

## Vocabulary for the Registration and Description of Research Data Repositories

---

Version 2.0

December 2012

doi:10.2312/re3.002

Authors: Paul Vierkant <sup>b</sup>, Shaked Spier <sup>b</sup>, Jessika Rücknagel <sup>b</sup>, Jens Gundlach <sup>c</sup>, David Fichtmüller <sup>d</sup>, Heinz Pampel <sup>a</sup>, Maxi Kindling <sup>b</sup>, Agnes Kirchhoff <sup>d</sup>, Hans-Jürgen Göbelbecker <sup>c</sup>, Jens Klump <sup>a</sup>, Roland Bertelmann <sup>a</sup>, Peter Schirmbacher <sup>b</sup>, Frank Scholze <sup>c</sup>

<sup>a</sup> GFZ German Research Centre for Geosciences, Library and Information Services (LIS)

<sup>b</sup> Humboldt-Universität zu Berlin, Berlin School of Library and Information Science (BLIS)

<sup>c</sup> Karlsruhe Institute of Technology (KIT), KIT Library

<sup>d</sup> Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin

### Contact

info@re3data.org

<http://www.re3data.org>

## Table of Contents

1	Introduction .....	3
1.1	re3data.org .....	3
1.2	The Vocabulary .....	4
1.3	Research Data Repository Registration Procedure .....	4
1.4	Version 2.0 Update .....	5
2	re3data.org Metadata Properties.....	5
3	XML Example .....	13
4	Appendix.....	16
4.1	Attribute Values and Controlled Vocabularies .....	16
4.2	Access To Research Data Repositories Explained.....	23

## 1 Introduction

### 1.1 re3data.org

Research data are valuable and ubiquitous. The permanent access to research data is a challenge for all stakeholders in the scientific community. The long-term preservation and the principle of open access to research data offer broad opportunities for the scientific community. More and more universities and research centres are starting to build research data repositories allowing permanent access to data sets in a trustworthy environment. Due to disciplinary requirements, the landscape of data repositories is very heterogeneous. Thus it is difficult for researchers, funding bodies, publishers and scholarly institutions to select appropriate repositories for storage and search of research data.

The goal of re3data.org is to create a global registry of research data repositories. The registry will cover research data repositories from different academic disciplines. re3data.org will present repositories for the permanent storage and access of data sets to researchers, funding bodies, publishers and scholarly institutions. In the course of this mission re3data.org aims to promote a culture of sharing, increased access and better visibility of research data. In the first phase of the project the following tasks are prioritized:

- the conception and construction of a web-based registry of research data repositories;
- the formulation of a vocabulary to describe research data repositories;
- the definition of selection criteria of research data repositories.

According to the schedule the registry will go live in autumn 2012.

re3data.org is funded by the German Research Foundation (DFG)<sup>1</sup>.

re3data.org is a joint project of the Berlin School of Library and Information Science (BLIS) at the Humboldt-Universität zu Berlin<sup>2</sup>, the GFZ German Research Centre for Geosciences<sup>3</sup> and the Karlsruhe Institute of Technology (KIT)<sup>4</sup>. The partners are actively involved in the German Initiative for Network Information (DINI)<sup>5</sup> and current research data management activities.

---

<sup>1</sup> <http://www.dfg.de>

<sup>2</sup> <http://www.ibi.hu-berlin.de>

<sup>3</sup> <http://bib.gfz-potsdam.de>

<sup>4</sup> <http://www.bibliothek.kit.edu>

<sup>5</sup> <http://www.dini.de/english>

## 1.2 The Vocabulary

The re3data.org vocabulary is a list of metadata properties describing a research data repository regarding its general scope, content and infrastructure as well as its compliance with technical, metadata and quality standards. The vocabulary includes required metadata properties and optional properties providing additional information. The vocabulary is designed to

- recommend a standard for describing a research data repository;
- provide the basis for interoperability between research data repositories and re3data.org;
- be one of the first steps towards the goal of a certificate for research data repositories.

The first version of the vocabulary was developed out of 20 randomly chosen repositories of a survey of 400 research data repositories. These findings as well as suggestions from project partners have led to version 1.0. After its publication in July 2012 we received feedback leading to substantial changes in the structure of the vocabulary. Version 2.0 of the vocabulary synthesizes all the responses as well as current developments in the realm of research data repositories. Future developments of the vocabulary will still depend on the feedback of the research data repository community. This open and transparent development process is to ensure a strong basis for a future standard for describing research data repositories supported by and rooted in the community. We appreciate your feedback and therefore would like to invite you to tell us what you think about our work done so far.

## 1.3 Research Data Repository Registration Procedure

Institutions can suggest research data repositories to be listed in the registry either via email ([info@re3data.org](mailto:info@re3data.org)) or a simple URL suggest form<sup>6</sup>. Soon after the launch research data repositories can be registered via an extended online form. The suggested repositories will be listed in re3data.org and marked as 'not reviewed'. The repository is checked by the re3data.org team according to the re3data.org minimum requirements (that the access to the data and repository as well as the terms of use must be clear), and will be marked as reviewed. This workflow ensures that all known research data repositories are indexed while distinguishing between reviewed and not reviewed repositories.

---

<sup>6</sup> <http://www.re3data.org/suggest/>

## 1.4 Version 2.0 Update

Version	Date of Publication	Changes to prior Version
1.0	July 2012	Initial Document
2.0	December 2012	structural changes, including addition of language attributes, addition of definitions of controlled vocabularies, changes in occurrences

## 2 re3data.org Metadata Properties

The table below provides a detailed description of the re3data.org properties that describe a research data repository registered in re3data.org. For an example of how the properties are expressed in XML format, please see the XML example provided (3).

A naming convention has been used for all properties and attributes as follows: properties, 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 *Occurrence (Occ)* 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). "Req" indicates that an attribute is required if the corresponding property is applied. Where the allowed values specify a "Controlled Vocabulary," the members of the list are regulated by the re3data.org team. The values of the controlled vocabularies are explained and defined in the appendix (4.1).

ID	re3data.org property	Definition	A/C	Occ	Allowed values, examples, other constraints
1	identifier	The ID is a unique string, which allows identifying the research data repository. The ID is assigned by re3data.org.		1	Auto-Value Example: r3d100000001
2	repositoryName	The full name of the research data repository.		1	The format is open. Example: Access to Archival Databases
2.1	language	Language of the research data repository name.	A	Req	Controlled vocabulary Allowed values from: ISO-639-3. Examples: eng, deu, fra
3	additionalName	The alternative name or acronym for the research data repository.		0-n	The format is open. Example: AAD

ID	re3data.org property	Definition	A/C	Occ	Allowed values, examples, other constraints
3.1	language	Language of the research data repository additional name.	A	Req	Controlled vocabulary Allowed values from: ISO-639-3. Examples: eng, deu, fra
4	repositoryURL	The URL, which gives reference to the research data repository.		1	URL Example: <a href="http://cdiac.esd.ornl.gov/home.html">http://cdiac.esd.ornl.gov/home.html</a>
5	repositoryIdentifier	A globally unique identifier that refers to the research data repository (e.g. DOI, URN, VIAF, GND).		0-n	Example: doi:10.1234/foo
6	description	A textual description containing additional information about the data repository (primary language is English).		0-1	The format is open (max. 1000 characters).
6.1	language	Language of the research data repository description.	A	Req	Controlled vocabulary Allowed values from: ISO-639-3. Examples: eng, deu, fra
7	type	The type of the research data repository.		1-n	Controlled vocabulary Allowed values: disciplinary institutional other
8	size	The number of items contained in the research data repository.		0-1	The format is open. Example: 50 datasets; 30 studies
8.1	updated	The date of the last update of the research data repository size.	A	Req	YYYY or YYYY-MM-DD or any other format described in W3CDTF (ISO-8601)
9	startDate	Releasing date of the research data repository.		0-1	YYYY or YYYY-MM-DD or any other format described in W3CDTF (ISO-8601)
10	endDate	Date when the research data repository stopped the ingest of new research data (still making the research data available).		0-1	YYYY or YYYY-MM-DD or any other format described in W3CDTF (ISO-8601)
11	repositoryLanguage	The user interface language of the research data repository.		1-n	Controlled vocabulary Allowed values from: ISO-639-3. Examples: eng, deu, fra

ID	re3data.org property	Definition	A/C	Occ	Allowed values, examples, other constraints
12	subject	The subject(s) of the research data repository.		1-n	The format is dependent on scheme. Examples: 11 Humanities 10101 Prehistory
12.1	subjectScheme	The subject scheme according to which the subject (see ID 12 subject) of the research data repository described.	A	Req	Controlled vocabulary Allowed values from: DFG
13	missionStatementURL	The URL of a mission statement describing the designated community of the research data repository.		0-1	URL Example: <a href="http://cdiac.ornl.gov/aboutcdiac.html">http://cdiac.ornl.gov/aboutcdiac.html</a>
14	contentType	All types of resources available in the research data repository.		0-n	The format is dependent on scheme. Examples: Raw data, Images, ...
14.1	contentTypeScheme	The name and/or URL of the subject scheme or classification code (e.g. parse)	A	Req	Controlled vocabulary Allowed values from: parse
15	providerType	The type of provider.		1-2	Controlled vocabulary Allowed values: dataProvider serviceProvider
16	keyword	English keyword(s) describing the subject focus of the research data repository.		0-n	Examples: computer science
17	institution	All institutions being responsible for funding, creating and/or running the research data repository (wrapper element).		1-n	-
17.1	institutionName	Name of the responsible institution.	C	1	Example: National Institute of General Medical Sciences, U.S. National Institutes of Health
17.1.1	language	Language of the institution name.	A	Req	Controlled vocabulary Allowed values from: ISO-639-3. Examples: eng, deu, fra
17.2	institutionAdditionalName	The alternative name or acronym for the responsible institution.	C	0-n	Example: NIGMS

ID	re3data.org property	Definition	A/C	Occ	Allowed values, examples, other constraints
17.2.1	language	Language of the institution additional name.	A	Req	Controlled vocabulary Allowed values from: ISO-639-3. Examples: eng, deu, fra
17.3	institutionCountry	Location of the responsible institution.	C	1	Controlled vocabulary Allowed Values from: ISO-3166-1 alpha-3. Examples: DEU, GBR, USA
17.4	responsibilityType	Type of responsibility for each responsible institution.	C	0-n	Controlled vocabulary Allowed values: funding general sponsoring technical
17.5	institutionType	Type of responsible institution.	C	0-1	Controlled vocabulary Allowed values: commercial non-profit
17.6	institutionURL	URL of the responsible institution.	C	0-1	Example: <a href="http://www.nigms.nih.gov/">http://www.nigms.nih.gov/</a>
17.7	institutionIdentifier	A globally unique identifier that refers to the research data repository (e.g. ORCID, VIAF, GND).	C	0-n	Example: doi:10.1234/foo
17.8	responsibilityStartDate	Start date of period of responsibility.	C	0-1	YYYY or YYYY-MM-DD or any other format described in W3CDTF (ISO-8601)
17.9	responsibilityEndDate	End date of period of responsibility.	C	0-1	YYYY or YYYY-MM-DD or any other format described in W3CDTF (ISO-8601)
17.10	contact	Email address of the contact or an online contact form URL.	C	0-n	The format is open. Example: john.doe@researchdata.com
18	policy	Any kind of policy (e.g. data policy, etc.) provided by the research data repository to clarify legal aspects (wrapper element).		0-n	-
18.1	policyName	Name of the policy.	C	1	The format is open. Example: Data policy of the information system PANGAEA
18.2	policyURL	URL of the policy.	C	1	Example: <a href="http://www.pangaea.de/curator/files/pangaea-data-policy.pdf">http://www.pangaea.de/curator/files/pangaea-data-policy.pdf</a>
19	databaseAccess	(wrapper element)		1	-



ID	re3data.org property	Definition	A/C	Occ	Allowed values, examples, other constraints
19.1	databaseAccessType	Type of access to the research data repository.	C	1	Controlled vocabulary Allowed values: open restricted closed
19.2	databaseAccessRestriction	All existing access restrictions to the research data repository. (Required if restricted is chosen).	C	0-n	Controlled vocabulary Allowed values: feeRequired registration other
20	databaseLicense	Database license of the research data repository (wrapper element).		0-n	-
20.1	databaseLicenseName	Name of database license.	C	1	Controlled vocabulary Allowed values: CC Copyrights ODC other
20.2	databaseLicenseURL	Database license URL.	C	1	Example: <a href="http://creativecommons.org/licenses/">http://creativecommons.org/licenses/</a>
21	dataAccess	(wrapper element)		1-n	-
21.1	dataAccessType	Type of access to data.	C	1	Controlled vocabulary Allowed values: open restricted closed
21.2	dataAccessRestriction	All existing access restrictions to the research data. (Required if restricted is chosen).	C	0-n	Controlled vocabulary Allowed values: feeRequired registration other
22	dataLicense	License of the data, existing in the research data repository. (wrapper element)		1-n	-
22.1	dataLicenseName	Name of the data license.	C	1	Controlled vocabulary Allowed values: CC Copyrights ODC OGL RL other none
22.2	dataLicenseURL	Data license URL.	C	1	Example: <a href="http://creativecommons.org/licenses/">http://creativecommons.org/licenses/</a>
23	dataUpload	(wrapper element)		1	-

ID	re3data.org property	Definition	A/C	Occ	Allowed values, examples, other constraints
23.1	dataUploadType	Type of data upload.	C	1	Controlled vocabulary Allowed values: open restricted closed
23.2	dataUploadRestriction	All existing restrictions to the data upload. (Required if restricted is chosen).	C	0-n	Controlled vocabulary Allowed values: feeRequired registration other
24	dataUploadLicense	The license for data upload (wrapper element).		0-n	-
24.1	dataUploadLicenseName	Name of the licence for data upload.	C	1	Example: RAD Submission of Array Data
24.2	dataUploadLicenseURL	Data upload license URL.	C	1	URL Example: <a href="http://cbil.upenn.edu/RAD/php/RAD-data_sub.php?">http://cbil.upenn.edu/RAD/php/RAD-data_sub.php?</a>
25	software	The software that is used to run the research data repository (wrapper element).		0-n	-
25.1	softwareName	Name of the research data repository software.	C	1	Controlled vocabulary Allowed values: CKAN DataVerse DigitalCommons dLibra DSpace EPrints eSciDoc Opus other unknown
26	versioning	The research data repository supports versioning of research data.		0-1	Controlled vocabulary Allowed values: yes / no
27	api	API URL		0-n	Example: <a href="http://www.datadryad.org/oa/">http://www.datadryad.org/oa/</a>

ID	re3data.org property	Definition	A/C	Occ	Allowed values, examples, other constraints
27.1	apiType	Type of API.	A	Req	Controlled vocabulary Allowed values: API FTP OAI-PMH REST SOAP SPARQL SWORD other
28	pidSystem	The persistent identifier system that is used by the research data repository.		1-n	Controlled vocabulary Allowed values: ARK DOI HDL PURL URN other none
29	citationGuidelineURL	The URL of the research data repository providing information on how to cite its research data.		0-1	Example: <a href="http://wiki.pangaea.de/wiki/Citation">http://wiki.pangaea.de/wiki/Citation</a> ; <a href="http://archaeologydataservice.ac.uk/archsearch/record.jsf?titleId=1686608">http://archaeologydataservice.ac.uk/archsearch/record.jsf?titleId=1686608</a>
30	qualityManagement	Any form of quality management concerning the data or metadata of the research data repository.		0-1	Controlled vocabulary Allowed values: yes / no / unknown
31	certificate	The certificate, seal or standard the research data repository complies with.		0-n	Controlled vocabulary: DIN 31644 DINI Certificate DSA ISO 16363 ISO 16919 TRAC WDS other
32	syndication	URL of the alerting service(s) provided by the research data repository.		0-n	Example: <a href="https://www.datadryad.org/feed/atom_1.0/10255/3">https://www.datadryad.org/feed/atom_1.0/10255/3</a>
32.1	syndicationType	Type of alerting service.	A	Req	Controlled Vocabulary: ATOM RSS
33	remarksIntern	Remarks for re3data.org internal use.		0-1	The format is open.
34	remarksExtern	Additional remarks that are visible to all users.		0-1	The format is open.

<i>ID</i>	<i>re3data.org property</i>	<i>Definition</i>	<i>A/C</i>	<i>Occ</i>	<i>Allowed values, examples, other constraints</i>
35	entryDate	The date the research data repository was indexed in re3data.org.		1	YYYY-MM-DD (ISO-8601)
36	lastUpdate	The date the metadata of the research data repository was updated.		1	YYYY-MM-DD (ISO-8601)
37	lastEditorIntern	The name of the last editor who edited the metadata of the research data repository.		1	Given Name Surname

### 3 XML Example

This is an XML example (of a fictional research data repository) illustrating how the re3data.org vocabulary can be used to describe research data repositories.

```
<?xml version="1.0" encoding="utf-8"?>
<!--re3data.org Vocabulary for the Registration and Description of Research Data Repositories. Version 2.0, December 2012. doi:10.2312/re3.002-->
<r3d:re3data xmlns:r3d=" http://www.re3data.org/schema/2-0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="
http://schema.re3data.org/2-0/re3dataV2.xsd">
<r3d:re3data xmlns:r3d="http://www.re3data.org">
<r3d:repository>
  <r3d:identifier>r3d000001001</r3d:identifier>
  <r3d:repositoryName language="eng">Global Research Data Repository</r3d:repositoryName>
  <r3d:additionalName language="eng">GReDaR</r3d:additionalName> <!-- optional --> <!-- multiple -->
  <r3d:additionalName language="deu">Globales Forschungsdatenrepositorium</r3d:additionalName> <!-- optional --> <!-- multiple -->
  <r3d:repositoryURL>http://www.globalresearchdatarepository.org</r3d:repositoryURL>
  <r3d:repositoryIdentifier>urn:0123456789/gredar</r3d:repositoryIdentifier> <!-- optional --> <!-- multiple -->
  <r3d:description language="eng">The Global Research Data Repository (GReDaR) is the place where researchers from all academic disciplines can
put their research data. Research data deposited in GReDaR</r3d:description> <!-- optional -->
  <r3d:type>disciplinary</r3d:type> <!-- multiple -->
  <r3d:size updated="2012-11-23">10.000 datasets; 323 studies</r3d:size> <!-- optional -->
  <r3d:startDate>2011-01-01</r3d:startDate> <!-- optional -->
  <r3d:endDate>2014-07-31</r3d:endDate> <!-- optional -->
  <r3d:repositoryLanguage>eng</r3d:repositoryLanguage> <!-- multiple -->
  <r3d:repositoryLanguage>deu</r3d:repositoryLanguage> <!-- multiple -->
  <r3d:repositoryLanguage>fra</r3d:repositoryLanguage> <!-- multiple -->
  <r3d:subject subjectScheme="DFG">31302 Oceanography</r3d:subject> <!-- multiple -->
  <r3d:subject subjectScheme="DFG">111 Social Sciences</r3d:subject> <!-- multiple -->
  <r3d:subject subjectScheme="DFG">20103 Cell Biology</r3d:subject> <!-- multiple -->
  <r3d:missionStatementURL>http://www.globalresearchdatarepository.org/mission_statement</r3d:missionStatementURL> <!-- optional -->
  <r3d:contentType contentTypeScheme="parse">Structured text</r3d:contentType> <!-- optional --> <!-- multiple -->
  <r3d:contentType contentTypeScheme="parse">Images</r3d:contentType> <!-- optional --> <!-- multiple -->
  <r3d:providerType>dataProvider</r3d:providerType> <!-- multiple -->
  <r3d:keyword>Earth Science</r3d:keyword> <!-- optional --> <!-- multiple -->
  <r3d:keyword>Measure Data</r3d:keyword> <!-- optional --> <!-- multiple -->
  <r3d:keyword>Social Science</r3d:keyword> <!-- optional --> <!-- multiple -->
  <r3d:keyword>Maps</r3d:keyword> <!-- optional --> <!-- multiple -->
  <r3d:institution> <!-- multiple -->
    <r3d:institutionName language="eng">Institute for Research Data</r3d:institutionName>
    <r3d:institutionAdditionalName language="eng">IRD</r3d:institutionAdditionalName> <!-- optional --> <!-- multiple -->
    <r3d:institutionAdditionalName language="deu">Institut für Forschungsdaten</r3d:institutionAdditionalName> <!-- optional -->
    <!-- multiple -->
    <r3d:institutionCountry>DEU</r3d:institutionCountry>
    <r3d:responsibilityType>general</r3d:responsibilityType> <!-- optional --> <!-- multiple -->
    <r3d:responsibilityType>technical</r3d:responsibilityType> <!-- optional --> <!-- multiple -->
    <r3d:institutionType>non-profit</r3d:institutionType> <!-- optional -->
    <r3d:institutionURL>http://www.ird.org</r3d:institutionURL> <!-- optional -->
  </r3d:institution>
</r3d:repository>
</r3d:re3data>
</r3d:re3data>
```

```

<r3d:institutionIdentifier>urn:47110815/ird</r3d:institutionIdentifier> <!-- optional --> <!-- multiple -->
<r3d:responsibilityStartDate>2011-01-01</r3d:responsibilityStartDate> <!-- optional -->
<r3d:responsibilityEndDate>2014-07-31</r3d:responsibilityEndDate> <!-- optional -->
<r3d:contact>John Doe, Institut für Forschungsdaten, Unter den Linden 6, 10099 Berlin, Germany</r3d:contact> <!-- optional -->
<!-- multiple -->
<r3d:contact>johndoe@globalresearchdatarepository.org</r3d:contact> <!-- optional --> <!-- multiple -->
<r3d:contact>http://www.globalresearchdatarepository.org/contactform/</r3d:contact> <!-- optional --> <!-- multiple -->
</r3d:institution>
<r3d:policy> <!-- optional --> <!-- multiple -->
<r3d:policyName>Data Management Policy</r3d:policyName>
<r3d:policyURL>http://www.globalresearchdatarepository.org/policy.html</r3d:policyURL>
</r3d:policy>
<r3d:databaseAccess>
  <r3d:databaseAccessType>restricted</r3d:databaseAccessType>
  <r3d:databaseAccessRestriction>registration</r3d:databaseAccessRestriction> <!-- optional --> <!-- multiple -->
</r3d:databaseAccess>
<r3d:databaseLicense> <!-- optional --> <!-- multiple -->
<r3d:databaseLicenseName>CC</r3d:databaseLicenseName>
<r3d:databaseLicenseURL>http://creativecommons.org/licenses/</r3d:databaseLicenseURL>
</r3d:databaseLicense>
<r3d:dataAccess> <!-- multiple -->
  <r3d:dataAccessType>restricted</r3d:dataAccessType>
  <r3d:dataAccessRestriction>registration</r3d:dataAccessRestriction> > <!-- optional --> <!-- multiple -->
</r3d:dataAccess>
<r3d:dataAccess> <!-- multiple -->
  <r3d:dataAccessType>closed</r3d:dataAccessType>
</r3d:dataAccess>
<r3d:dataLicense> <!-- multiple -->
  <r3d:dataLicenseName>CC</r3d:dataLicenseName>
  <r3d:dataLicenseURL>http://creativecommons.org/licenses/</r3d:dataLicenseURL>
</r3d:dataLicense>
<r3d:dataLicense> <!-- multiple -->
  <r3d:dataLicenseName>Copyrights</r3d:dataLicenseName>
  <r3d:dataLicenseURL>http://www.globalresearchdatarepository.org/copyrights/dataset1234</r3d:dataLicenseURL>
</r3d:dataLicense>
<r3d:dataUpload>
  <r3d:dataUploadType>restricted</r3d:dataUploadType>
  <r3d:dataUploadRestriction>registration</r3d:dataUploadRestriction> <!-- optional --> <!-- multiple -->
</r3d:dataUpload>
<r3d:dataUploadLicense> <!-- optional --> <!-- multiple -->
  <r3d:dataUploadLicenseName>GReDaR Data Deposit License</r3d:dataUploadLicenseName>
  <r3d:dataUploadLicenseURL>http://www.globalresearchdatarepository.org/deposit_license/</r3d:dataUploadLicenseURL>
</r3d:dataUploadLicense>
<r3d:software> <!-- optional --> <!-- multiple -->
  <r3d:softwareName>EPrints</r3d:softwareName>
</r3d:software>
<r3d:versioning>yes</r3d:versioning> <!-- optional -->
<r3d:api apiType="OAI-PMH">http://www.globalresearchdatarepository.org/oai/</r3d:api> <!-- optional --> <!-- multiple -->
<r3d:api apiType="REST">http://www.globalresearchdatarepository.org/rest</r3d:api> <!-- optional --> <!-- multiple -->

```

```
<r3d:pidSystem>DOI</r3d:pidSystem> <!-- multiple -->
<r3d:citationGuidelineURL>http://www.globalresearchdatarepository.org/how-to-cite.html</r3d:citationGuidelineURL> <!-- optional -->
<r3d:qualityManagement>yes</r3d:qualityManagement> <!-- optional -->
<r3d:certificate>DSA</r3d:certificate> <!-- optional --> <!-- multiple -->
<r3d:certificate>ISO16919</r3d:certificate> <!-- optional --> <!-- multiple -->
<r3d:syndication syndicationType="RSS">http://www.globalresearchdatarepository.org/news.rss</r3d:syndication> <!-- optional -->
<!-- multiple -->
<r3d:remarksIntern>This is internal information for the re3data.org team.</r3d:remarksIntern> <!-- optional -->
<r3d:remarksExtern>This is additional information on the research data repository visible to all users.</r3d:remarksExtern>
<!-- optional -->
<r3d:entryDate>2012-11-15</r3d:entryDate>
<r3d:lastUpdate>2012-12-21</r3d:lastUpdate>
<r3d:lastEditorIntern>John Doe</r3d:lastEditorIntern>
</r3d:repository>
</r3d:re3data>
```

## 4 Appendix

### 4.1 Attribute Values and Controlled Vocabularies

The following tables describe and define the allowed attribute values as well as the controlled vocabularies of the re3data.org metadata properties (see section 2).

#### 2.1 language

<i>Value</i>	<i>Definition</i>
Examples: eng, deu, fra	ISO-639-3; <a href="http://www.sil.org/iso639-3/codes.asp">http://www.sil.org/iso639-3/codes.asp</a>

#### 3.1 language

<i>Value</i>	<i>Definition</i>
Examples: eng, deu, fra	ISO-639-3; <a href="http://www.sil.org/iso639-3/codes.asp">http://www.sil.org/iso639-3/codes.asp</a>

#### 6.1 language

<i>Value</i>	<i>Definition</i>
Examples: eng, deu, fra	ISO-639-3; <a href="http://www.sil.org/iso639-3/codes.asp">http://www.sil.org/iso639-3/codes.asp</a>

#### 7 type

<i>Value</i>	<i>Definition</i>
disciplinary	Disciplinary repository (Subject repository): "This is a collection of research outputs with a common link to a particular subject discipline. Subject repositories are likely to cover one broad-based discipline, with contributors from many different institutions supported by a variety of funders; the repositories themselves are likely to be funded from one or more sources within the subject community. Although for some subject repositories the funding may be fragile, if they are of enough importance to the community then funding crises are usually weathered." <sup>7</sup>

<sup>7</sup> Jones, C et al. "Report of the Subject and Institutional Repositories Interactions Study." (2008): pag. 5



<i>Value</i>	<i>Definition</i>
institutional	Institutional Repository: "This is a collection of research outputs with a common link to a particular institution, usually by authorship. These repositories are likely to cover more than one research discipline, to have funders in many if not all the Research Councils and support communities who have different approaches to research dissemination. Whether deposit of content is mandatory is a decision that will be made by each institution. The institutions may have many requirements for the content of the repository, from open access dissemination, through metrics, marketing to strategic planning. It is likely that many of these processes in the past were undertaken through collection of bibliographic information." <sup>8</sup>
other	Other (Repository Type): A research data repository that is neither institutional nor disciplinary, e.g. a funder repository: "This is a collection of research outputs with a common link to one (or more) funder(s). These are likely to cover the funder's remit, which is usually subject-based but can become indistinct at subject boundaries, and will have authors from many institutions. Deposit of content is usually mandatory and can include project-related material, such as completion project reports. The funders will have requirements for the content of repository, from metrics through to strategic planning." <sup>9</sup>

### 8.1 updated

<i>Value</i>	<i>Definition</i>
YYYY or YYYY-MM-DD	ISO-8601; <a href="http://en.wikipedia.org/wiki/ISO_8601">http://en.wikipedia.org/wiki/ISO_8601</a>

### 9 startDate

<i>Value</i>	<i>Definition</i>
YYYY or YYYY-MM-DD	ISO-8601; <a href="http://en.wikipedia.org/wiki/ISO_8601">http://en.wikipedia.org/wiki/ISO_8601</a>

### 10 endDate

<i>Value</i>	<i>Definition</i>
YYYY or YYYY-MM-DD	ISO-8601; <a href="http://en.wikipedia.org/wiki/ISO_8601">http://en.wikipedia.org/wiki/ISO_8601</a>

### 11 repositoryLanguage

<i>Value</i>	<i>Definition</i>
Examples: eng, deu, fra	ISO-639-3; <a href="http://www.sil.org/iso639-3/codes.asp">http://www.sil.org/iso639-3/codes.asp</a>

<sup>8</sup> Jones, C et al. "Report of the Subject and Institutional Repositories Interactions Study." (2008): pag. 5

<sup>9</sup> Jones, C et al. "Report of the Subject and Institutional Repositories Interactions Study." (2008): pag. 5

## 12.1 subjectScheme

<i>Value</i>	<i>Definition</i>
DFG	DFG Classification of Subject Area, Review Board, Research Area and Scientific Discipline (using Level 1-4). <a href="http://www.dfg.de/download/pdf/dfg_im_profil/gremien/fachkollegien/dfg_fachsystematik_en_08_11.pdf">http://www.dfg.de/download/pdf/dfg_im_profil/gremien/fachkollegien/dfg_fachsystematik_en_08_11.pdf</a>

## 14 contentType

<i>PARSE.Insight type</i>	<i>Examples of File Formats</i>
Standard office documents	text documents, spreadsheets, presentations
Networkbased data	web sites, email, chat history, etc.
Databases	DBASE, MS Access, Oracle, MySQL, etc.
Images	JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.
Structured graphics	CAD, CAM, 3D, VRML, etc.
Audiovisual data	WAVE, MP3, MP4, Flash, etc.
Scientific and statistical data formats	SPSS, FITS, GIS, etc.
Raw data	device specific output
Plain text	TXT in various encodings
Structured text	XML, SGML, etc.
Archived data	ZIP, RAR, JAR, etc.
Software applications	modelling tools, editors, IDE, compilers, etc.
Source code	scripting, Java, C, C++, Fortran, etc.
Configuration data	parameter settings, logs, library files
Other	-

## 14.1 contentTypeScheme

<i>Value</i>	<i>Definition</i>
parse	For the values of parse see the PARSE.Insight types in contentType 14. For the source see page 5 in <a href="http://www.parse-insight.eu/downloads/PARSE-insight_survey_questions_research.pdf">http://www.parse-insight.eu/downloads/PARSE-insight_survey_questions_research.pdf</a>

## 15 providerType

<i>Value</i>	<i>Definition</i>
dataProvider	A research data repository is a data provider if it offers research data and its metadata (ideally exposing metadata via interfaces).
serviceProvider	A research data repository is a service provider if it harvests the metadata of research data from data providers as a basis for building value-added services.

## 17.1.1 language

<i>Value</i>	<i>Definition</i>
Examples: eng, deu, fra	ISO-639-3; <a href="http://www.sil.org/iso639-3/codes.asp">http://www.sil.org/iso639-3/codes.asp</a>

### 17.2.1 language

<i>Value</i>	<i>Definition</i>
Examples: eng, deu, fra	ISO-639-3; <a href="http://www.sil.org/iso639-3/codes.asp">http://www.sil.org/iso639-3/codes.asp</a>

### 17.3 institutionCountry

<i>Value</i>	<i>Definition</i>
Examples: DEU, GBR, USA	ISO-3166-1 Alpha-3 code; <a href="http://en.wikipedia.org/wiki/ISO_3166-1#Current_codes">http://en.wikipedia.org/wiki/ISO_3166-1#Current_codes</a>

### 17.4 responsibilityType

<i>Value</i>	<i>Definition</i>
funding	Funding responsibility refers to the institution that funds / funded the setup or running of the research data repository.
general	General responsibility refers to the institution that is generally responsible for the providing and running of the research data repository.
sponsoring	Sponsoring responsibility (in contrast to funding responsibility) refers to financial support of a research data repository in return for e.g. advertising on the website.
technical	General responsibility refers to the institution that is responsible for the technical matters of the research data repository, e.g. hosting.

### 17.5 institutionType

<i>Value</i>	<i>Definition</i>
commercial	A commercial institution is an institution that distributes surplus revenues as profit or dividends.
non-profit	The institution does not aim at economic profits but at non-profit (e.g. social, cultural and academic) goals of its members.

### 17.8 responsibilityStartDate

<i>Value</i>	<i>Definition</i>
YYYY or YYYY-MM-DD	ISO-8601; <a href="http://en.wikipedia.org/wiki/ISO_8601">http://en.wikipedia.org/wiki/ISO_8601</a>

### 17.9 responsibilityEndDate

<i>Value</i>	<i>Definition</i>
YYYY or YYYY-MM-DD	ISO-8601; <a href="http://en.wikipedia.org/wiki/ISO_8601">http://en.wikipedia.org/wiki/ISO_8601</a>

### 19.1 databaseAccessType

<i>Value</i>	<i>Definition</i>
open	There are no access barriers.
restricted	External users can overcome access barriers.
closed	External users cannot overcome access barriers.

## 19.2 databaseRestriction

<i>Value</i>	<i>Definition</i>
feeRequired	A single or regular payment is required.
registration	A free registration is required.
other	Other restrictions beside registration or payment such as a written consent are required.

## 20.1 databaseLicenseName

<i>Value</i>	<i>Definition</i>
CC	The database is licensed under a Creative Commons license. URL: <a href="http://creativecommons.org/licenses/">http://creativecommons.org/licenses/</a>
ODC	The database is made available under an Open Data Commons license. URL: <a href="http://opendatacommons.org/licenses/">http://opendatacommons.org/licenses/</a>
Copyrights	The database is copyright-protected. Any public or private use of this data is subject to prevailing copyright laws. Please contact the content provider of this data for permission requests. <sup>10</sup>
other	Licenses that are not explicitly listed.

## 21.1 dataAccessType

<i>Value</i>	<i>Definition</i>
open	There are no access barriers.
restricted	External users can overcome access barriers.
closed	External users cannot overcome access barriers.

## 21.2 dataAccessRestriction

<i>Value</i>	<i>Definition</i>
feeRequired	A single or regular payment is required.
registration	A free registration is required.
other	Other restrictions beside registration or payment such as a written consent are required.

## 22.1 dataLicenseName

<i>Value</i>	<i>Definition</i>
CC	The data is licensed under a Creative Commons license. URL: <a href="http://creativecommons.org/licenses/">http://creativecommons.org/licenses/</a>
ODC	The data is made available under an Open Data Commons license. URL: <a href="http://opendatacommons.org/licenses/">http://opendatacommons.org/licenses/</a>
OGI	The research data is licensed under an Open Government Licence. URL: <a href="http://www.nationalarchives.gov.uk/doc/open-government-licence/">http://www.nationalarchives.gov.uk/doc/open-government-licence/</a>

<sup>10</sup> Based on: "Copyrighted. Use Requirements to Be Specified Upon Request of the Copyright Holder." *The Cell: An Image Library - License*. Web. 13 Nov. 2012.

Value	Definition
RL	The research data is licensed under a Restrictive License. URL: <a href="http://www.ausgoal.gov.au/restrictive-licence-template">http://www.ausgoal.gov.au/restrictive-licence-template</a>
Copyrights	The data(base) is copyright-protected. Any public or private use of this data is subject to prevailing copyright laws. Please contact the content provider of this data for permission requests. <sup>11</sup>
other	Licenses that are not explicitly listed.
none	No license exists or was found.

### 23.1 dataUploadType

Value	Definition
open	There are no access barriers.
restricted	External users can overcome access barriers.
closed	External users cannot overcome access barriers.

### 23.2 dataUploadRestrictionType

Value	Definition
feeRequired	A single or regular payment is required.
registration	A free registration is required.
other	Other restrictions beside registration or payment such as a written consent are required.

### 25.1 softwareName

Value	Definition
CKAN	<a href="http://ckan.org">http://ckan.org</a>
Digital Commons	<a href="http://digitalcommons.bepress.com">http://digitalcommons.bepress.com</a>
DSpace	<a href="http://www.dspace.org">http://www.dspace.org</a>
EPrints	<a href="http://www.eprints.org">http://www.eprints.org</a>
eSciDoc	<a href="https://www.escidoc.org">https://www.escidoc.org</a>
OPUS	<a href="http://www.opus-repository.org/index.html">http://www.opus-repository.org/index.html</a>
dLibra	<a href="http://dlibra.psnc.pl">http://dlibra.psnc.pl</a>
other	The repository is using software, which is not listed in the controlled vocabulary.
unknown	The software of the research data repository is unknown.

### 27.1 apiType

Value	Definition
API	<a href="http://en.wikipedia.org/wiki/API">http://en.wikipedia.org/wiki/API</a>
FTP	<a href="http://en.wikipedia.org/wiki/FTP">http://en.wikipedia.org/wiki/FTP</a>
OAI-PMH	<a href="http://en.wikipedia.org/wiki/OAI-PMH">http://en.wikipedia.org/wiki/OAI-PMH</a>

<sup>11</sup> Based on: "Copyrighted. Use Requirements to Be Specified Upon Request of the Copyright Holder." *The Cell: An Image Library - License*. Web. 13 Nov. 2012.

<i>Value</i>	<i>Definition</i>
REST	<a href="http://en.wikipedia.org/wiki/REST">http://en.wikipedia.org/wiki/REST</a>
SOAP	<a href="http://en.wikipedia.org/wiki/SOAP">http://en.wikipedia.org/wiki/SOAP</a>
SPARQL	<a href="http://en.wikipedia.org/wiki/Sparql">http://en.wikipedia.org/wiki/Sparql</a>
SWORD	<a href="http://swordapp.org/about/a-brief-history/">http://swordapp.org/about/a-brief-history/</a>
other	An API that is not explicitly listed.

### 28 pidSystem

<i>Value</i>	<i>Definition</i>
ARK	<a href="https://wiki.ucop.edu/display/Curation/ARK">https://wiki.ucop.edu/display/Curation/ARK</a>
DOI	<a href="http://en.wikipedia.org/wiki/Digital_object_identifier">http://en.wikipedia.org/wiki/Digital_object_identifier</a>
hdl	<a href="http://www.handle.net/">http://www.handle.net/</a>
PURL	<a href="http://en.wikipedia.org/wiki/PURL">http://en.wikipedia.org/wiki/PURL</a>
URN	<a href="http://en.wikipedia.org/wiki/Uniform_Resource_Name">http://en.wikipedia.org/wiki/Uniform_Resource_Name</a>
other	A persistent identifier system that is not explicitly listed.
none	No persistent identifier is used to identify a research data set.

### 31 certificate

<i>Value</i>	<i>Definition</i>
DIN 31644	<a href="http://www.nabd.din.de/cmd?level=tpl-art-detailansicht&amp;artid=133059734&amp;committeeid=54738855&amp;bcrumblevel=2&amp;languageid=de">http://www.nabd.din.de/cmd?level=tpl-art-detailansicht&amp;artid=133059734&amp;committeeid=54738855&amp;bcrumblevel=2&amp;languageid=de</a>
DINI Certificate	<a href="http://nbn-resolving.de/urn:nbn:de:kobv:11-100182800">http://nbn-resolving.de/urn:nbn:de:kobv:11-100182800</a>
DSA	<a href="http://www.datasealofapproval.org/">http://www.datasealofapproval.org/</a>
ISO 16363	<a href="http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56510">http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56510</a>
ISO 16919	<a href="http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57950">http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57950</a>
TRAC	<a href="http://www.dcc.ac.uk/resources/tools-and-applications/trustworthy-repositories">http://www.dcc.ac.uk/resources/tools-and-applications/trustworthy-repositories</a>
WDS	<a href="http://www.icsu-wds.org/">http://www.icsu-wds.org/</a>
other	A certificate that is not explicitly listed.

### 32.1 syndicationType

<i>Value</i>	<i>Definition</i>
ATOM	<a href="http://en.wikipedia.org/wiki/Atom_%28standard%29">http://en.wikipedia.org/wiki/Atom_%28standard%29</a>
RSS	<a href="http://en.wikipedia.org/wiki/Rss">http://en.wikipedia.org/wiki/Rss</a>

### 35 entryDate

<i>Value</i>	<i>Definition</i>
YYYY or YYYY-MM-DD	ISO-8601; <a href="http://en.wikipedia.org/wiki/ISO_8601">http://en.wikipedia.org/wiki/ISO_8601</a>

36 lastUpdate

<i>Value</i>	<i>Definition</i>
YYYY or YYYY-MM-DD	ISO-8601; <a href="http://en.wikipedia.org/wiki/ISO_8601">http://en.wikipedia.org/wiki/ISO_8601</a>

## 4.2 Access To Research Data Repositories Explained

Beside accessibility of research data the terms of use are probably the most important issue of a research data repository. To define the general accessibility of a research data repository the term access has to be split into three categories representing different levels of access. First there is the access to the research data repository defining whether users can access the database in general. Secondly there is the access to the research data sets in the research data repository in particular. Thirdly there is the access to upload research data sets to the research data repository. The access to each level can be open, restricted and/or closed. Open means that there are no access barriers. Restricted means that external users can overcome access barriers. Closed means external users cannot overcome access barriers.

As shown in the matrix below the access to the research data repository is the basic level to define the general framework of access to the actual research data sets, e.g. a research data set in a research data repository that is restricted through registration cannot be open but just restricted or even closed; a research data set in a closed research data repository which underlies access restrictions that cannot be overcome by external users is neither open nor restricted but only closed. Thus to define the general accessibility of a research data repository the values highlighted in orange are needed (values in braces can occur).

<i>Access (property)</i>	<i>Open Access</i>	<i>Restricted Access</i>	<i>Closed Access</i>
Access to Repository (19.1 databaseAccessType)	open	open or restricted	closed
Access to Data (21.1 dataAccessType)	open (restricted, closed)	restricted (closed)	closed
Data Upload (23.1 dataUploadType)	open or restricted	open or restricted	-