



Conceptual Model for User Stories

Version 1.0, August 2021

doi: <https://doi.org/10.48440/re3.012>

Authors: [Paul Vierkant](#)¹, [Roland Bertelmann](#)², [Helena Cousijn](#)¹, [Kirsten Elger](#)⁴, [Lea Maria Ferguson](#)², [Hans-Jürgen Goebelbecker](#)⁵, [Maxi Kindling](#)³, [Gabriele Kloska](#)⁵, [Thanh Binh Nguyen](#)⁵, [Heinz Pampel](#)², [Vivien Petras](#)³, [Rouven Schabinger](#)⁵, Edeltraud Schnepf⁵, Angelika Semrau⁵, [Dorothea Strecker](#)³, [Margarita Trofimenko](#)⁵, [Robert Ulrich](#)⁵, [Arne Upmeier](#)⁵, [Nina Weisweiler](#)², [Yi Wang](#)³, [Michael Witt](#)⁶

¹ [DataCite - International Data Citation Initiative e. V.](#)

² [Helmholtz-Gemeinschaft / Helmholtz Association](#), Helmholtz Open Science Office, Germany

³ [Humboldt-Universität zu Berlin](#), Berlin School of Library and Information Science, Germany

⁴ [GFZ German Research Centre for Geosciences / Helmholtz-Zentrum Potsdam – Deutsches GeoForschungsZentrum \(GFZ\)](#), Germany

⁵ [Karlsruher Institut für Technologie / Karlsruhe Institute of Technology \(KIT\)](#), Germany

⁶ [Purdue University](#), United States

Contact:

info@re3data.org

<https://www.re3data.org>

Except where otherwise noted, this work is licensed under <http://creativecommons.org/publicdomain/zero/1.0/>



Table of Contents

Introduction	2
Description of re3data	2
History and governance	2
Editorial board	3
Technical infrastructure	3
Metadata	4
Target groups using re3data	6
User stories	7
Search and Discover Research Data Repositories	7
Reuse re3data metadata	8
Administrate re3data records	9
Reference re3data metadata	10
Conclusion	10

1. Introduction

The primary goal of the conceptual model is to have a structured description of the service's architecture to improve and increase the usage of re3data (and its metadata) by third party services. Embedding re3data within the various stakeholder communities will advance re3data as a global information infrastructure for trustworthy research data repositories. A research data repository (RDR) is a subtype of a sustainable information infrastructure providing long-term storage of and access to research data. Research data are information objects generated in scholarly contexts, for example through experiments, measurements, surveys or interviews.

re3data is a global registry of research data repositories that covers research data repositories from all academic disciplines. It presents repositories for the permanent storage of and access to research data sets for researchers, funding bodies, publishers, and scholarly institutions. re3data promotes a culture of sharing, increased access to, and better visibility of research data and other research outputs. re3data is a partner service of DataCite¹, a global not-for-profit organization that is actively involved in several initiatives to improve the availability and citation of research outputs.

The re3data Conceptual Model for User Stories describes what kind of value the registry offers to its users. The model depicts the interactions between the service offered and its users based on use cases. It was released in 2021 in the context of the DFG-funded project re3data COREF.²

2. Description of re3data

a. History and governance

The registry re3data was developed within a research project of the same name funded by the German Research Foundation (DFG)³ and went online in the fall of 2012. The original project partners were the Berlin School of Library and Information Science at the Humboldt-Universität zu Berlin⁴, the Helmholtz Open Science Office of the Helmholtz Association, the Library and Information Services department (LIS) of the GFZ German Research Centre for Geosciences⁵, and the KIT Library at the Karlsruhe

¹ <https://datacite.org/>

² <https://gepris.dfg.de/gepris/projekt/422587133?language=en>

³ <https://www.dfg.de/>

⁴ <https://www.ibi.hu-berlin.de/>

⁵ <https://bib.telegrafenberg.de/>

Institute of Technology (KIT)⁶. In March 2014, re3data merged with Databib, a similar initiative at Purdue University⁷, into one service that is partnered with DataCite since the end of 2015. The aim of this cooperation was to serve the research community with a single, sustainable registry of RDR that incorporates the best features of both initiatives.

b. Editorial board

The re3data international editorial board indexes, reviews and updates the metadata of all RDR registered in re3data. Researchers, research data specialists and librarians from various countries are members of the international editorial board.

c. Technical infrastructure

Technical infrastructure is the backbone of re3data's service. It offers interfaces through which different kinds of users can access and use the information provided by the registry. Since re3data was designed primarily for researchers, users can access the registry through a web user interface (WUI). To enable other stakeholders in the realm of research data management, re3data offers information through an application programming interface (API). Through these interfaces, third party services can extract customized information about RDRs and make it available to their users.

i. Web user interface (WUI)

The re3data WUI offers four main functionalities, searching, filtering, browsing and metrics:

1. Users can search the registry right on the home page using the single search field or using the search functionality on the results page. Both search fields allow keyword searches in any field of the re3data metadata schema. The search result displays a list of research data repositories showing relevant information like a short description, subjects, content types and the country the repository operator is based in. By clicking on the name of the research data repository users reach the landing page of the respective service that provides general information about the repository, its hosting institutions and the terms and standards of the service. To facilitate finding appropriate RDRs, re3data developed a set of icons. These icons are displayed for a

⁶ <https://www.bibliothek.kit.edu/>

⁷ <https://www.lib.purdue.edu/>

respective entry if an RDR provides important information concerning the repository, e.g., if it acquired a certificate or supports the provision of persistent identifiers.

2. The search results can be filtered according to different aspects of a repository, e.g., subjects, content types, data upload etc.
3. In addition to searching and filtering, users can browse through the registry by subject. This browsing can be done either visually through an infographic or by choosing a subject from a list. Browsing the registry by content type is also possible by choosing one from a controlled list. The third option to browse is by country. This can be done by either choosing a country from a map or from a list of countries.
4. In the metrics section users are provided with charts based on single metadata fields. Users can see how many repositories provide information about different aspects, e.g., PID systems, access terms, quality management. etc.

ii. API

In addition to the WUI, re3data offers an API which allows users or institutions to retrieve content from the registry.⁸ Currently the platform offers a simple open search implementation as well as a basic version of a RESTful interface.

iii. Metadata schema

The re3data Metadata Schema contains metadata properties describing an RDR, such as its general scope, content, and infrastructure as well as its compliance with technical, quality, and metadata standards. The schema includes required metadata properties and optional properties providing additional information. The schema serves the purpose of:

- recommending a standard for describing an RDR;
- providing the basis for interoperability between RDRs, re3data and other infrastructures (e.g. DataCite Commons, OpenAIRE, etc.);
- helping RDRs move towards shared standards and common practices.

⁸ <https://www.re3data.org/api/doc>

The initial version 1.0 of the schema (formerly called vocabulary) was developed and tested based on a small sample of RDRs. After version 1.0 of the schema was published in July 2012, the re3data project team issued a public request for comments. Version 2.0 of the schema considered all responses as well as current developments in the area of RDRs, resulting in substantial changes in the structure of the schema. Versions 2.1 and 2.2 introduced minor changes, as well as an outline of the re3data Registration Policy.⁹ Version 3.0 included changes of the re3data Registration Policy as well as structural adjustments to better reflect changes within the landscape of RDRs.¹⁰ Following another public request for comments, the current version 3.1 includes a differentiation of certification information as well as the option to create profiles, in addition to minor changes.¹¹

iv. Curation workflow

Anyone can suggest RDRs to be listed in re3data via an application form providing information such as the name and URL of the RDR.¹² This information is reviewed and enhanced by the editorial board.¹³ Members of the editorial board analyze the RDR's website and describe it based on an indexing handbook. A repository is indexed if the minimum requirements of the re3data Registration Policy are met.¹⁴ Before a new record of an RDR is published in re3data, all gathered information is reviewed by a second editor.

To update a re3data entry, a user can use the web form suggesting additions and/or changes concerning the entry. After a review by the re3data editorial board, the entry will be updated.

The major purpose of the registry is to improve the discoverability of RDRs. An RDR is a subtype of a sustainable information infrastructure providing long-term storage of and access to research data. Research data are information objects generated in scholarly contexts, for example through experiments, measurements, surveys or interviews.

An RDR listed in re3data is either:

- a data provider if it offers research data and its metadata (ideally exposing metadata via interfaces),

⁹ <https://www.re3data.org/schema/2-2>

¹⁰ <https://www.re3data.org/schema>

¹¹ <https://doi.org/10.48440/re3.010>

¹² <https://www.re3data.org/suggest>

¹³ <https://www.re3data.org/editorialboard>

¹⁴ <https://www.re3data.org/suggest>

or

- a service provider (e.g., a portal) if it harvests the metadata of research data from data providers as a basis for building value-added services,

or

- both.

To be registered in re3data, an RDR must comply with the following minimum requirements:

A repository must

- have a focus on research data;
- be operated by a legal entity with an organizational framework that provides sustainability (e.g., library, university);
- clarify access conditions to the repository and research data and provide terms of use.

v. PID for re3data repository records

Within re3data each record is assigned a DOI in order to enable persistent citation for the metadata record of the respective repository indexed in re3data. Each DOI is displayed in the WUI below the repository description as part of a citation recommendation or can be retrieved via the API.

3. Target groups using re3data

re3data serves the needs of different target groups. The service is primarily aimed at:

- Researchers looking for either an adequate repository to upload their data or for research data to work with or reuse.
- Information specialists as e.g., librarians, data scientists, and data stewards
- RDR operators seeking to curate and update their service's re3data record.
- Publishers employing re3data as a reference for suggesting RDRs on their websites and in their policies, either by pointing to the registry or by using the API to display a list of repositories on their websites.

- Funders suggesting RDRs to their researchers either by linking to the registry or by connecting their research data management planning tools to the re3data API.
- Third parties like other registries, databases or organizations using re3data as a source to enrich the content of their services with the information provided in re3data's records.

4. User stories

a. Search and Discover Research Data Repositories

i. Existing:

As a **researcher**, I want to find an appropriate RDR so that I can **deposit my research data** in compliance with funder, institutional or publisher requirements. Given that researchers are looking for relevant filtering criteria when searching for RDR, re3data provides **filtering and browsing options**.

1. As a **researcher**, I want to find RDR of my research field so that I can **find research data relevant to my research** (e.g., reuse, reproduction, inspiration). Given that researchers are looking for relevant filter criteria when searching for RDR, they are provided with **filtering and browsing options**.
2. As an **information specialist**, I want to find an appropriate RDR so that I can **recommend RDR** to researchers or other end users. Given that information specialists are looking for relevant filter criteria when searching for RDR, they are provided with **filtering and browsing options**.

ii. Future:

1. As a **researcher**, I want to find an appropriate RDR so that I can **deposit my research data** in compliance with funder, institutional or publisher requirements. Given that researchers are looking for relevant filter criteria when searching for RDR, they should be provided with **filtering and browsing options** as well as **predefined subsets of criteria** (profiles, e.g., FAIR criteria).

2. As a **researcher**, I want to find RDR of my research field so that I can **find research data relevant to my research** (e.g., reuse, reproduction, inspiration). Given that researchers are looking for relevant filter criteria when searching for RDR, they should be provided with **filtering and browsing options** as well as **predefined subsets of criteria (profiles)**.
3. As an **information specialist**, I want to find an appropriate RDR so that I can **recommend RDR** to researchers or other end users. Given that information specialists are looking for relevant filter criteria when searching for RDR, they should be provided with **filtering and browsing options** as well as **predefined subsets of criteria (profiles)**.

b. Reuse re3data metadata

i. Existing:

1. As a **registry, database or discovery system**, I want to **reuse the re3data metadata** about RDR so that I can **enrich the metadata of my service**. Given that these services expect up-to-date data when fetching the metadata, this metadata is provided by re3data in an **automated process via an API**.
2. As a **publisher, funder or organization**, I want to **reuse the re3data metadata** about RDR so that I can **provide a list of recommended RDR** to my authors, grantees or affiliates. Given that these stakeholders expect up-to-date data when fetching the metadata, this metadata is **provided by re3data in an automated process via an API**.
3. As a **funder or organization**, I want to **reuse the re3data metadata** about RDR so that I can **monitor the RDR of my grantees or organization**. Given that these stakeholders expect up-to-date data when fetching the metadata, this metadata is **provided by re3data in an automated process via an API**.

ii. Future:

1. As a **publisher, funder or organization**, I want to **reuse the re3data metadata** about RDR so that I can **provide a list of recommended RDR based on predefined subsets of criteria** (profiles, e.g., FAIR criteria) to my authors, grantees or affiliates. Given that these

stakeholders expect up-to-date data when fetching the metadata, this metadata should be **provided by re3data in an automated process via an API**. The **created profiles should be fed back into re3data** so that users can reuse them.

c. Administrate re3data records

i. Existing:

1. As a **user**, I want to be able to **suggest changes to an entry** of an RDR in re3data so that I can improve its metadata and visibility. Given that users make these suggestions on an ad-hoc basis when searching or browsing the registry, this **suggestion function is possible via WUI**.

ii. Future:

1. As a **researcher**, I want to be sure that the **metadata in re3data is authenticated** so that I can trust it. Given that users want quality assured metadata when searching for RDR, the **administration of RDR entries in re3data should be attributable via ORCID**.
2. As a **repository operator**, I want to be able to **edit the entry of my RDR** so that I can improve its metadata and visibility. Given that repository operators have different capabilities when editing their metadata, then this process should be **authenticated and possible via WUI and API**.
3. As a **repository operator**, I want to be able to **add standardized organizational metadata of my RDR** so that I disambiguate my organization that runs the RDR. Given that repository operators want to uniquely identify their organization when editing their metadata, then this process should be possible **using the ROR ID**.
4. As a **certification agency**, I want to be able to push metadata into an entry of RDR so that I can **enrich the entry with metadata that creates trustworthiness**. Given that manual editing of RDR metadata in re3data is time consuming when updating RDR metadata, the updated information should be **pushed into re3data in a quality-assured automated process via an API**.

d. Reference re3data metadata

i. Existing:

1. As a **researcher**, I want to be able to **refer to an RDR** so that I can attribute the location of my research data properly. Given that researchers have difficulties when referencing their research data, **referencing the respective re3data landing page of the RDR is made as easy as possible using an identifier.**
2. As a **repository operator**, I want researchers to be able to **refer to the RDR of my organization** so that my service gets credit and higher visibility. Given that it is difficult to uniquely identify a RDR, **referencing the respective re3data landing page of the RDR is made as easy as possible using an identifier.**
3. As a **registry, database, publisher, or discovery system**, I want to be able to **uniquely identify a RDR so that I can reference that RDR.** Given that it is difficult to uniquely identify a RDR, **the reference of the respective RDR via its re3data record is made as easy as possible using an identifier.**

5. Conclusion

For almost a decade re3data has been a global source of information on RDR. This model is to make sure that the registry is embedded in the future landscape of open science tools and serves the needs of all of its users, institutions and individuals alike. To reflect ongoing changes in this landscape and the design of RDRs, re3data's conceptual model is revised iteratively. This iterative process is based on constant feedback from the RDR community that is evaluated regularly by re3data.