<i>Person/institution responsible for finding, gathering/collecting data under the guidelines of the author(s) or Principal Investigator (PI).</i>
DataCurator	<i>Person tasked with reviewing, enhancing, cleaning, or standardizing metadata and the associated data submitted for storage, use, and maintenance within a data centre or repository.</i>
HostingInstitution	<i>Typically, the organisation allowing the resource to be available on the internet through the provision of its hardware/software/operating support.</i>
ProjectLeader	<i>Person officially designated as head of project team or sub-project team instrumental in the work necessary to development of the resource.</i>
ProjectManager	<i>Person officially designated as manager of a project. Project may consist of one or many project teams and sub-teams.</i>
ProjectMember	<i>Person on the membership list of a designated project/project team. All persons with a contract in the context of the project which produced the resource.</i>
Researcher	<i>A person involved in analyzing data or the results of an experiment or formal study. May indicate an intern or assistant to one of the authors who helped with research but who was not so “key” as to be listed as an author.</i>
RightsHolder	<i>Person or institution owning or managing property rights, including intellectual property rights over the resource. Mandatory for non-free licenses; person or institution that owns the rights listed in field <b>Rights</b>.</i>
WorkPackageLeader	<i>The Work Package Leader is responsible for ensuring the comprehensive contents, versioning, and availability of the Work Package during the development of the resource.</i>

## Example

```

<contributors>
  <contributor contributorType="ProjectLeader">
    <contributorName>Ludwig, Ralf</contributorName>
    <givenName>Ralf</givenName>
    <familyName>Ludwig</familyName>
    <nameIdentifier
      nameIdentifierScheme="ORCID"
      schemeURI="http://orcid.org/">
      0000-0002-4225-4098
    </nameIdentifier>
    <affiliation xml:lang="de">
      Department für Geographie, Ludwig-Maximilians-Universität München
    </affiliation>
  </contributor>
  <contributor contributorType="RightsHolder">
    <contributorName>Štědranská, Markéta</contributorName>
    <givenName>Markéta</givenName>
    <familyName>Štědranská</familyName>

```

```

    <nameIdentifier
      nameIdentifierScheme="GND"
      schemeURI="http://d-nb.info/gnd/">
        141321350
    </nameIdentifier>
    <affiliation xml:lang="de">
      Institut für Musikwissenschaft, Universität Wien
    </affiliation>
  </contributor>
</contributors>

```

## Date

- This field is recommended.
- Time periods can be specified by specifying the start date and the end date separated by a slash.
- The following types should be filled-out by the data producer:
  - Collected: Time range when the resource was arranged (not necessarily identical to the time range when the resource was created).
  - Created: First version of a resource, cannot be identical with updated.
  - Updated: If the resource is a newer version, cannot be identical with created.
- The following types are set by the publisher:
  - Submitted: point in time when the data were recieved by the data publisher.
  - Accepted: point in time when the data publisher decides to publish the data.
  - Issued: long format of the field **PublicationYear**, point in time when a publisher publishes the data; should be set.
  - Available: only use in the context of embargo periods (this is not recommended).
  - Withdrawn: point in time when the publisher retracts the data publication.
- For dates describing the period the resource covers use “Other” and add “Coverage” as a description.
- It is recommended to use the attribute *dateInformation* for disambiguation, if multiple dates with the same type are specified.
- “CopyRighted” as a *dateType* will not be used.

### Example

```

<dates>
  <date dateType="Created">2016</date>
  <date
    dateType="Other"
    dateInformation="coverage">
    2050-09-01T00:00:00/2050-09-30T23:59:59
  </date>
</dates>

```

## RelatedIdentifier

- This field is recommended.

- The relations are always specified from the perspective of the described resource (if the metadata describe a publication A and this publication cites another publication B, the relation is A cites B and not B IsCitedBy A).
- Only use these *relationTypes*:

relationType	Definition (Derived from DataCite Metadata Schema v4.3)	Example	Usage
Describes	indicates A describes B	a README “describes” a dataset	
IsDescribedBy	indicates A is described by B	a dataset “IsDescribedBy” a README	
HasPart	indicates A is a portion of B; may be used for elements of a series	a container “HasPart” a dataset	use to create hierarchical structures
IsPartOf	indicates A includes the part B	a dataset “IsPartOf” a container	use to create hierarchical structures
HasMetadata	indicates resource A has additional metadata B	a dataset “HasMetadata”	use to add additional metadata to a dataset e.g. domain-specific metadata
IsMetadataFor	indicates additional metadata A for a resource B	a metadata file “IsMetadataFor” a dataset	use to add additional metadata to a dataset e.g. domain-specific metadata
HasVersion	indicates A has a version B	a dataset “HasVersion”	use for relating a non-versioned or abstract resource to a versioned resource e.g. a snapshot; alternative usage: Use if exact order of the versions is not known

relationType	Definition (Derived from DataCite Metadata Schema v4.3)	Example	Usage
IsVersionOf	indicates B is a version of A	a dataset “IsVersionOf” of a dataset	use for relating a versioned resource (e.g. a snapshot) to a non-versioned or abstract resource; alternative Usage: Use if exact order of the versions is not known
IsNewVersionOf	indicates A is a new edition of B, where the new edition has been modified or updated	a dataset “IsNewVersionOf” a dataset	use to order versions
IsPreviousVersionOf	indicates B is a previous edition of A	a dataset “IsPreviousVersionOf” a dataset	use to order versions
IsSourceOf	indicates A is a source upon which B is based	a dataset A “IsSourceOf” a dataset B	use to express that A has been refined or modified by B
IsDerivedFrom	indicates B is a source upon which A is based	a dataset B “IsDerivedFrom” a dataset A	use to express that B is the result of refining or modifying the content of A
References	indicates B is used as a source of information for A	an article “References” a dataset	
IsReferencedBy	indicates A is used as a source of information by B	a dataset “IsReferencedBy” an article	
IsVariantFormOf	indicates A is a variant or different form of B	a dataset (xml-file) “IsVariantFormOf” a dataset (csv-file) with the same content	
IsIdenticalTo	indicates that A is identical to B, for use when there is a need to register two separate instances of the same resource	a dataset in location A “IsIdenticalTo” a dataset in location B	use for a resource that is the same as the registered resource but is saved in another location, maybe another institution
IsSupplementTo	indicates that A is a supplement to B	a dataset “IsSupplementTo” an article	

relationType	Definition (Derived from DataCite Metadata Schema v4.3)	Example	Usage
IsSupplementedBy	indicates that B is a supplement to A	an article “IsSupplementedBy” a dataset	
Documents	indicates A is documentation about B; e.g. points to software documentation	a data management plan (DMP) “Documents” a dataset	
IsDocumentedBy	indicates B is documentation about/explaining A; e.g. points to software documentation	a dataset “IsDocumentedBy” a data management plan (DMP)	

- Make sure, that links from the described resource are mirrored in in the referenced resource. For example, make sure that a paper referencing the data includes a formal link to the data set (and vice versa). This also applies to all reciprocal *relationTypes* (e.g. “IsNewVersionOf” and “IsPreviousVersionOf”).
- The *relationIdentifierType* offers a list containing the most important identifier types to specify the relation (e.g. DOI, ISSN, PURL, URN).
- If *relationType* “HasMetadata” is chosen, please specify additionally the *relatedMetadataSchema*, *schemeURI* and *schemeType* (e.g. XSD, DDT) of the related identifier.
- The publisher may curate the list of *relatedIdentifiers* (e.g. adding relevant related resources over time on a best effort basis).

### Example

```

<relatedIdentifiers>
  <relatedIdentifier relatedIdentifierType="DOI" relationType="IsDescribedBy">
    10.1175/JAMC-D-18-0021.1
  </relatedIdentifier>
  <relatedIdentifier
    relatedIdentifierType="DOI"
    relationType="HasMetadata"
    relatedMetadataSchema="DDI-L"
    schemeURI="http://www.ddialliance.org/[...]/3.1/XMLSchema/instance.xsd">
    10.1234/567890
  </relatedIdentifier>
</relatedIdentifiers>

```

## Description

- This field is mandatory, whereas the DataCite standard only recommends it: There has to be at least one entry of type “Abstract” in English.
- Always specify the used language (*xml:lang attribute*) of each description.
- If there are descriptions in more than one language, the content may be different (no literal translation required).
- Each description has a limit of 300 words.
- Description of *descriptionType* “Methods” is optional. Best practice: use keywords from [this controlled list](#), separated by comma. This list will either be extended in the future or replaced by a controlled vocabulary (in this case mappings will be provided).
- Description of *descriptionType* “TechnicalInfo” is optional. Best practice: use keywords from [this controlled list](#), separated by comma. This list will either be extended in the future or replaced by a controlled vocabulary (in this case mappings will be provided). Additionally, data producers could consider creating a README and link it via the [RelatedIdentifier](#) field.
- These types are not recommended:
  - SeriesInformation
  - TableOfContents
  - Other

### Example

```
<descriptions>
  <description xml:lang="en" descriptionType="Abstract">
    The "Kritische Ausgabe der Werke von Richard Strauss", a long-term editorial
    project, has been under way at the Institut für Musikwissenschaft of the
    Ludwig-Maximilians-Universität Munich since 2011; it is directed by ...
  </description>
  <description xml:lang="de" descriptionType="Abstract">
    Das Langzeit-Editionsprojekt „Kritische Ausgabe der Werke von Richard Strauss"
    wird seit Februar 2011 unter der Leitung von Prof. Dr. Hartmut Schick am
    Institut für Musikwissenschaft der Ludwig-Maximilians-Universität München ...
  </description>
  <description xml:lang="en" descriptionType="Methods">
    digital editing, software/application development
  </description>
</descriptions>
```

## GeoLocation

- This field is recommended.
- Describes the resource (e.g. where an image has been taken or where a sensor is located), *not* the related project or institute, if the former is not applicable, do *not* use it for the latter.
- *geoLocationPlace* must be identical to payload of the according GeoNames field in the subjects, consult the [geotagging](#) subsection.
- Canonical source for coordinates is the [GeoNames Service](#) [\[external link\]](#).

### Examples

- geoLocationPlace and geoLocationPolygon:

```

<geoLocations>
<geoLocation>
  <geoLocationPlace>Höslwang</geoLocationPlace>
  <geoLocationPolygon>
    <polygonPoint>
      <pointLatitude>47.9231796264648</pointLatitude>
      <pointLongitude>12.2860469818115</pointLongitude>
    </polygonPoint>
    <polygonPoint>
      <pointLatitude>47.9231796264648</pointLatitude>
      <pointLongitude>12.3512439727784</pointLongitude>
    </polygonPoint>
    <polygonPoint>
      <pointLatitude>47.9707412719727</pointLatitude>
      <pointLongitude>12.3512439727784</pointLongitude>
    </polygonPoint>
    <polygonPoint>
      <pointLatitude>47.9707412719727</pointLatitude>
      <pointLongitude>12.2860469818115</pointLongitude>
    </polygonPoint>
    <polygonPoint>
      <pointLatitude>47.9231796264648</pointLatitude>
      <pointLongitude>12.2860469818115</pointLongitude>
    </polygonPoint>
  </geoLocationPolygon>
</geoLocation>
</geoLocations>

```

- geoLocationPlace and geoLocationBox:

```

<geoLocations>
<geoLocation>
  <geoLocationPlace>Hall in Tirol</geoLocationPlace>
  <geoLocationPoint>
    <pointLongitude>11.51667</pointLongitude>
    <pointLatitude>47.28333</pointLatitude>
  </geoLocationPoint>
  <geoLocationBox>
    <westBoundLongitude>11.5272636413574</westBoundLongitude>
    <eastBoundLongitude>11.4707803726196</eastBoundLongitude>
    <southBoundLatitude>47.2697830200196</southBoundLatitude>
    <northBoundLatitude>47.2893867492676</northBoundLatitude>
  </geoLocationBox>
</geoLocation>
</geoLocations>

```



## Language

- This field is optional.
- Main language of the resource, not the metadata.
- Recommendation: use ISO 639-1 (listed on [Wikipedia \[external link\]](#)).

### Example

```
<language>en</language>
```

## AlternateIdentifier

- This field is optional.
- These alternate identifiers identify the resource, meaning that it is also known under the following identifiers.
- Possible values for *alternateIdentifierType*:
  - project-specific identifier: an ID that has meaning inside the project that created the data.
  - application-specific identifier: an ID that has meaning in the context of an application that is used to process the data.
  - institution-specific identifier: an ID that has meaning in the context of the institution that provides, funded or created the data.

### Example

```
<alternateIdentifiers>  
  <alternateIdentifier alternateIdentifierType="project-specific identifier">  
    S15645_v4_182  
  </alternateIdentifier>  
</alternateIdentifiers>
```

## Size

- This field is optional.
- Specify the size in Bytes, with a trailing B.
- If the data are compressed, specify the size of the compressed file/archive.
- If the data consist of several units (without using an archival software), specify their combined sizes.
- Do not use semantical units (records, pages, rows, etc.), put them in the description if they are relevant.

### Example

```
<sizes>  
  <size>6197732B</size>  
</sizes>
```

## Format

- This field is optional.

- Use MIME type format as specified in [RFC 2646 \[external link\]](#), possible values should be taken from [the IANA list of MediaTypes \[external link\]](#).
- Specify in this order (skip if it does not imply):
  1. If files are compressed, specify the mime type of the compression first as a suffix to the mime type (e.g. text/xml+zip).
  2. If files are in an archive, specify the mime type of the archive format.
  3. Specify each mime-type in its own field in alphabetical order, do not repeat mime-types.

#### Example

```
<formats>
  <format>application/tar+gzip</format>
  <format>application/netcdf</format>
  <format>text/plain</format>
  <format>text/csv</format>
</formats>
```

#### Version

- This field is optional.
- The versioning information is set according to the policies of the data provider (data publishers do not change/use this field).
- It is recommended to use [semantic versioning \[external link\]](#).

#### Example

```
<version>4.2.1</version>
```

#### Rights

- This field is mandatory, whereas the DataCite standard specifies it as optional.
- *RightsURI* must be set.
- Use only ONE license for data and code each (to avoid inconsistencies).
- Use [SPDX \[external link\]](#) to unambiguously specify the license used.

Guidance for using a license:

- Recommendation: [Creative Commons \(CC\) \[external link\]](#) as license for data and [Apache 2.0 license \[external link\]](#) for software.
- You should not use CC licenses with the NC or ND extension (although submissions with these extensions are accepted).
- It is recommended to use only one type of licence, not several at the same time.
- In case both research data and software code need to be licensed in one publication (which is not recommended, consider two data publications, see [ResourceType](#)), specify the licenses in the order data, software.
- For further license options consult [Choose a license \[external link\]](#) or the [CC license helper \[external link\]](#).

#### Example

```

<rightsList>
  <rights
    xml:lang="en-US"
    schemeURI="https://spdx.org/licenses/"
    rightsIdentifierScheme="SPDX"
    rightsIdentifier="CC0-1.0"
    rightsURI="http://creativecommons.org/publicdomain/zero/1.0/"
  />
</rightsList>

```

## FundingReference

- This field is optional.
- This is the place to add information about the project and its funding.
- *funderName*: See [How should I Specify an Institution](#).
- Use [Cordis \(EU\)](#) [external link], [Gepris \(DFG\)](#) [external link], [FWF \(Austria\)](#) [external link] et al. for identification of fundings.
- *awardTitle* is the name of the grant, not the funding line (Förderlinie).

### Example

```

<fundingReferences>
  <fundingReference>
    <funderName>Deutsche Forschungsgemeinschaft (DFG)</funderName>
    <funderIdentifier funderIdentifierType="ROR">
      https://ror.org/018mejw64
    </funderIdentifier>
    <awardNumber>253900505</awardNumber>
    <awardURI>http://gepris.dfg.de/gepris/projekt/253900505</awardURI>
    <awardTitle xml:lang="de">
      VERBA ALPINA. Der alpine Kulturraum im Spiegel seiner Mehrsprachigkeit
    </awardTitle>
  </fundingReference>
</fundingReferences>

```

## Examples

### Bayerisches Musiker-Lexikon Online (BMLO)

The Digital Encyclopedia of Bavarian Musicians (Bayerisches Musiker-Lexikon Online, short BMLO) refers as a musicological model project. The BMLO offers a digital biographical dictionary focussing on music science, furthermore it enriches the presented personalities of Bavarian history of music by implementing further information, gathered from biographical literature, archives, libraries and digital collections. In this way, the BMLO constitutes the core of an interconnected, virtual cluster for history of music. Currently, 24621 out of a total of 27818 records is presented on the web. As parts of this semantic network should also be mentioned the Munich Dictionary of Musik (Münchner Musiklexikon, short MUK), which serves since 2010 as a encyclopedia for

music corporations with a linkage to Munich, as well as LOCI, a geographic database for music, culture and history, founded in 2012.

[BMLO metadata example](#)

## VerbaAlpina

The project VerbaAlpina seeks to investigate the Alpine region in its cultural and historical linguistic unity from a transnational perspective through selective analyses. For this purpose, a comprehensive, multilingual research environment will be set up using up-to-date media technology (database, georeferenced cartography, internet, data collection with social software). The guiding intention is to analyze already existing data from linguistic atlases and dictionaries (partly difficult of access) which is first structured and linguistically annotated according to certain principles to fit the database of VerbaAlpina.

[Verba Alpina metadata example 1](#)

[Verba Alpina metadata example 2](#)

## ClimEx

The ClimEx project investigates the effects of climate change on meteorological and hydrological extreme events and implications for water management in Bavaria and Québec.

It especially consists of two new aspects:

- An ensemble of 50 transient runs of the canadian general circulation model CanESM2 (~200km resolution) from 1950 to 2100, resulting in 7500 years of modelled climate. As each of these runs is initialized with only slightly altered starting conditions, this ensemble can be interpreted as (modelled) natural variability. The CanESM2 then drives the regional climate model CRCM5 (~11km resolution) for a domain that covers most of central Europe. Both models are internationally established and widely used in the climate science community.
- A physically based hydrological model (WaSiM) is driven by this climate input for the entire hydrological Bavaria at very high temporal and spatial resolution of 3 hours and 500m to investigate both climate change impacts and natural variability of extreme events, especially floods.

ClimEx further strengthens the international collaboration between Bavaria and Québec as research facilities, universities and public water agencies intensify their former cooperation approaches.

[ClimEx metadata example \[external link\]](#)