re3data Stakeholder Survey and Workshop Report

August 2021

DOI: https://doi.org/10.48440/re3.013

Authors:

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

[1] Humboldt-Universität zu Berlin, Berlin School of Library and Information Science, Germany
[2] Helmholtz Association, Helmholtz Open Science Office, Germany
[3] DataCite - International Data Citation Initiative e. V.
[4] German Research Centre for Geosciences (GFZ)
[5] Karlsruher Institut für Technologie / Karlsruhe Institute of Technology (KIT), Germany
[6] Purdue University, United States

re3data COREF project partners:

DataCite - International Data Citation Initiative e. V.
Berlin School of Library and Information Science at the Humboldt-Universität zu Berlin
Helmholtz Association's Open Science Office at the GFZ German Research Centre for Geosciences
KIT Library at the Karlsruhe Institute of Technology (KIT)

Contact:

info@re3data.org
https://www.re3data.org

Table of Content

1. Introduction to the re3data Service

Established in 2012, the Registry of Research Data Repositories - re3data[1] is a globally used service providing extensive information about more than 2700 research data repositories (RDR) in August 2021. The service aims to make data repositories accessible in a web-based directory and thus provide orientation on existing data collections. The central concern of re3data is to give researchers an orientation in the heterogeneous landscape of research data repositories. It is aimed both at users who want to make their research data available or are requested to do so, e.g., by their funding agencies or scientific journals, and at information seekers, such as researchers, looking for reusable datasets. It also provides an overview of the landscape of research data repositories to infrastructure services such as data centers, computing centers, and libraries. In addition, the service can be used by funding organisations: If they ask their beneficiaries to make the research data they collect openly available, a search of re3data can help locate potential repositories to store the research data. Furthermore, re3data is an important metadata resource for third party services that process and reuse re3data content to support other services (e.g., the European Open Science Monitor[2]).

2. The re3data COREF Project

From January 2020, the German Research Foundation (DFG) has been funding the project “Community Driven Open Reference for Research Data Repositories” (re3data COREF[3]) over a period of 36 months. The main goals of the project are connecting re3data as reference for research data repositories with other services and infrastructures and developing and enhancing its features according to the needs of the community. The fine-grained re3data Metadata Schema, which contains numerous properties to describe RDR – such as their general scope, content, and infrastructure, or their compliance with technical, quality, and metadata standards – provides a widely accepted standard for the description of RDR and already serves as a basis for interoperability among RDR, re3data, and other infrastructures.

During the project period, re3data COREF will implement the following measures to achieve the project’s objectives:

---

[1] https://www.re3data.org/ [Retrieved on 30.08.2021]

[2] https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/facts-and-figures-open-research-data_en [Retrieved on 30.08.2021]

[3] https://coref.project.re3data.org/project [Retrieved on 30.08.2021]

- Provide customizable and extendable core repository descriptions that are persistently identifiable and can be referred to and cited in an appropriate manner. For this purpose, the project members continue to work on the update and expansion of the current re3data metadata schema (see the latest version 3.1[4]).
- Develop a model of trust for authorized editing and adding of metadata, which will enable further options for automated data exchange between services and machine-to-machine communication (e.g., to automatically include certification information from repository certification organizations like CoreTrustSeal[5]).
- Advance the overall infrastructure and build easily embeddable widgets and tools that take recent requirements of stakeholders using the re3data metadata and API into account.
- Provide more sophisticated functions for monitoring and recommendation, for example of repositories that support the implementation of the FAIR Data Principles.
- Expand and strengthen collaboration and communication with stakeholders within the research data community.
- Conduct a study on the status quo of quality assurance measures and standards in the context of research data repositories. Results from this study will guide further actions for the advancement of the re3data metadata schema and editorial process.
- Implement the use of authority files and persistent identifier systems like the Open Researcher and Contributor ID (ORCID[6]) or the Research Organization Registry (ROR)[7].

re3data COREF is a joint project of the Berlin School of Library and Information Science[8] at the Humboldt-Universität zu Berlin, DataCite e.V.[9], the Helmholtz Open Science Office[10] at the German Research Centre for Geosciences (GFZ), and the KIT Library[11] at the Karlsruhe Institute of Technology (KIT).

---

[4] Strecker, D., Bertelmann, R., Cousijn, H., Elger, K., Ferguson, L. M., Fichtmüller, D., Goebelbecker, H.-J., Kindling, M., Kloska, G., Nguyen, T. B., Pampel, H., Petras, V., Schabinger, R., Schnepf, E., Semrau, A., Trofimenko, M., Ulrich, R., Upmeier, A., Vierkant, P., Weisweiler, N. L., Wang, Y., Witt, M. (2021): Metadata Schema for the Description of Research Data Repositories: version 3.1. https://doi.org/10.48440/re3.010

[5] https://www.coretrustseal.org/ [Retrieved on 30.08.2021]

[6] https://orcid.org/ [Retrieved on 30.08.2021]

[7] https://ror.org/ [Retrieved on 30.08.2021]

[8] https://www.ibi.hu-berlin.de/ [Retrieved on 30.08.2021]

[9] https://datacite.org/ [Retrieved on 30.08.2021]

[10] https://os.helmholtz.de/ [Retrieved on 30.08.2021]

[11] https://www.bibliothek.kit.edu/ [Retrieved on 30.08.2021]

3. re3data Stakeholder Survey and Workshop

In order to channel and align the efforts within the COREF project, re3data is revising its conceptual service model according to the most important use cases of the various stakeholders working with re3data. Adopting and reflecting current developments in the research data landscape, the update of the service architecture in COREF is based on a bottom-up approach taking up the results from a stakeholder survey and a stakeholder workshop in November 2020 (see appendices for the questionnaire and slides).

In the three workshop sessions in November 2020, the participants were invited to hold short presentations, following the basic structure of the survey questionnaire and elaborating further on their individual use case scenarios and requirements. The presentations led into group discussions and allowed for detailed questions and discussions that helped the re3data COREF team gain a deeper understanding of the survey results.

Stakeholders were invited to share their experiences, impressions, and requirements based on typical use cases involving the re3data Metadata Schema, technical interfaces, and other features. The findings from the survey and workshops were incorporated in the process of developing a conceptual model for user stories, which embeds the registry within the research community and the infrastructure landscape to meet the emerging needs for a trusted repository reference. This ensures that the interests of the re3data community and stakeholders are adequately represented within the updated service architecture and helps the re3data COREF team to coordinate development decisions.

As a result of this preliminary work, version 1.0 of the Conceptual Model for User Stories[12] has been released together with the present report in August 2021.

4. Participants

The re3data COREF project invited a number of partner organizations to participate in the survey and workshops, who have used or continue to use re3data metadata or who have collaborated with re3data in various other ways. A total of 21 stakeholders participated in the survey and 16 of them joined the workshops. Most of the participants represent research infrastructures, as re3data serves as an

[12] Vierkant, P., Bertelmann, R., Cousijn, H., Elger, K., Ferguson, L. M., Goebelbecker, H.-J., Kindling, M., Kloska, G., Nguyen, T. B., Pampel, H., Petras, V., Schabinger, R., Schnepf, E., Semrau, A., Strecker, D., Trofimenko, M., Ulrich, R., Upmeier, A., Weisweiler, N. L., Wang, Y., Witt, M. (2021): Conceptual Model for User Stories: version 1.0. https://doi.org/10.48440/re3.012

important metadata resource for service providers. A central goal of the re3data COREF project is to further connect re3data – as the reference point for research data repositories – with other relevant services and infrastructures. Accordingly, the needs of these stakeholders are central to further developing the re3data service.

The following partners participated in the survey and the workshop sessions:

Stakeholder	**Participation**
American Geophysical Union (AGU)	Survey & Workshop
Canadian Association of Research Libraries and the Portage Network	Survey & Workshop
CoreTrustSeal	Survey & Workshop
Database Information System (DBIS)	Survey & Workshop
Digital Research Infrastructure for the Arts and Humanities (DARIAH)	Survey & Workshop
CLARIN European Research Infrastructure Consortium (ERIC)	Survey & Workshop
FAIRsFAIR Project	Survey & Workshop
Leibniz Information Centre for Economics (ZBW)	Survey & Workshop
Open Access Office of Berlin	Survey & Workshop
OpenAIRE	Survey & Workshop
ORCID	Survey & Workshop
Research Data Management Organiser (RDMO)	Survey & Workshop
SciCrunch (Research Resource Identification (RRID) Initiative)	Survey & Workshop
Springer Nature	Survey & Workshop
University of Basel / University Library	Survey & Workshop
ZB MED - Informationszentrum Lebenswissenschaften	Survey & Workshop
B2FIND	Survey
EOSC / FAIR Working Group	Survey
Imperial College London	Survey
RADAR (FIZ Karlsruhe)	Survey
Research Organization Registry (ROR)	Survey
Swiss National Science Foundation (SNSF), Open Research Data-Working Group (ORD group)	Survey

Table 1: re3data stakeholder survey and workshop participants

5. Survey Design

To capture the heterogeneous needs of re3data users regarding its service, a retrospective gap analysis served as a starting point for the revision of the service. As a first step, existing use cases were analyzed and resulted in a categorization and clustering of user groups and their interactions with and expectations towards re3data. This preliminary work determined the structure of the questionnaire for the stakeholder survey. The survey contained open questions on the following areas of use:

- Seeking information about research data repositories
- Data export from re3data
- Data import into re3data
- Monitoring of the repository landscape
- Referencing of research data repositories
- Recommendation of research data repositories
- Other (to be specified)

Subsequent to the survey, participants were invited to the joint virtual workshop where they presented their individual use cases and discussed them with the re3data COREF team and the other participants. The workshop was divided into three sessions, each with 4 to 6 stakeholders. The sessions were held online on November 24 and 26, 2020.

6. Key Findings and Take-Aways

Over the course of the COREF project, updates to the re3data service architecture will be implemented to reflect and incorporate current developments in the research data landscape. Several key observations, needs, and recommendations for the re3data service emerged from analyzing the survey. These were clustered and summarized to provide an overview of the requirements related to the user stories and applications of the re3data service.

6.1 The Bigger Picture: How re3data Contributes to the International Network of Research Data Infrastructures

re3data serves several functions for the research data community, not only by enabling a comprehensive discovery of research data repositories, but also by adding value to other infrastructure services. Some of the general contributions that could be derived from the workshops and survey responses are:

- re3data increases the visibility of repository services towards the research community.
- re3data facilitates monitoring of the research data infrastructure landscape.
- re3data is recommended to research communities and universities to support best practices in research data management.
- re3data metadata is reused to build repository recommendation tools with domain-specific, regional, or other focus.
- re3data is used to retrieve metadata about RDR as a basis for various other purposes and infrastructure services.
- The re3data subject classification (which is based on the "DFG Fachsystematik"[13]) has been adopted by several other services.
- re3data fosters the implementation of the FAIR Data Principles through close collaboration with initiatives, organisations, and projects from the research data community.
- re3data metadata is reused in DMP-Tools to support researchers in the creation of data management plans.
- API and Web User Interface (WUI) are equally important entry points for accessing re3data content and adding new information to the registry.

6.2 Zooming In: Requirements and Recommendations for re3data

One goal of the survey and workshop was to conduct a gap analysis and capture stakeholder requirements not currently covered by the re3data service. These recommendations provide an important orientation for re3data's development roadmap on a more detailed level. Table 2 lists the general recommendations from the survey responses and workshop discussions along with the proposed solutions. Table 3 displays recommendations for adjustments of and additions to the re3data Metadata Schema.

[13] https://www.dfg.de/dfg_profil/gremien/fachkollegien/faecher/

6.2.1 General Recommendations

Recommendation	**Method of resolution**
Enhance and improve the re3data API (subset export, enable queries with OR operator, CoreTrustSeal integration).	Will be addressed in re3data COREF, Work Package 3 (Technical Infrastructure).
Review and enhance the subject classification.	Review of the subject classification is currently in progress, as part of re3data COREF Work Package 2 (Metadata Schema).
Provide more information about a repository's curation activities and data quality management. Include a clear statement about the level of storage / curation / preservation provided by the repository.	As part of re3data COREF, the survey "Data quality management at research data repositories"[14] has been conducted. The findings will inform the revision of the re3data metadata schema.
Provide usage statistics.	Usage figures are not currently collected for data protection reasons.
Closer alignment with EOSC service catalogs.	re3data is seeking to be registered in the EOSC Portal.
Add a dashboard with CSV export option.	With the further development of the data model of re3data and the planned implementation of the RDF standard, further export formats such as JSON-LD and CSV will be made possible, including the necessary user interfaces.
(Automated) periodic updates of metadata (e.g., number of items contained in the RDR).	Requires suitable repository APIs. This option is being discussed as part of the re3data COREF project work.
Allow direct editing and addition of community-generated content for authorized third parties (and make visible which data has been externally validated	re3data COREF works towards the development of a Metadata Model of Trust to allow for direct integration of repository information via authorized third parties, e.g. CoreTrustSeal.

[14] https://coref.project.re3data.org/blog/how-to-ensure-good-data-a-presentation-at-open-repositories-2021 [Retrieved on 30.08.2021]

using a standard model for third-party validation of repository metadata, particularly with respect to certification information). Enable verification of indexed repository attributes (especially names) to keep the information up to date.	
Allow creation / flagging of subset collections, based on filtering criteria as determined by authorized stakeholders (e.g. endorsement by funders, journals etc).	re3data COREF develops the foundations to enable the creation of community profiles within re3data. Such profiles will allow the selection of subsets of RDR according to community-defined criteria.
Add a community subsite for users to connect and find out more about re3data's activities and partners.	Will be addressed in re3data COREF, Work Package 5 (Community Building).
Improve usability of Repository Finder Tool and re3data.	The Repository Finder will be integrated into DataCite Commons.[15]
Implement ROR IDs for institutions in re3data.	Already integrated for DataCite member organizations. Will be further addressed in re3data COREF.
Implementation of MakeDataCount, and the relevant metadata concerning average citations per repository, or other usage metric.	This idea is being discussed among the re3data COREF team for the revision of the metrics field in the MetaData Schema.

Table 2: General recommendations for the further development of the re3data service

[15] See Ulrich, R., Weisweiler, N. L., Wimalaratne, S., & Witt, M. (2021). 17th RDA Plenary Poster Presentation: re3data - Discovering FAIR-enabling Repositories. RDA 17th Plenary Meeting (RDA VP17), Edinburgh (Virtual). https://doi.org/10.5281/zenodo.4705209

6.2.2 Recommendations for Adjustments of the re3data Metadata Schema

Some of the following recommendations could already be implemented in version 3.1 of the schema, which has been published recently[16]. All remaining suggestions will be analyzed and considered for the upcoming revisions.

Extend the vocabulary for content types.
Add information about whether a repository is equipped to handle sensitive data.
Link out to community and cross-domain (metadata) standards.
Specify relationships between repositories.
Add ORCID, IGSN, and ROR to PID System options.
Align repository type with CoreTrustSeal[17] (or create a common vocabulary).
Add information on any service fees charged by a repository.
Add information on how quickly the data can be curated / preserved.
Add information about the source of data submission (e. g. data is submitted by a person or an instrument).
Distinguish between machine-readable and human-readable interfaces (e. g., SPARQL endpoint vs. SPARQL Query Editor).
Specify subject areas of metadata standards.
Add more precise geo-location information.
Add information on the terms of use, the designated user community, and preconditions to use the repository.
Add information on whether a repository tracks data use and on how / if the data is being (re-)used.

Table 3: Recommended additions for the re3data Metadata Schema

[16] https://coref.project.re3data.org/blog/releasing-version-3-1-of-the-re3data-metadata-schema [Retrieved on 30.08.2021]

[17] https://www.coretrustseal.org/why-certification/specialists-generalists-technical-repository-service-providers/ [Retrieved on 30.08.2021]

7. Integration into the Conceptual Model for User Stories for re3data

The various use scenarios and requirements communicated via the survey and workshop discussions formed an important basis for the development of re3data's Conceptual Model for User Stories[18]. The model provides a structured description of the service's architecture and depicts the possible interactions between the service and its users; numerous user stories illustrate existing and potential future use cases of the re3data service. During the model development process, the usage areas were restructured and clustered into the following high-level interactions covering all identified re3data use cases:

- Search and discover Research Data Repositories
- Reuse re3data metadata
- Administrate re3data records
- Reference re3data metadata

Based on the needs of the stakeholder community, the model will serve to guide the development of the re3data service during the re3data COREF project and beyond. Nevertheless, the conceptual model will not be fixed, but will be adjusted based on new input from the community and partner organizations.

We, the re3data COREF team, thank all participants of the survey and workshop sessions for their valuable contributions!

[18] Vierkant, P., Bertelmann, R., Cousijn, H., Elger, K., Ferguson, L. M., Goebelbecker, H.-J., Kindling, M., Kloska, G., Nguyen, T. B., Pampel, H., Petras, V., Schabinger, R., Schnepf, E., Semrau, A., Strecker, D., Trofimenko, M., Ulrich, R., Upmeier, A., Weisweiler, N. L., Wang, Y., Witt, M. (2021): Conceptual Model for User Stories: version 1.0. https://doi.org/10.48440/re3.012

Appendices

A. Survey Questionnaire

1. Introduction

1.1 What is your current role / position at your institution?

2. Use cases

2.1 Information

2.1.1 What information on research data repositories do you need at your organization?

2.1.2 What information on research data repositories is currently missing (i.e., in general / for your organization) in re3data?

2.2 Data export from re3data

2.2.1 How do you currently access repository descriptions in re3data?

2.2.2 How would you prefer to access repository descriptions in re3data (API, dashboards, website, subsets based on (which) criteria, spreadsheet export ...)?

2.3 Data import to re3data

2.3.1 What information could your organization add to repository descriptions in re3data?

2.3.2 How would you prefer to add information to the service?

2.4 Monitoring

2.4.1 Do you use re3data to monitor the repository landscape?

2.4.2 How could re3data improve to support your analysis?

2.5 Reference

2.5.1 Do you use re3data to persistently refer to repositories? If yes, how (e.g., repository name, re3data PID, etc.)?

2.5.2 How could the reference to repositories be improved?

2.5.3 How would you use a PID for repositories?

2.6 Recommendation

2.6.1 What kind of information (i.e., metadata fields) do you use to select recommended repositories?

2.6.2 Do you need predefined recommendations (e.g., based on the FAIR Data Principles, journal requirements, community standards)?

2.6.3 How would you like to export these recommendations from re3data?

2.7 Other use cases

2.7.1 Are there any other use cases or requirements we did not address so far? Do you have any other concluding remarks?

3. Prioritization

3.1 Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?

3.2 Which of the service gaps mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?

B. Presentation Slides

re3data coref

re3data Use Case Analysis: AGU

Shelley Stall, (https://orcid.org/0000-0003-2926-8353), 24 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

Karlsruher Institut für Technologie

DFG

Introduction

Senior Director,
Data Leadership for the American Geophysical Union

In consultation with the co-chairs of the recent ESIP Session: Supporting the Earth Science Community in Repository Discovery

Amber E Budden, PhD, Director of Learning and Outreach
National Center for Ecological Analysis and Synthesis

Daniella Lowenberg, Data Publishing Product Manager, University of California, Product Manager for Dryad at California Digital Library, and member of the Make Data Count team

Denise Hills, Director, Energy Investigations, Geological Survey of Alabama

re3data coref

AGU
ADVANCING EARTH AND SPACE SCIENCE

Information

re3data coref

1. Information for our authors and members on the best repository for their data to be preserved.
2. A way for authors to know which repositories provide the services that best meet the Enabling FAIR Data criteria.
3. Repositories that provide automated credit and attribution for preserved datasets.
4. Capture information about how paper peer reviewers can confidentiality access embargoed data.
5. Track terms of service information
    a. The conditions under which a dataset can be updated or removed,
    b. The expectation that the metadata will be present and maintained even when the data are purged/deleted,
    c. The conditions under which data are made open following an embargo.
6. Qualitative metadata to help with repository selection by researchers.
7. Easier way to compare domain, Institutional Repositories (IRs), and general repositories.
8. Current, managed, accurate information about repositories.

Data Export from re3data

Information on which journals commonly have data citations by repository.

Data Import to re3data

To the best of our knowledge, AGU has no information to provide.

Monitoring

We have not been able to use re3data to monitor repository landscape changes.

If this were possible, we would be interested in determining which repositories provide specialized curation for data types used within the Earth, space, and environmental sciences and which data types were not supported.

It would also be useful to know depending on a researcher's geographic location (or other limiting characteristics that repositories have) if they were disadvantaged.

For instance if you live in the US, you have access to one or more domain repositories for your research community, but if you live in Peru, your situation changes. This is important to understand equity and inclusion issues.

Reference

For AGU, we don't use the PID assigned to the repository widely. However, when we supported the development the most recent General Repository Comparison Chart, we did list the PID for each repository being compared - both re3data and FAIRSharing.org

We also created a FAIRSharing Collection to go with this chart.

Stall, Shelley, Martone, Maryann E., Chandramouliswaran, Ishwar, Crosas, Mercè, Federer, Lisa, Gautier, Julian, … Zigoni, Alberto. (2020, July 15). Generalist Repository Comparison Chart. Zenodo. http://doi.org/10.5281/zenodo.3946720

Recommendation

re3data coref

In addition to the information section…

- For repositories, the TRUST Principles should be included, the elements of the FAIR Data Principles that apply to repositories, and the elements of the CARE Principles for Indigenous Data Governance that apply to repositories.
- Implementation of MakeDataCount, and the relevant metadata concerning average citations per record, or other usage metric.
- A significant issue at AGU is the language of the repository. We require that our papers be in English and that content supporting the paper also be in English. We are working hard with repositories that are not in English to help our peer reviewers. This isn't always possible and a challenge for all involved. Some editors insist on English language repositories only. The repository should indicate the supported languages.
- It would be helpful to our authors to search by country where they live/work to determine which repositories support them.
- Any fees should also be made clear and part of the search criteria.

Other Use Cases

None

re3data coref

Prioritization of Use Cases

re3data coref

Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1. Researchers in search of a repository that is the best place for their data, needs to be a priority. All of the recommendations and information have this premise at the core.
2.
3.
4.
5.

Prioritization of Service Gaps

re3data coref

Which of the service gaps mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1. The current re3data tool is not intuitive to a researcher. The filtering and facets need to be considered from the point of view of a researcher who does not know the services of a repository or why they are important.
2. The information about each repository needs to be current and accurate.
3.
4.
5.

Thank You!

Tell us your use case story:

info@re3data.org

www.re3data.org

@re3data

doi.org/10.17616/R3D

re3data coref

re3data Use Case Analysis: Canada

Kelly Stathis (https://orcid.org/0000-0001-6133-4045), Discovery and Metadata Coordinator, Portage Network
Peter Webster (https://orcid.org/0000-0001-6801-0852), Associate University Librarian, St. Mary's University

24 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

DFG

Introduction

re3data coref

Portage Network

- Initiative of the Canadian Association of Research Libraries
- Dedicated to the shared stewardship of research data in Canada through:
  - Developing a national research data culture
  - Fostering a community of practice for research data
  - Building national research data services and infrastructure
- Network of Experts
  - Expert/Working Groups
- Infrastructure platforms
  - Includes the Federated Research Data Repository (FRDR)

CARL ABRC
CANADIAN ASSOCIATION OF RESEARCH LIBRARIES
ASSOCIATION DES BIBLIOTHÈQUES DE RECHERCHE DU CANADA

portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

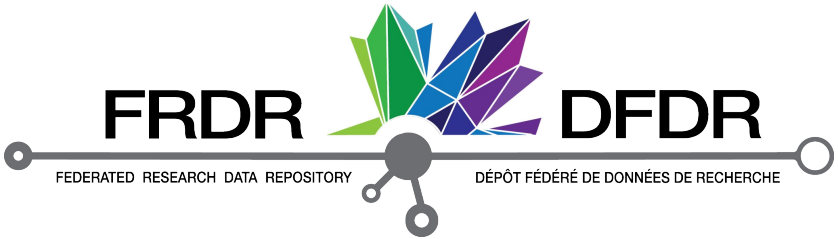
Introduction

re3data coref

Federated Research Data Repository

- “FRDR”
- Collaboration between Portage and Compute Canada
- Includes national discovery layer
- Harvests over 70 Canadian research data repositories
- Helps to:
  - Improve discovery of Canadian research (meta)data
  - Break down repository siloes
  - Drive traffic to existing repository sites
  - Create interoperability between platforms

portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA



compute canada | calcul canada

Information

re3data coref

What information do we need about research data repositories?

Answering the question, *Can we harvest this repository for inclusion in FRDR?*

- ✔️ Basic and contextual information
  - repositoryName, repositoryURL, repositoryContact, subject, etc.
- ✔️ Eligibility: Is the data repository Canadian?
  - institutionCountry
- ❓ Technical: Can we harvest the repository?
  - What platform does the repository use?
  - Does the repository have an API?
  - What metadata standard do they use?
- ❓ Content: Does the repository include metadata at the dataset level?
  - Or is it more of a “database”?

Information

re3data coref

What information is missing?

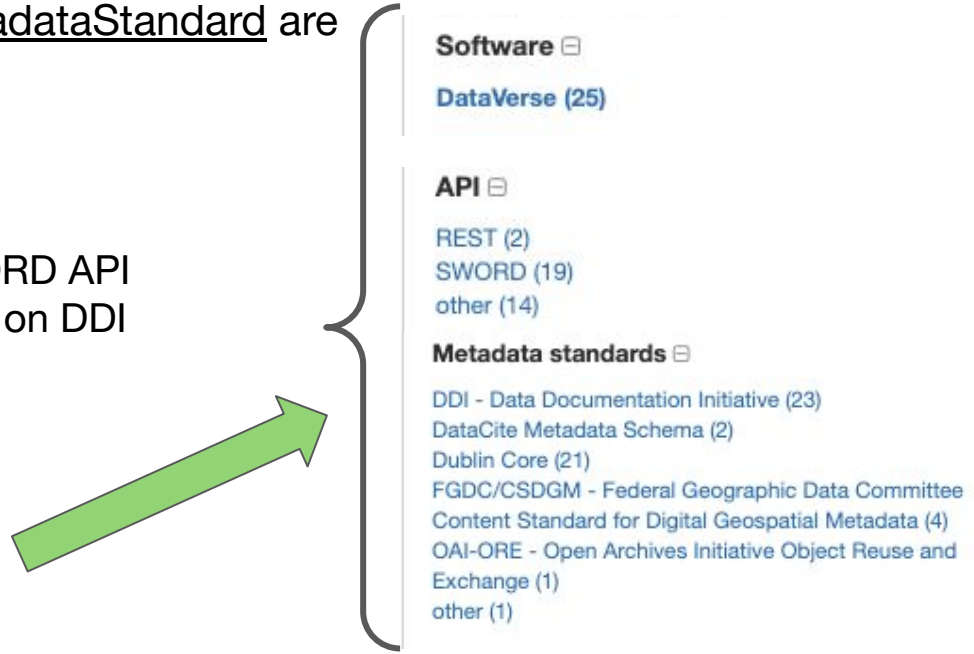
❓ Technical: Can we harvest the repository?

The fields software, apiType, and metadataStandard are intertwined.

Example: Dataverse repositories

- Software: Dataverse
- APIs: same set of REST APIs, SWORD API
- Metadata standard: schema based on DDI

How this is reflected in re3data:



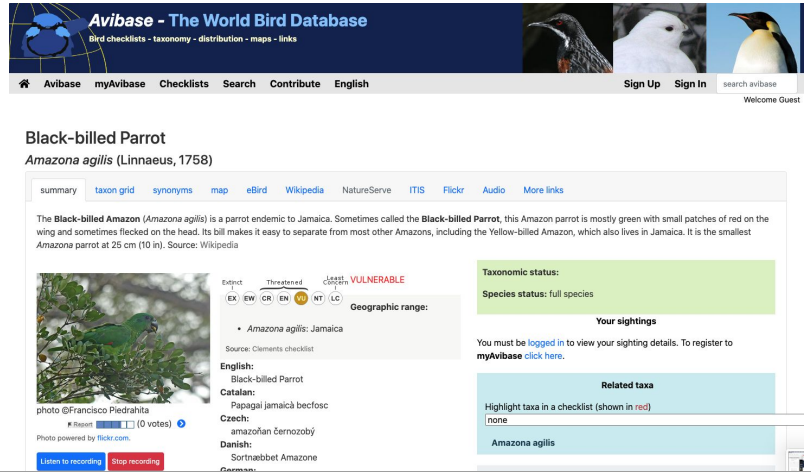
Information

re3data coref

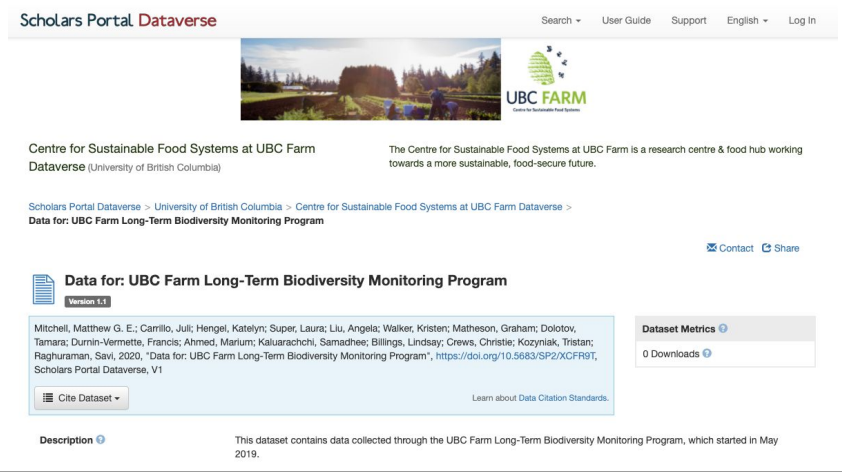
What information is missing?

❓ Content: Does the repository include metadata at the dataset level?

database: stores individual records, may not have any data files



data repository: hosts datasets consisting of files and metadata

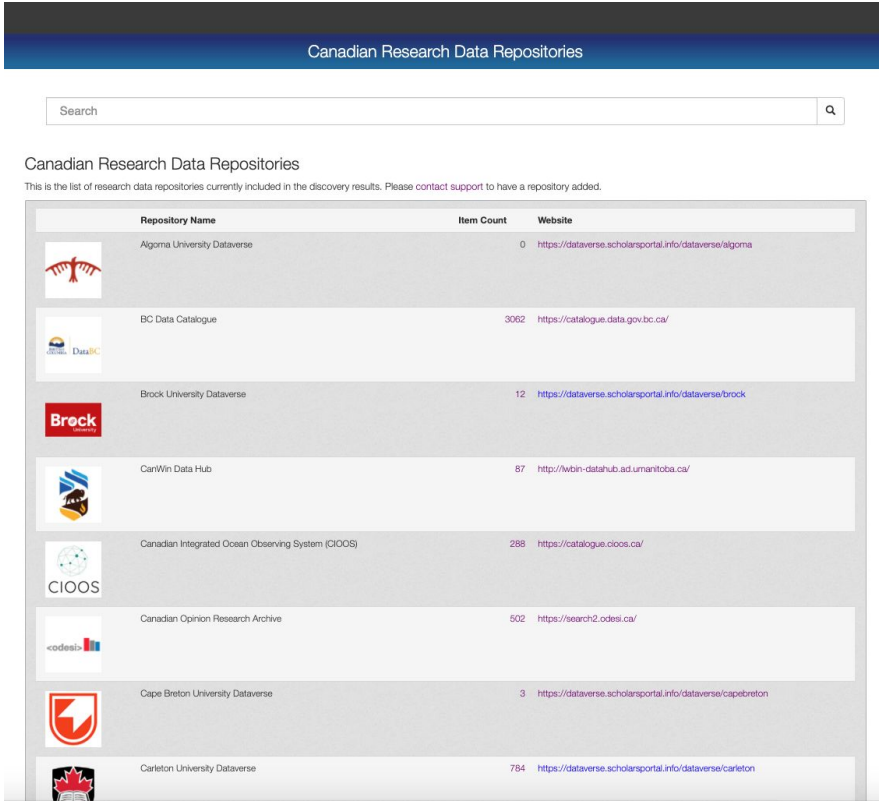


Reference

re3data coref

Do you use re3data to persistently refer to repositories?

- When evaluating repositories for harvesting, useful to have a unique ID for tracking
- In the future, we may add re3data PIDs to FRDR's repository list



Data Export from re3data

re3data coref

How do you currently access repository descriptions in re3data?

- API:
  - Export full list using query by country
- Website:
  - Used for reviewing individual repositories in the list

Data Import to re3data

re3data coref

What valuable information could your organization add to repository descriptions in re3data?

- Adding API and platform details
- Indicating repositories that are now offline
- Submitting new repositories

How would you prefer to add information to the service (e.g., via API, metadata upload, etc.)?

- Web interface to directly edit/suggest edits (like Wikipedia)
- API or metadata upload

Monitoring

re3data coref

How do you use re3data to monitor the repository landscape?

- Because Canadian institutions have worked to create an inventory of repositories in re3data, FRDR has been able to use re3data to identify new-to-us repositories for harvesting
- FRDR has also had access to metadata on repositories using data formats that can be easily harvested

How could re3data improve to support your analysis?

1. Completeness and currency
2. Distinguishing between data repositories and databases
3. Identifying relationships between repositories
4. Tracking the data coverage of repositories by country

Recommendation

re3data coref

What kind of information do you use to select recommended repositories?

- The Portage Network does not provide repository recommendations
- **Canadian librarians** may use re3data to monitor the landscape → informs their recommendations

Do you need predefined recommendations?

- Standards-based recommendations (e.g. based on FAIR principles) would benefit the Canadian RDM community

Other Use Cases

re3data coref

National Data Source Inventory / Reference Source

- Canadian agencies worked to collect/consolidate a relatively complete list of Canadian data repositories in re3data
- A sustained collection process was needed to create a reliable inventory - in preference to the random process of data repositories self-submitting
- With a reliable national inventory, re3data is being used by Canadian librarians as a reference source for locating available data providers
- It is also used by data creators as a reference for locating discipline specific data sources, and locations for data deposit

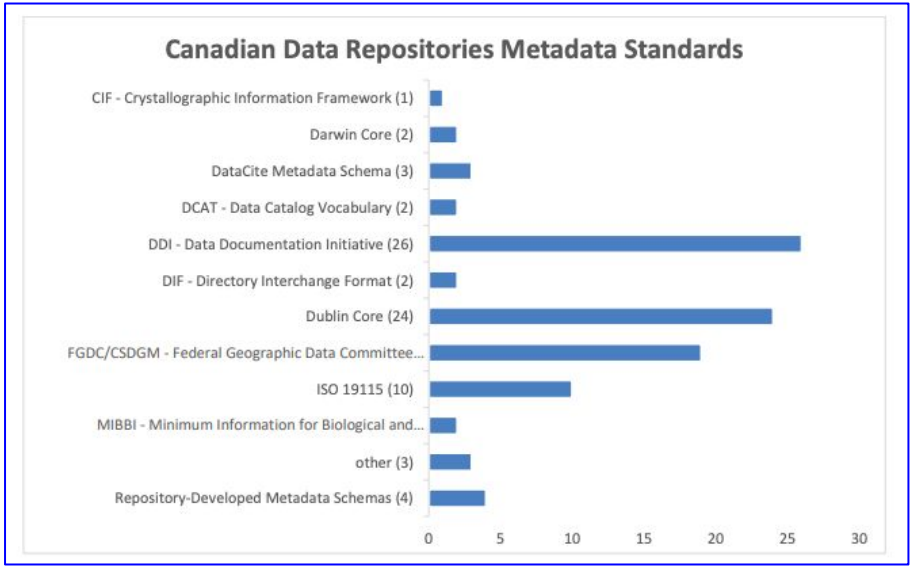
Other Use Cases

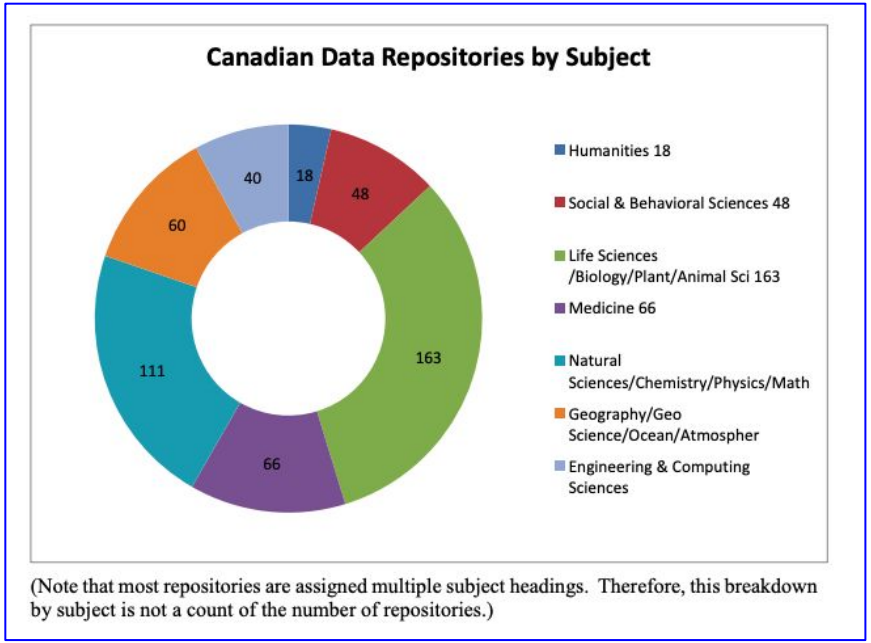
re3data coref

National Data Repository Landscape Analysis

- Once re3data contained a reliable national inventory, it became a source for analysis of the Canadian data repository landscape
- It provides statistics on:
  - the number of data repositories in different subject disciplines
  - metadata schemas being used to describe data
  - the data repository software and methods
- This information is invaluable in planning, promoting and funding the future development of research data infrastructure in Canada

Repository Landscape Analysis Examples





(Note that most repositories are assigned multiple subject headings. Therefore, this breakdown by subject is not a count of the number of repositories.)

Prioritization of Use Cases

re3data coref

Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?	
1. Monitoring (Inventory Collection)	Identifying repositories
2. Information	Evaluating repositories
3. Data import to re3data	Sharing our findings
4. Data export from re3data	Getting a list to evaluate
5. Recommendation	Recommending repositories to researchers

Prioritization of Service Gaps

re3data coref

Which of the service gaps mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?	
1. Enhanced, wiki-like interface for data import to re3data	Import
2. Distinguishing between data repositories and databases	Information
3. Making consistent distinctions between platforms, APIs and metadata standards	Information
4. Metadata to identify relationships between repositories	Information
5. Systems to assist Canadian data agencies in keeping repository entries up to date (e.g. checking if a repositories are active)	Monitoring

Thank you!

NDRIO | NOIRN

New Digital Research Infrastructure Organization

Nouvelle organisation d'infrastructure de recherche numérique

Funding in support of the Portage Network's stewardship of research data within Canada is administered through the New Digital Research Infrastructure Organization (NDRIO)

Contact:
Kelly Stathis, kelly.stathis@carl-abrc.ca
Peter Webster, peter.webster@smu.ca

Tell us your use case story:

info@re3data.org

www.re3data.org

@re3data

doi.org/10.17616/R3D

re3data

CoreTrustSeal

re3data Integration Recommendations

Wim Hugo on behalf of CoreTrustSeal

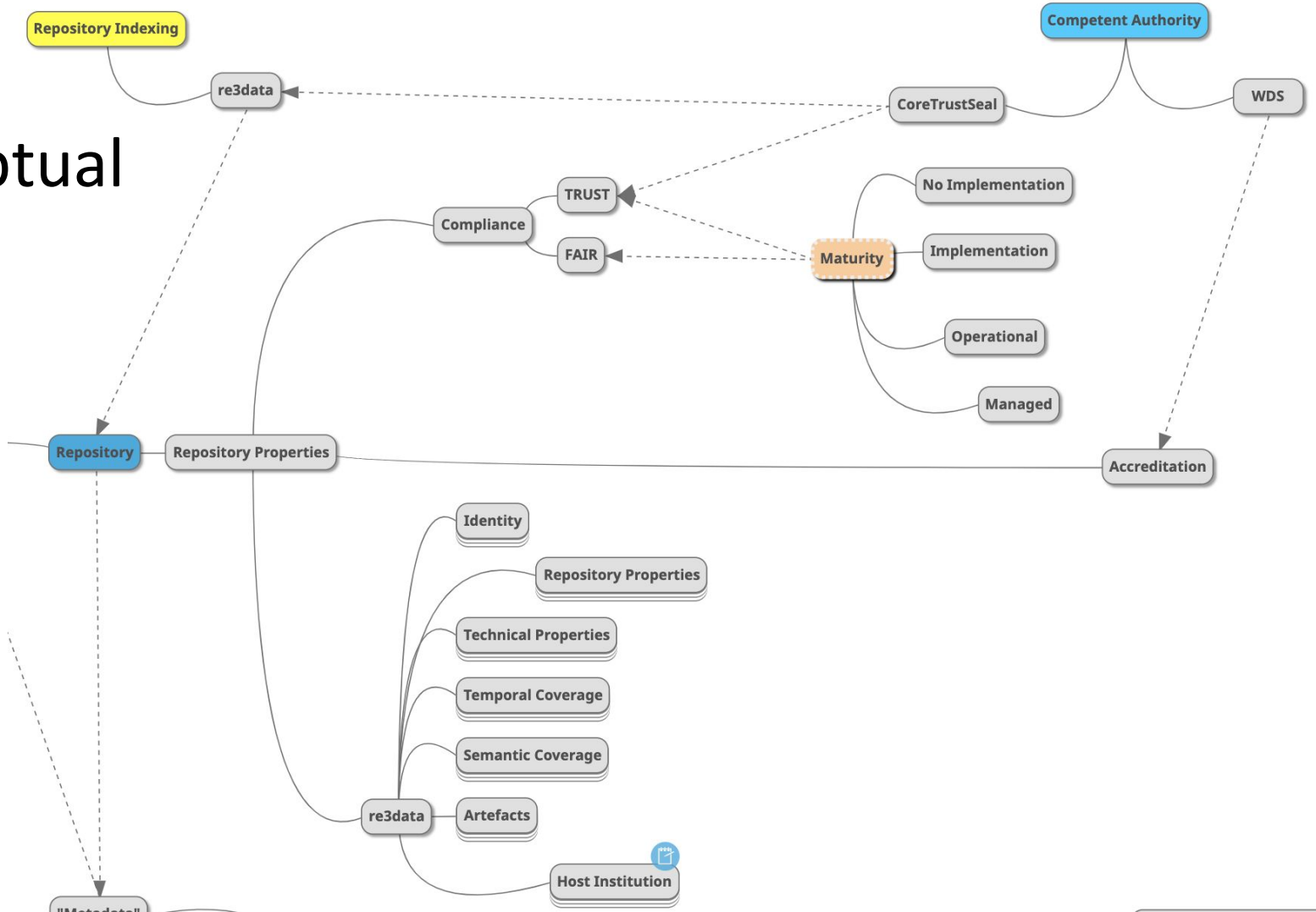
Broad Scope of CoreTrustSeal Recommendations

- The ability to provide persistent identifiers for (trustworthy) data repositories;
- The ability to highlight/maintain accreditation and certification status in re3data;
- The ability to provide a search facility that would list trustworthy repositories, and
  - direct the user to
    - Information about certification and/ or accreditation;
      - CoreTrustSeal and/or WDS;
    - The website of the repository;
  - be available as a machine readable list/ API output.

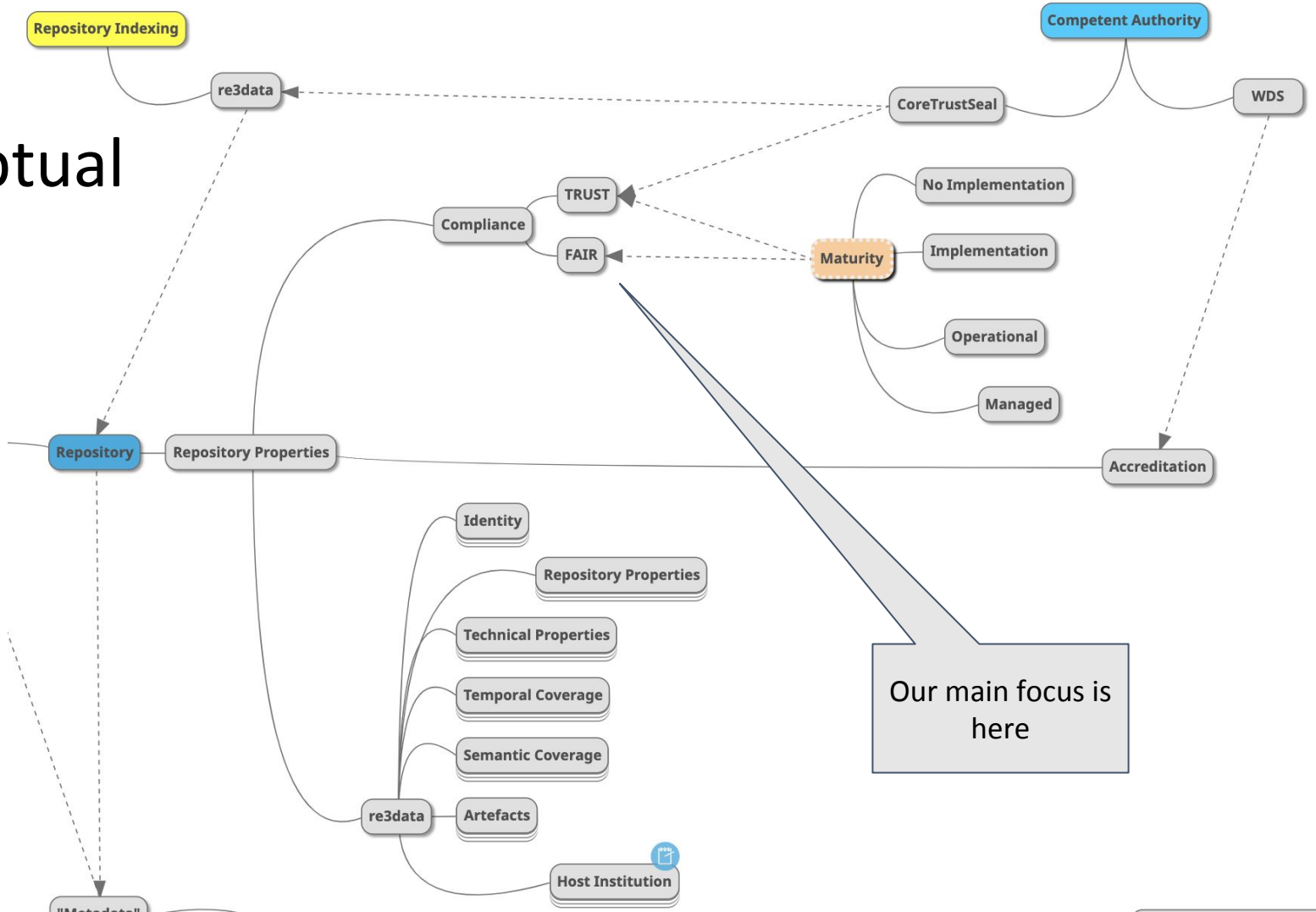
- Alignment with certification properties and vocabulary
- The ability to update repository information in the re3data registry should information about the repository change at any given point

Allow new applicants to populate CTS Profile from re3data

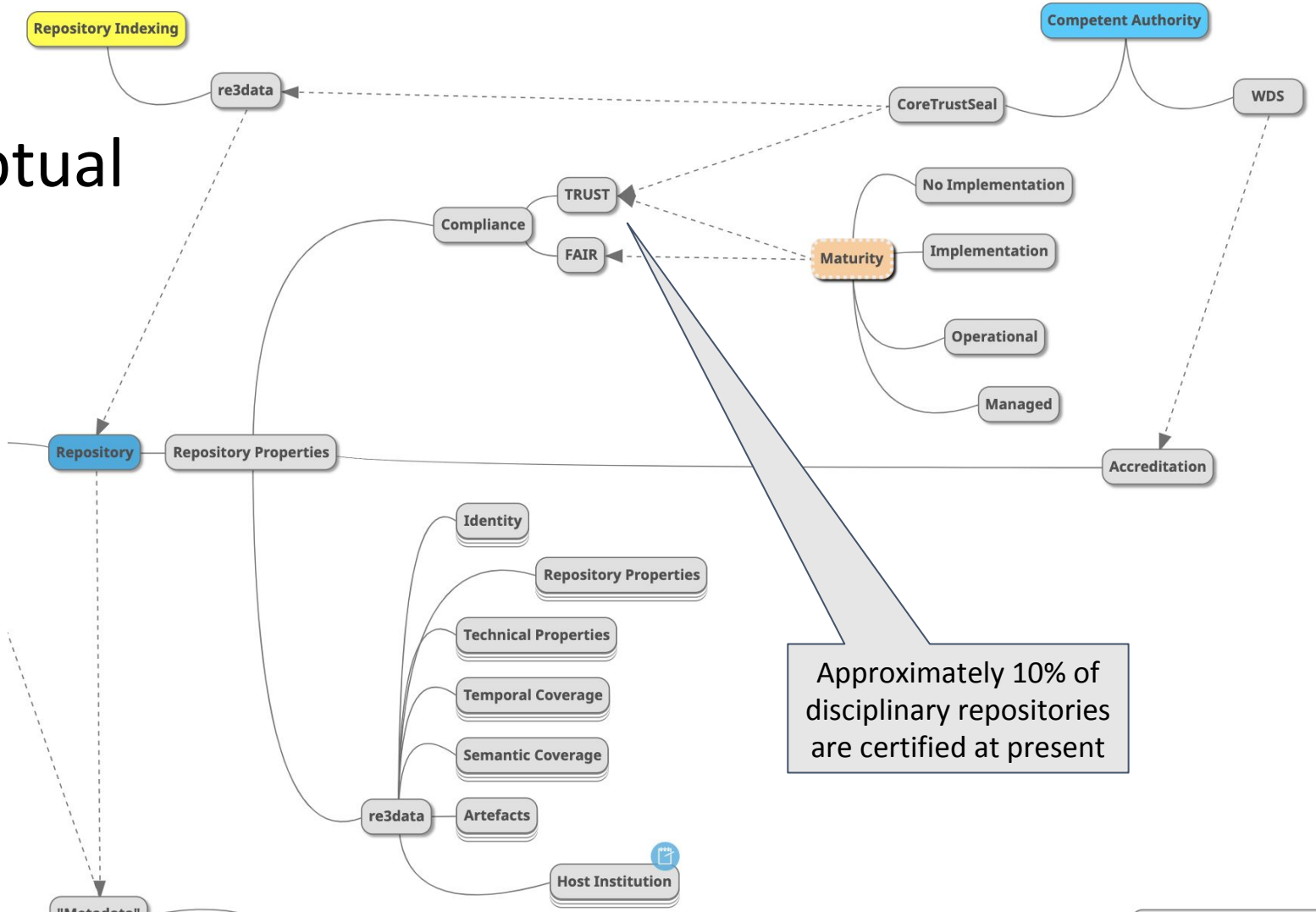
Conceptual Model



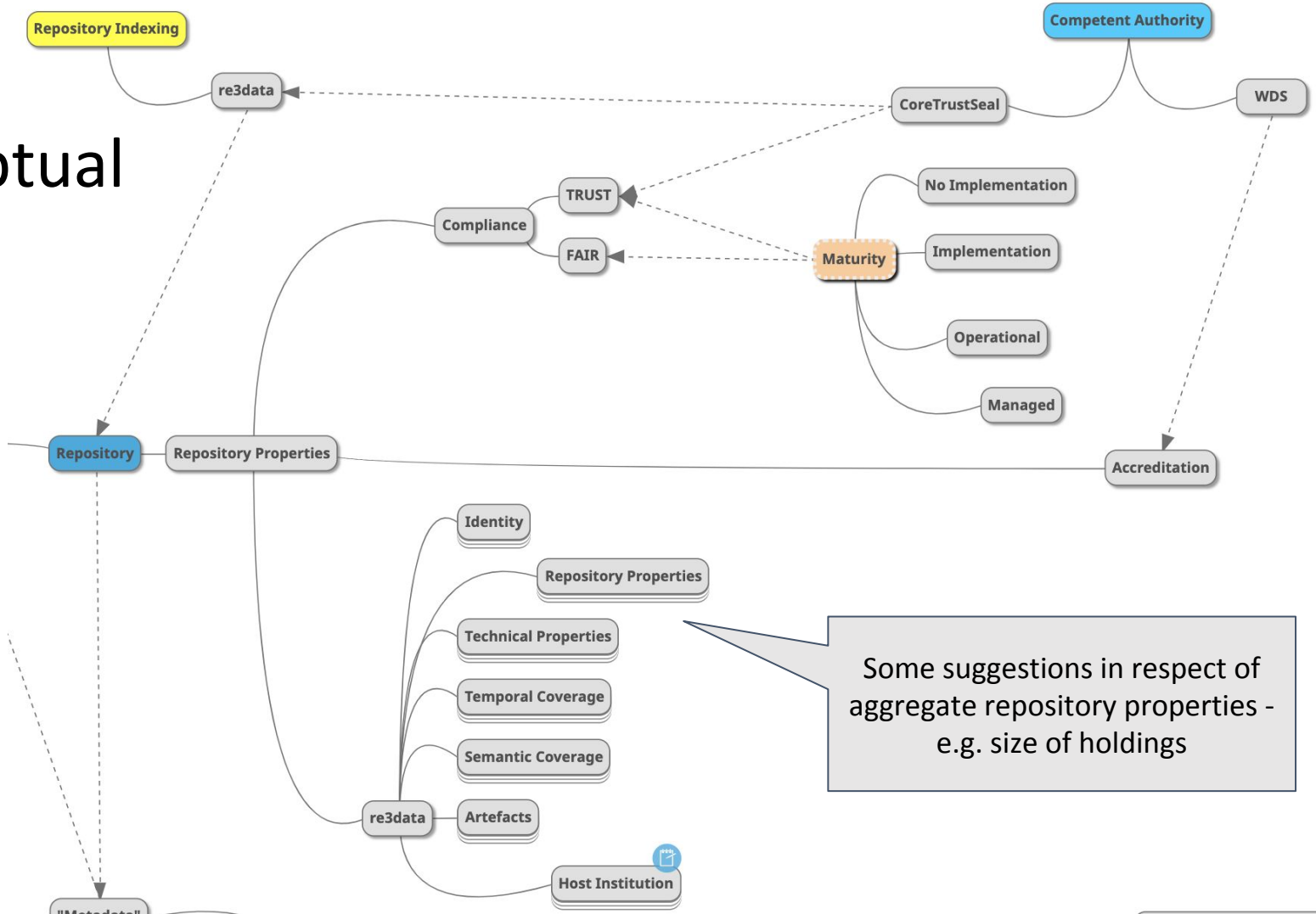
Conceptual Model



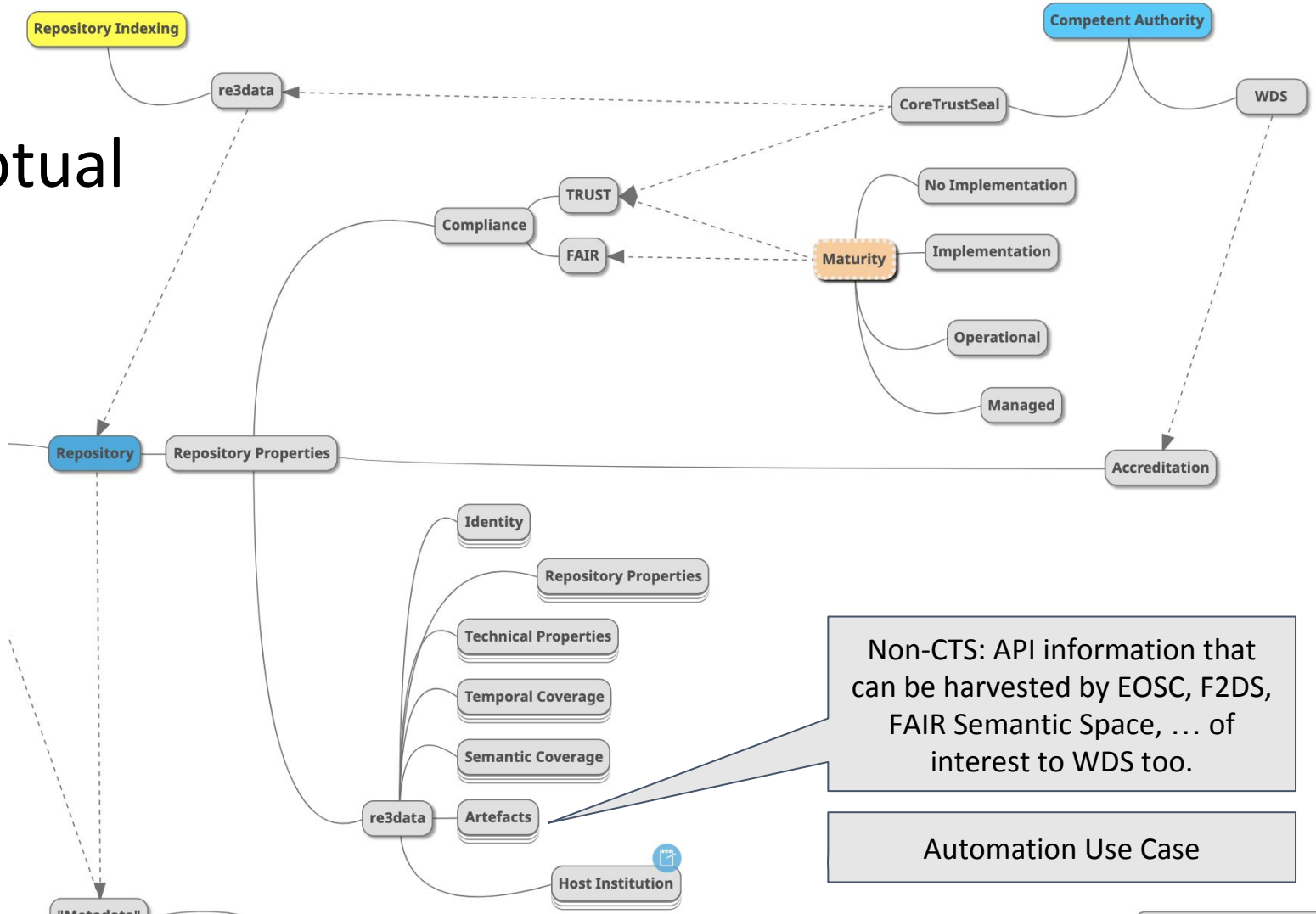
Conceptual Model



Conceptual Model



Conceptual Model



Vocabularies and Properties

CoreTrustSeal Contextual Information

- Repository Type:
  - Domain or subject-based repository; Institutional repository; National repository system, including governmental; Publication repository; Library; Museum; Archive; Research project repository; Other
- Brief Description of Repository
- Brief Description of the Designated Community
- Level of Curation Performed,
  - A. Content distributed as deposited;
  - B. Basic curation – e.g., brief checking, addition of basic metadata or documentation;
  - C. Enhanced curation – e.g., conversion to new formats, enhancement of documentation;
  - D. Data-level curation – as in C above, but with additional editing of deposited data for accuracy
- Insource/Outsource Partners

Trustworthy Data Repositories Requirements (10.5281/zenodo.3638211)

Vocabularies and Properties (continued)

It would be valuable from our perspective for re3data to

1. Align with CoreTrustSeal options for repository type and curation level or to at least work with CoreTrustSeal to create a standard vocabulary for them.
2. Develop/ confirm a typology for discipline/ domain or adopt an existing one.
3. Include a clear statement about the level of storage/ curation/ preservation provided by the repository.
4. Include a clear statement about whether the repository holds sensitive data (including personal data).

Vocabularies and Properties (continued)

5. Whether a repository is certified and against what should be made more explicit, and should include the date of certification (***see details later on***).
6. What standards are being employed by the repository.
7. Add ORCID, IGSN, and ROR to PID Systems options.
8. Integrate property definitions into Search and Change Request interfaces.

   For example, enhancedPublication means 'The RDR offers the interlinking between publications and data.', but that is not obvious unless reading the schema documentation (***human-friendly labels for lists***)

Data Exchange: API

WDS (also representing CoreTrustSeal) has an an ongoing discussion to align the data held by re3data with certification or accreditation data held by the relevant authority (e.g., WDS or CoreTrustSeal) so that repository entries are automatically updated from CTS/ WDS data.

In particular, repositories registering in re3data can claim certification, but this ideally should be verified manually or automatically by the certification authority.

Both self-declaration and certification validated by a third party must be supported though, since it is not always possible to do the latter (c.f., ISO27001).

- Any methodology employed should be built on a standard model for third party validation of repository metadata,
- and it should be clear when information is self-declared versus that validated by a third party (stating who the third party is). This should apply to all types of certification, including any future FAIR, ISO, other certifications.

I***t is expected that CoreTrustSeal would not remain on the self-declaration list once a validation system is in place.***

Data Exchange (API)

CoreTrustSeal has the ability to provide
any of the fields in its Organizational Profiles and approved applications form,
as well as supply information on certification status,
effective time period (certification and renewal dates),
and a link to the public certification itself.

Data Exchange (API)

CoreTrustSeal has the ability to provide
any of the fields in its Organizational Profiles and approved applications form,
as well as supply information on certification status,
effective time period (certification and renewal dates),
and a link to the public certification itself.

We make a technical proposal in the next slides since the ***schema of certification data*** is not predictable in a generic use case

Data Exchange (API)

CoreTrustSeal has the ability to provide
any of the fields in its Organizational Profiles and approved applications form,
as well as supply information on certification status,
effective time period (certification and renewal dates),
and a link to the public certification itself.

CoreTrustSeal needs a technical solution for getting the repository CoreTrustSeal certification status into the re3data repository metadata, and would prefer that ***certification cannot be claimed publicly without verification*** from the authority.

CoreTrustSeal is also strongly connected to the European FAIRsFAIR project, which is considering an integrated CoreTrustSeal+FAIR certification and FAIR digital object assessment. Specifically, the project is exploring how to provide possible badging solutions to FAIR-enabling repositories and FAIR objects, and this requires a solution that can take into account a change in repository or object status.

‘Certificate’ Element

#	Property	Description	O/C	Vocabulary/ Values
34	certificate	The certificate, seal or standard the research data repository complies with. Can be more than one.	0-n	Controlled vocabulary: CLARIN-B? CoreTrustSeal DIN 31644 ... ISO 16363 ... other
34.1	certificateURL	URL to a widget or webpage providing detail about accreditation. If ID is blank, URL is displayed.	0-n	Query string parameters: id=r3d100000000
34.2	certificateWidget	Determines support for widgets (badges)	0-n	Value=true, false
34.3	certifiedUntil	The date at which certification expires	0-n	DateTime

Priority: Setting up a system where one can differentiate between self-declared metadata and metadata that have been validated by a third party, particularly CoreTrustSeal.

‘Certificate’ Element

#	Property	Description		
34	certificate	The certificate, seal or standard the researc repository complies with. Can be more tha		... ISO 16363 ... other
34.1	certificateURL	URL to a widget or webpage providing detail about accreditation. If ID is blank, URL is displayed.	0-n	Query string parameters: id=r3d100000000
34.2	certificateWidget	Determines support for widgets (badges)	0-n	Value=true, false
34.3	certifiedUntil	The date at which certification expires	0-n	DateTime

Using just one reference to an external plug-in (widget or badge) the responsibility for providing non-re3data properties, level of certification, etc. are deferred to the authority

Priority: Setting up a system where one can differentiate between self-declared metadata and metadata that have been validated by a third party, particularly CoreTrustSeal.

Widgets (Badges)

#	Property	Description	O/C	Vocabulary/ Values
34	certificate	The certificate, seal or standard the research data repository complies with. Can be more than one.	0-n	Controlled vocabulary: CLARIN-B ? CoreTrustSeal

Outcomes	id=blank	Id= r3d100000000
certificateWidget=false, blank	URL=Authority Webpage	URL=Member/ Repo Webpage
certificateWidget =true	URL=Authority Widget	URL=Member/ Repo Widget

		blank, URL is displayed		
34.2	certificateWidget	Determines support for widgets	0-n	Value=true, false
34.3	certifiedUntil	The date at which certification expires	0-n	DateTime

Monitoring and Metrics

CTS predominantly uses re3data to

1. Discover data repositories that serve a particular discipline.
2. Check the listed certifications:
    a. Which ones, how many types, the number of repositories with certifications, whether those listed as being WDS or CoreTrustSeal accredited/ certified are correct, whether any WDS or CoreTrustSeal certifications are missing, and so on.

Improvements include:

- Refinement of the Metrics section to be able to show metrics by certification type. For example, when selecting the Database Licenses metric, have colours to indicate the certification type. Or be able to filter and only get metrics for a certain certification type/. This would enable WDS/ CoreTrustSeal to contrast certified repositories with non-certified ones more easily.
- Similarly, it would help to be able to filter metrics results by a specific country or discipline.
- A ‘nice’ option would be to have a separate dashboard to show certified repositories and their disciplines.

The API already allows most of this on a technical level in the beta implementation.

Filterable API Responses

/api/beta/repositories

A GET-Request will return a an XML-Document following the HATEOAS-Principles

Example:

```
<list>
 <repository>
  <id>r3d100000000</id>
  <name>Cat-alog - Meow, meow, meow</name>
  <link href="/api/beta/repository/r3d100000000" rel="self"/>

  <certificate>CTS</certificate>
 </repository>
</list>
```

CTS confirms the need for extension of the API to allow filtering on the ‘certificate’ element, using one of the permissible values - add to V1

Additional API Uses

WDS and CoreTrustSeal do not use re3data to select recommended repositories. However, the Subject(s) field enables Board members to select repositories in their personal capacities.

For predefined recommendations, we believe certifications according to the TRUST Principles and FAIR Principles should be available as filters, and as they become available, any certifications based on the CARE Principles. More specifically, we would want to see recommendations for all hybrid CoreTrustSeal certifications that are developed in the future (i.e.,CoreTrustSeal + ‘X’), with the CoreTrustSeal+FAIR as the starting point.

The ability to again export via an API would be an advantage to WDS and CoreTrustSeal.

Add an ‘Accreditation’ element

/api/beta/repositories

A GET-Request will return a an XML-Document following the HATEOAS-Principles

Example:

```
<list>
 <repository>
  <id>r3d100000000</id>
  <name>Cat-alog - Meow, meow, meow</name>
  <link href="/api/beta/repository/r3d100000000" rel="self"/>

  <accreditation>WDS</accreditation>
 </repository>
</list>
```

WDS, CLARIN, … - more like accreditations and not certifications

‘Accreditation’ Element

#	Property	Description	O/C	Vocabulary/ Values
3x	accreditation	The organisation that the repository is accredited with. Can be more than one.	0-n	Controlled vocabulary: CLARIN? ... DRAMBORA ... RatSWD TRAC ... WDS other
3x.1	accreditationURL	URL to a widget or webpage providing detail about accreditation. If ID is blank, URL is displayed.	0-n	Query string parameters: id=r3d100000000
3x.2	accreditationWidget	Determines support for widgets	0-n	Value=true, false
3x.3	accreditedUntil	The date at which accreditation expires	0-n	DateTime

API Methods

- Create

  in cases where a member is accredited or a repository is certified and it is not known to re3data

  When requesting a CoreTrustSeal Application, a repository must create an Organisation Profile. As part of this Profile, we have a field asking for their re3data PID; or, if they do not have one, we encourage them to contact re3data to set up their entry and thus obtain a re3data PID

- Update

  update certification or accreditation details

Additional Use Cases

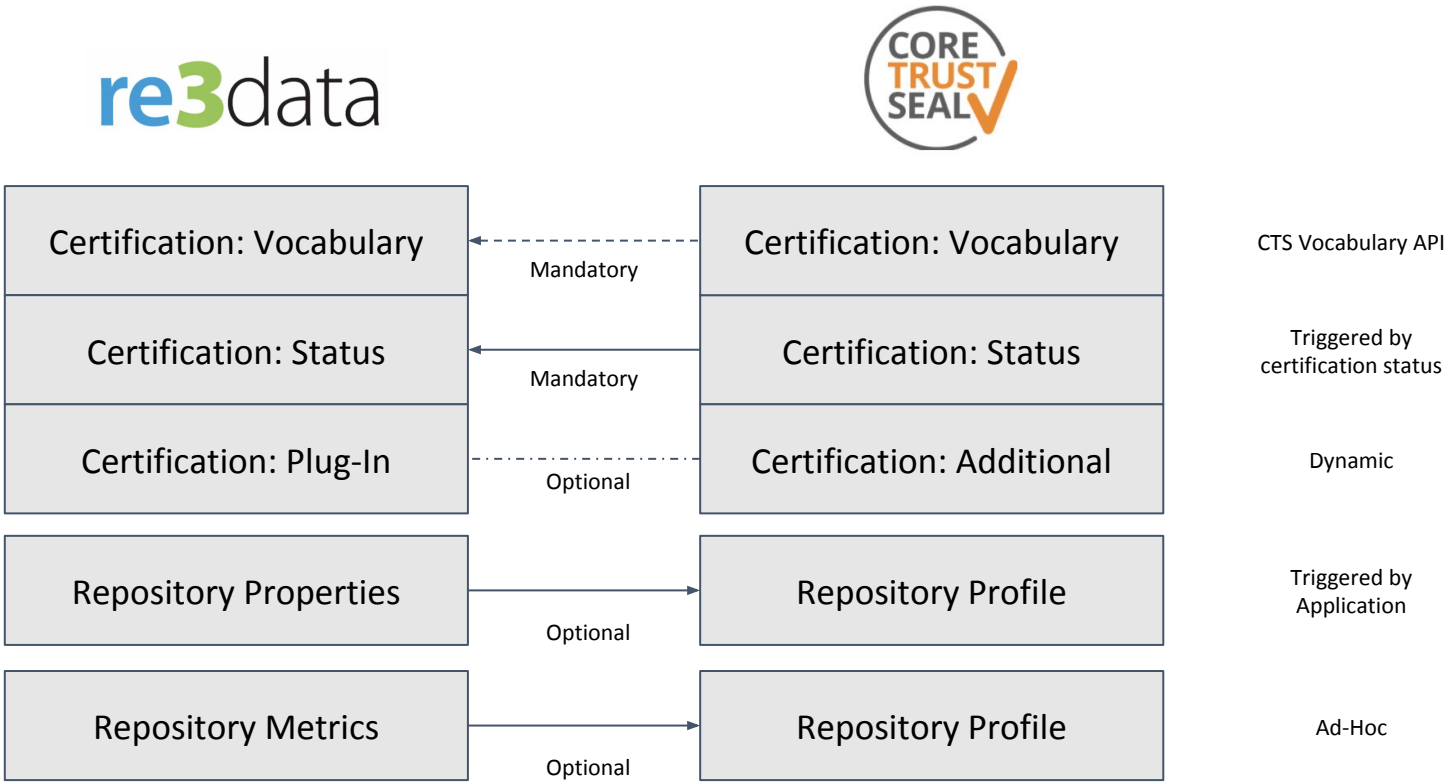
The following two use cases have not been mentioned and are of increasing importance:

1. Endorsement by funders as a repository to be used for grant-funded research output deposit and for inclusion into DMPs.
2. Sensitive data capabilities at a repository (e.g., indigenous data, personal data, etc.).

We have no other comments at this time. As you should be aware, and is briefly touched upon above, WDS is part of the DataCite re3data WG, and so we will be happy to elaborate on the points made above and continue the discussions through that forum.

Additional (non-CTS): Semantic Artefact Repositories, Protocol and Methodology Repositories, ….

Some Thoughts on Architecture



Thank you!

re3data coref

re3data Use Case Analysis: DBIS

Dr. Brigitte Doß, (https://orcid.org/0000-0002-9113-9424) & team, 26 November 2020

Dr. Gernot Deinzer, (https://orcid.org/0000-0002-7462-3847)

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

DFG

DBIS | Database Information System

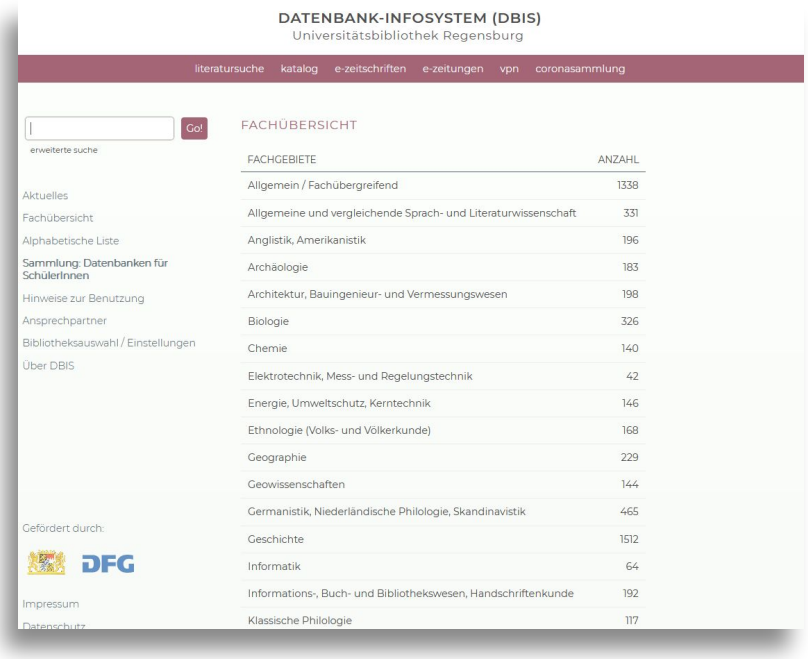
re3data coref

Large Index for scientific databases

13750 registered databases in total

5810 freely available databases

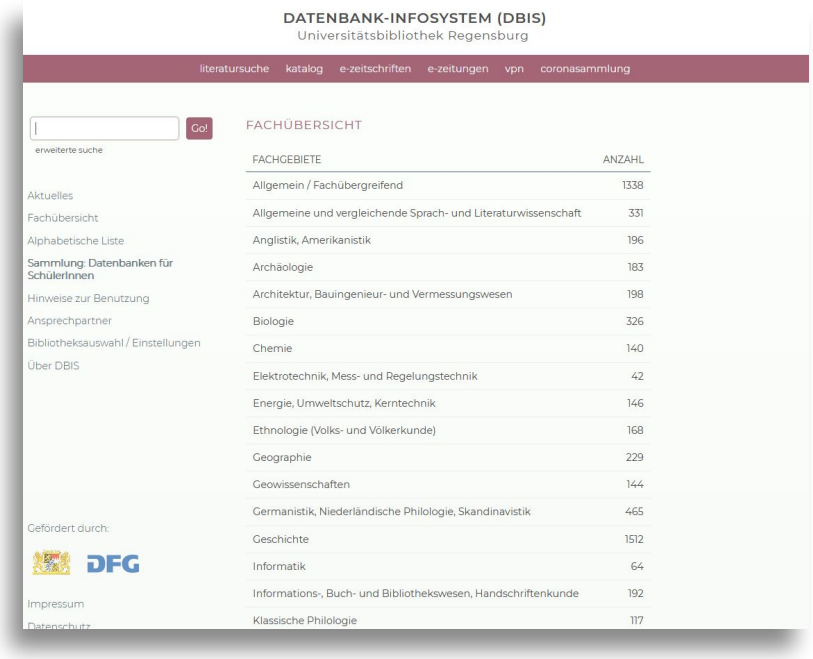
342 participating libraries



DBIS | Database Information System

re3data coref

Developed since **2002** at the **University Library of Regensburg** by editorial and IT teams



Information | Status Quo

re3data coref

1) **re3data** *is a* registered database in **DBIS**

2) Some data repositories from **re3data** are registered in **DBIS**

   - Redundant effort
   - Less than 10% of relevant data repositories from **re3data** are registered in **DBIS**

Information | Vision

re3data coref

Further development of **DBIS** within **DFG**-funded project

Planned integration of **re3data** repository metadata into **DBIS**

- (Alternative) Titles
- Topics
- URIs
- Languages etc.

Data Import & Export

re3data coref

re3data → DBIS

- Crawling the **Web-API** or using an **Export File**
- Including queries to **Web-API** into **DBIS** queries

DBIS → re3data

- **DBIS** database descriptions meet **high quality standards**
- **Enhanced API** as a main feature of **future development**

Monitoring

re3data coref

If using crawler for data import into **DBIS**:

- Monitor updates on registered repositories
- Trigger partial index updates on relevant items

Referencing

re3data coref

Hyperlink **DBIS** search results to **re3data** using PID

Link resolver from **re3data** PID to valid repository URL

Recommendation

re3data coref

According to admission criteria of **DBIS**

- **Repository type**: only disciplinary, not institutional
- **Repository quality**: meeting scientific standards

Thank You!

Stay tuned!

info.dbis@ur.de

 dbis.ur.de/projekt

 @ubreg

DARIAH-EU
Digital Research Infrastructure for the Arts and Humanities

re3data coref

The DDRS* as re3data use case

Stefan Buddenbohm (https://orcid.org/0000-0002-3469-6101)
Yoann Moranville (https://orcid.org/0000-0002-3502-1667)

Humanities at Scale has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 675570.

24 November 2020

Source: Buddenbohm, Stefan/Moranville, Yoann/De Jong, Maaike/Minel, Jean-Luc (in publishing): *Find Research Data Repositories for the Humanities - The Data Deposit Recommendation Service*. International Journal of Digital Humanities. Springer, Heidelberg.
Current preprint (in validation step, soon online): https://hal.archives-ouvertes.fr/hal-03020703v1

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

DFG

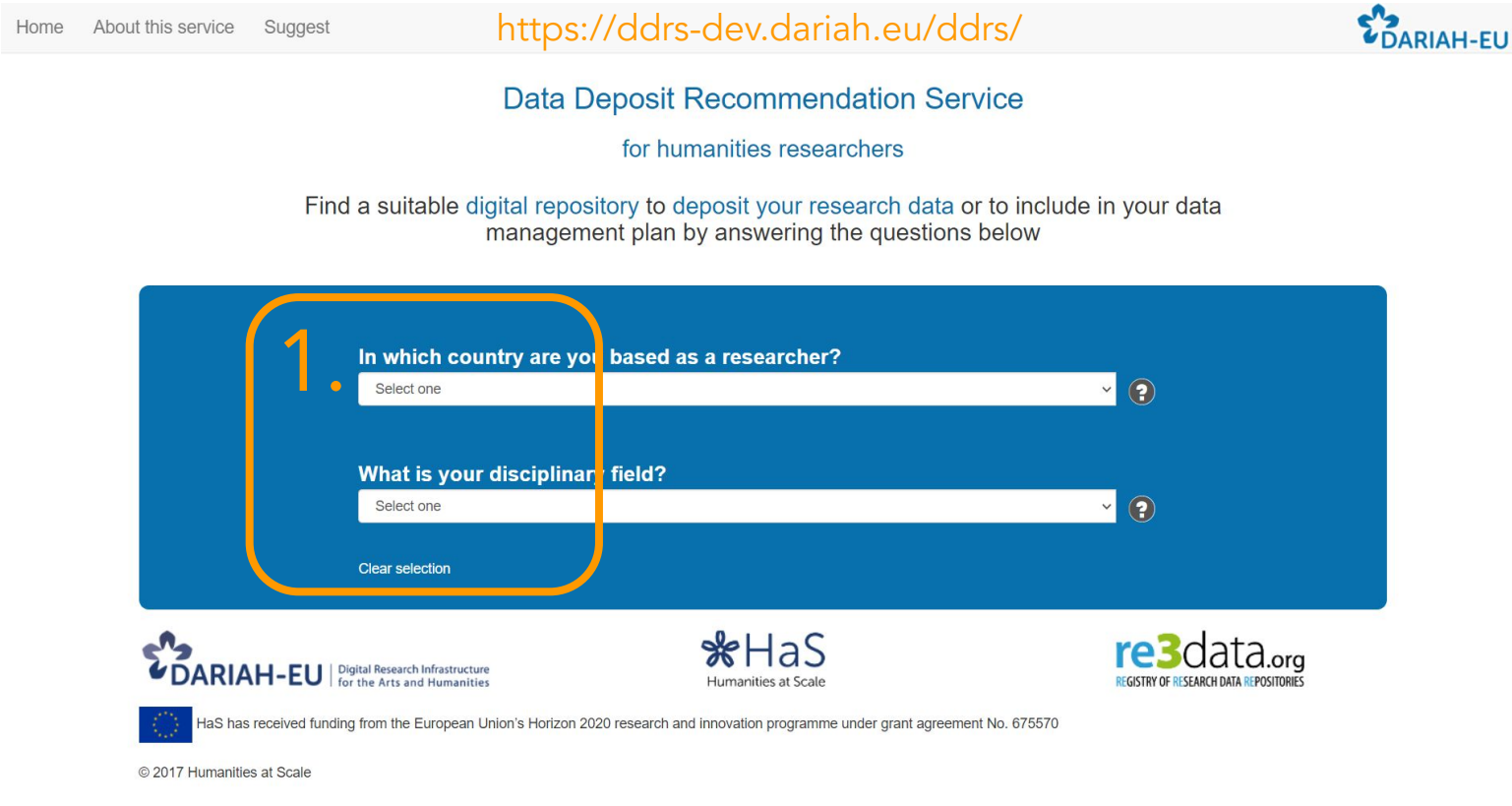
* Data Deposit Recommendation Service

re3data coref

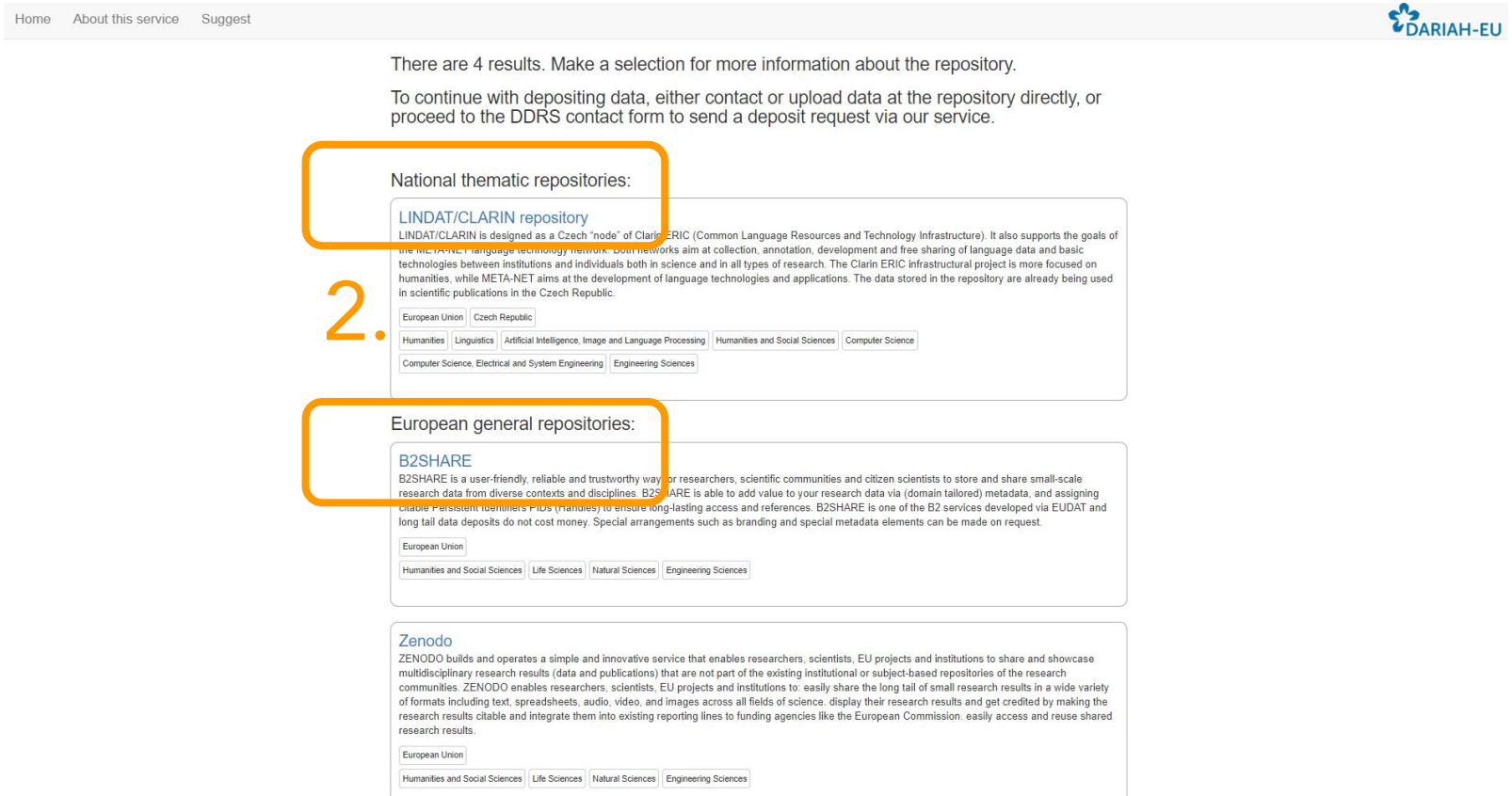
What problem solves the DDRS and for whom?

The DDRS **recommends research data repositories** to **humanities** researchers searching for deposit services for their research data, which **comply to criteria** such as PIDs, funders' requirements, disciplinary scope or language preferences.

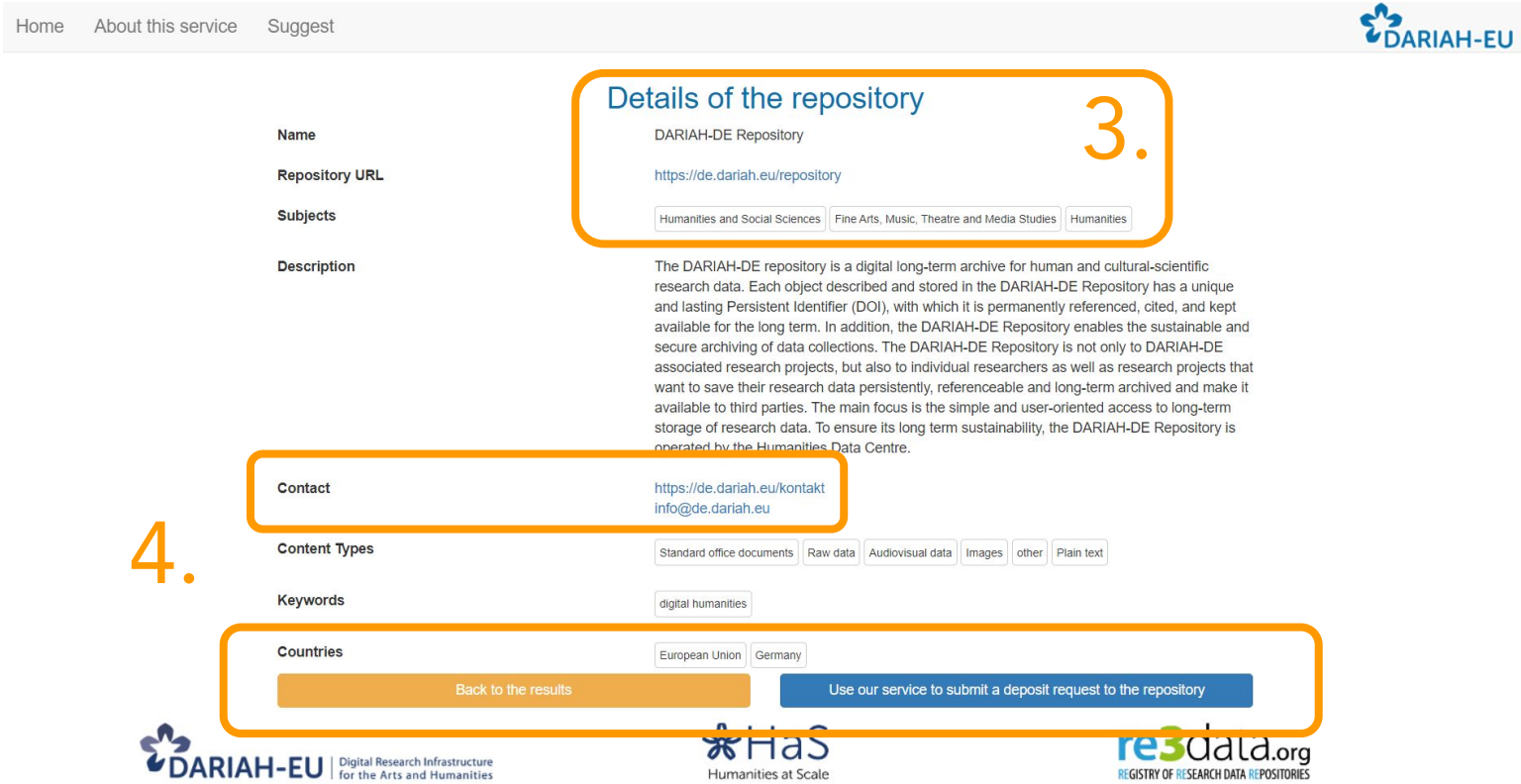
How does it look like?



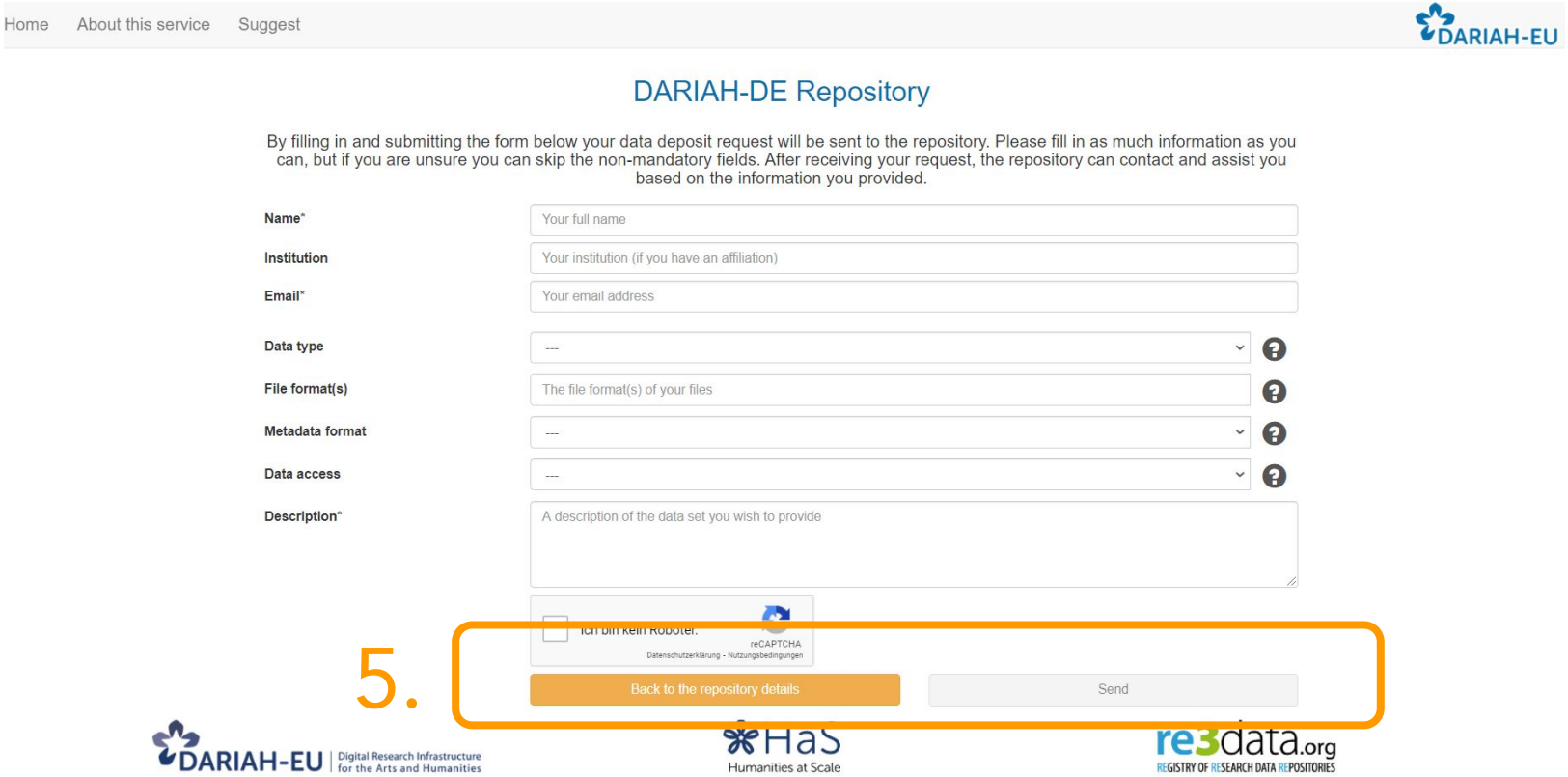
How does it look like?



How does it look like?



How does it look like?



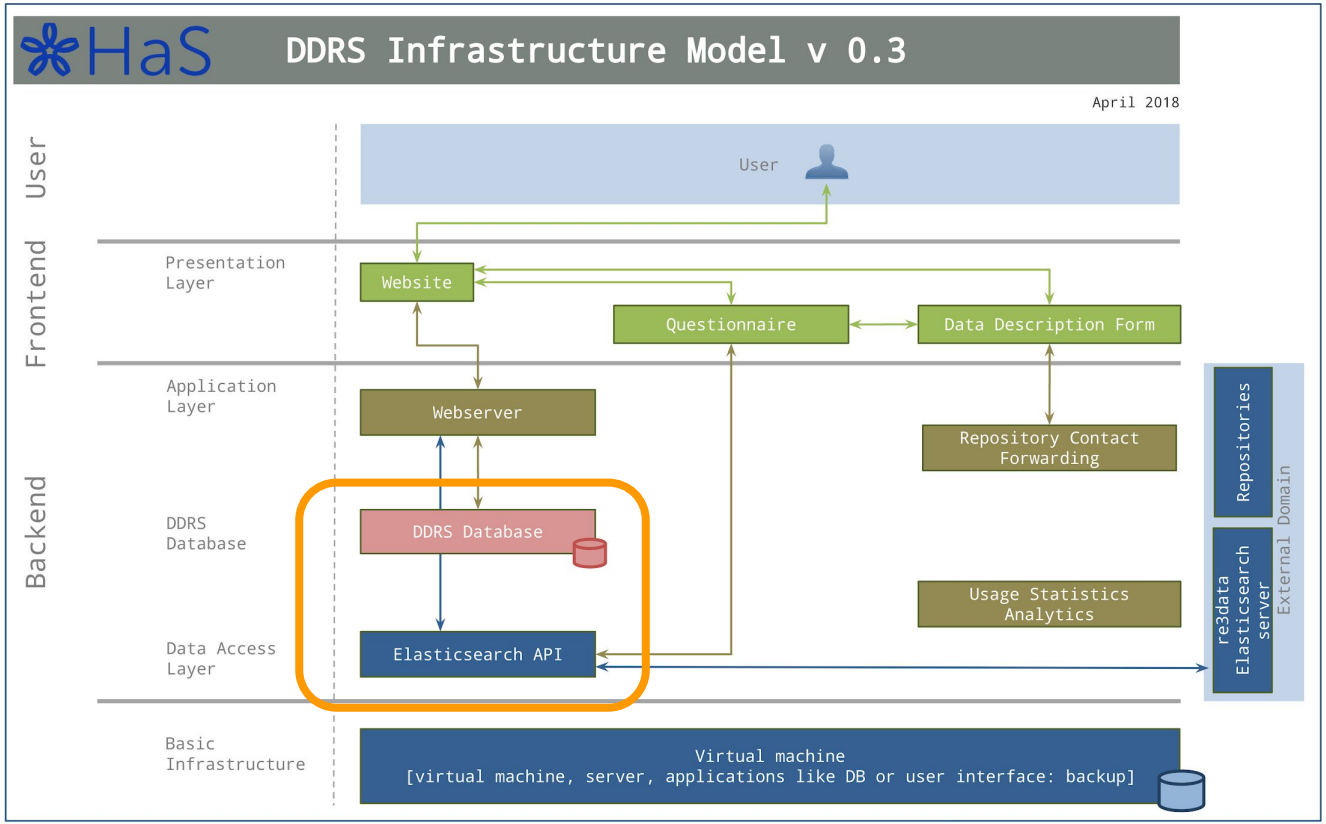
5.

DARIAH-EU | Digital Research Infrastructure for the Arts and Humanities

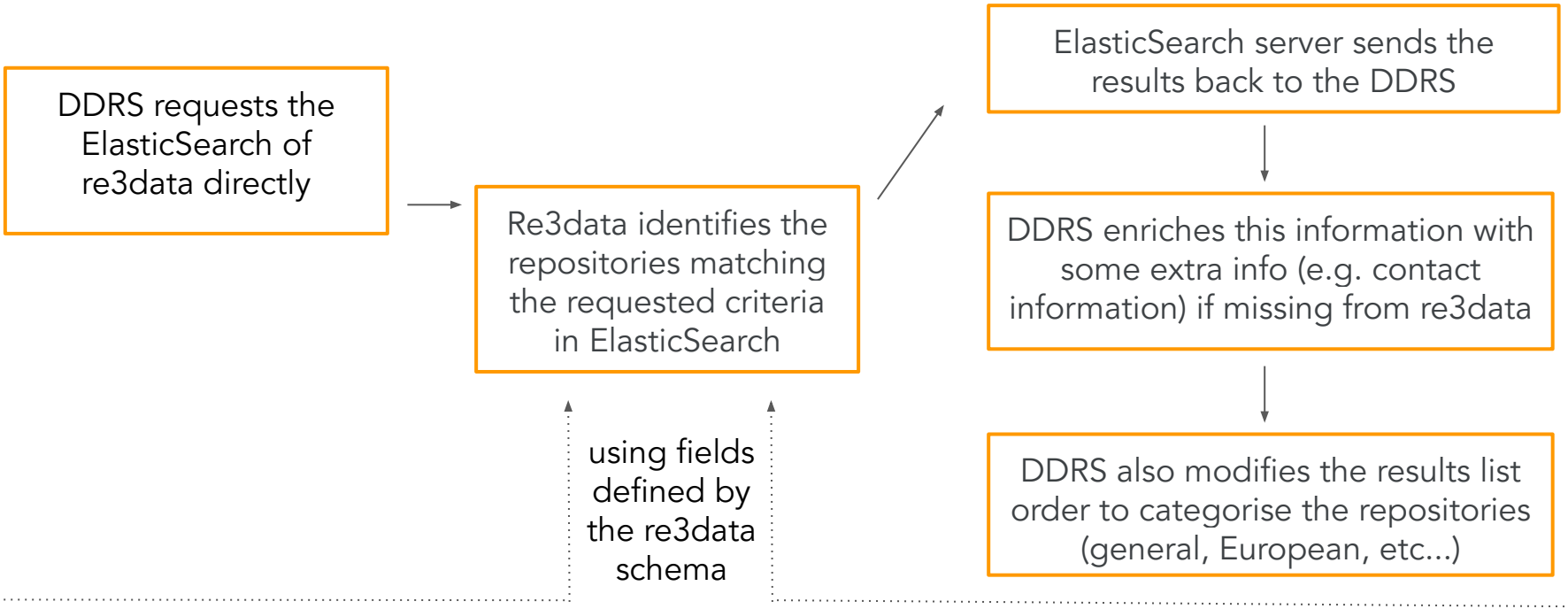
HaS Humanities at Scale

re3data.org REGISTRY OF RESEARCH DATA REPOSITORIES

How is it technically working?



Information retrieval



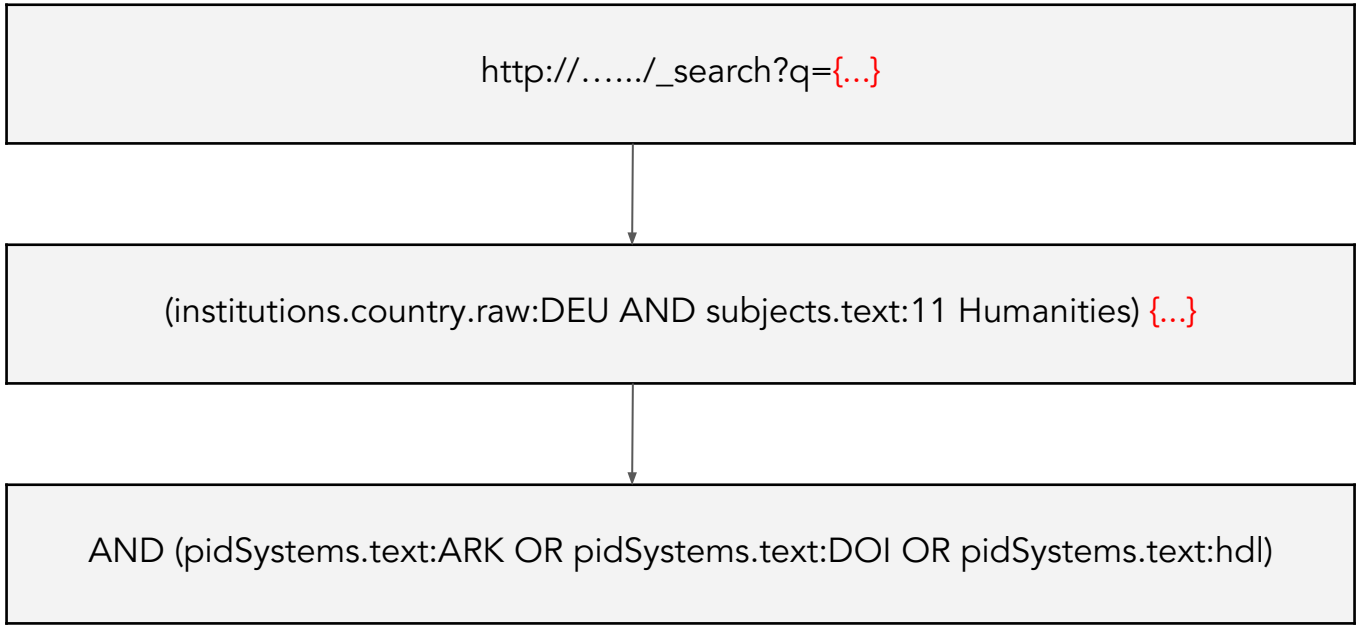
re3data schema 2.2 http://doi.org/10.2312/re3.006
e.g. ID 22.1 dataAccessType (open, embargoed, restricted, closed); ID 35.1 metadataStandardName; ID 23.1 dataLicenseName (for instance: CC0)

Information retrieval

DDRS requests the ElasticSearch of re3data directly

From the users' input: e.g. Germany and Humanities

Additional filters for refinement, e.g here for PIDs



http://……/_search?q=(institutions.country.raw:DEU AND subjects.text:11 Humanities) AND (pidSystems.text:ARK OR pidSystems.text:DOI OR pidSystems.text:hdl)

Problem: OR queries

re3data coref

Excerpt from our questionnaire:

"How do you currently access repository descriptions in re3data? How would you prefer to access repository descriptions in re3data (e.g., API, dashboards, website, subsets based on (which) criteria, spreadsheet export ...)?

- *Access directly via Elastic Search API. The reason being that the public API was not allowing us to make queries as we wished, especially "OR" queries (e.g. I want all repositories who provide "DOI" OR "Handle", in contrast to "DOI" AND "Handle")*
- *Maybe an API that could replicate the ES API would be best? Or a subset of it at least. So that we would not need to use a special API for our needs but directly the public API*
- *Regarding accessing the database as human and browsing through the registry: clean and useful user interface! Keep the UI as simple as possible, this is an asset of Re3data."*

re3data coref

Why re3data.org ?

REGISTRY OF RESEARCH DATA REPOSITORIES

- ➔ Available and geared for our purpose. Openness of the data and collegiality and the team. Useful and not just “politically or strategically interested”.
- ➔ Re3data’s resembles our own use case. > *find repositories*
- ➔ Granularity of their data and its metadata schema seems future-proof to us. *> re3data.org schema 3.0*
- ➔ Illustrates reuse and adoption of other resources. *> important wrt DARIAH*
- ➔ Sustainability. *> technically quite easy from DARIAH side*

re3data coref

Conclusions

- DDRS…
  - is a demonstrator focussing at humanities researchers and service providers.
  - shows how a quite simple use case from a user's perspective may be transformed into an easy-to-use web service.
  - it encourages the re-use of already existing resources.
- Re3data…
  - serves as enabler for the DDRS web service and coincidentally guarantees its actuality - and adaptability (e.g. for other disciplines).
  - demonstrates the potential for re-use of its database and metadata schema.
  - illustrates openness and collegiality with DDRS as example.

re3data coref

Thank you for listening to: The DDRS as re3data coref

Stefan Buddenbohm (https://orcid.org/0000-0002-3469-6101) sbudden@gwdg.de
Yoann Moranville (https://orcid.org/0000-0002-3502-1667) yoann.moranville@dariah.eu

24 November 2020

Source: Buddenbohm, Stefan/Moranville, Yoann/De Jong, Maaike/Minel, Jean-Luc (in publishing): *Find Research Data Repositories for the Humanities - The Data Deposit Recommendation Service*. International Journal of Digital Humanities. Springer, Heidelberg.
Current preprint (in validation step, soon online): https://hal.archives-ouvertes.fr/hal-03020703v1

Humanities at Scale has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 675570.

re3data coref

re3data Use Case Analysis: CLARIN

Dieter Van Uytvanck, 26 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

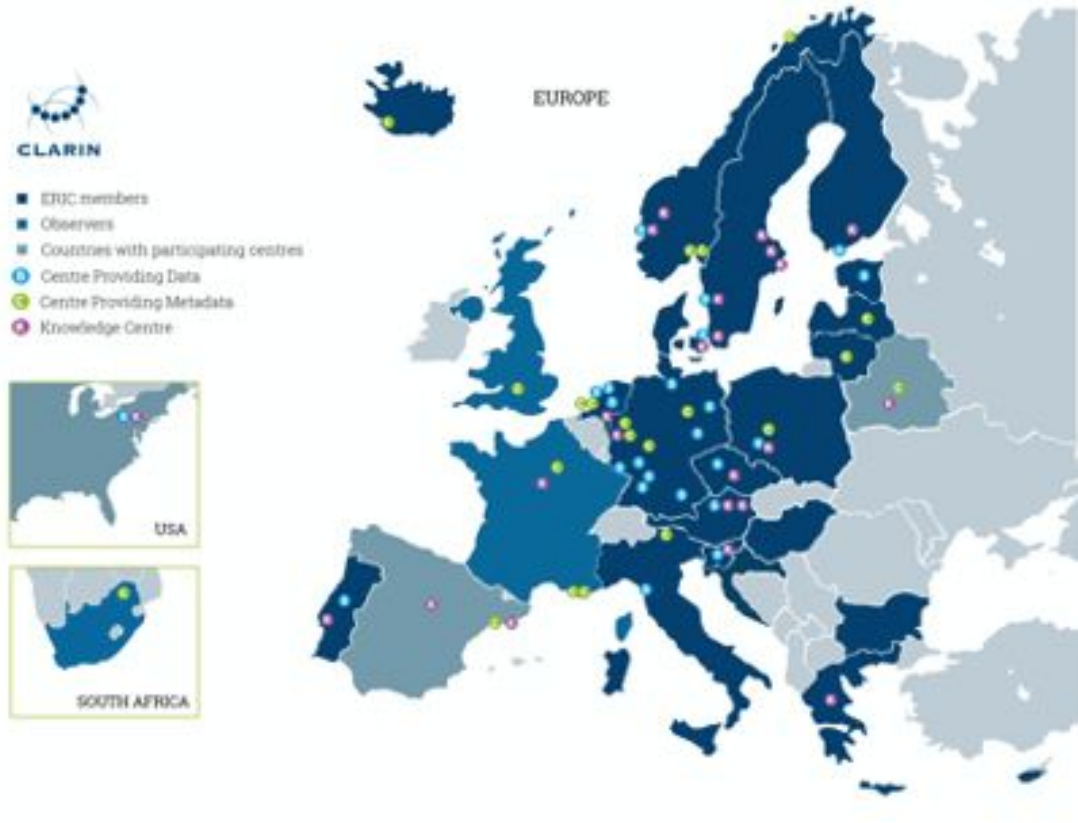
DFG

Introduction

We represent a European Research Infrastructure Consortium (ERIC): CLARIN, the Common Language Resources and Technology Infrastructure with 21 member countries and 3 observers.

As a network in Europe we represent data repositories and services providers throughout Europe and some partners outside of Europe, representing key stakeholders providing data and services for scholars in the humanities and natural language processing domains.

On the European level we synchronize activities, technical protocols, standards, and best practices and contribute to the development in these fields.



Information

re3data coref

What information on research data repositories do you generally need at your organization?

Information about the repositories offered through our research infrastructure. We already keep this information in the CLARIN centre registry but for completeness, visibility, usability, and cooperation we also like to see this information stored in and provided through re3data.

What information on research data repositories is currently missing (i.e., in general / for your organization) in re3data?

More up to date information about the CLARIN entries. Some entries are a bit out of sync (e.g. certification status).

Data Export from re3data

How do you currently access repository descriptions in re3data?

We are mostly aware about external users searching for CLARIN repositories via re3data. Many CLARIN centres find it important to have a good visibility to the research community through as many channels as possible, including re3data.

How would you prefer to access repository descriptions in re3data (e.g., API, dashboards, website, subsets based on (which) criteria, spreadsheet export ...)?

Mainly through the website.

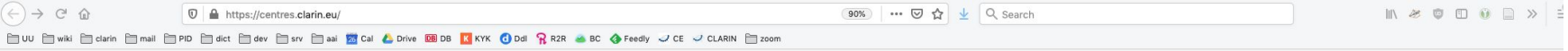
Data Import to re3data

What valuable information could your organization add to repository descriptions in re3data?

All repository-related information as accessible via the CLARIN centre registry.

How would you prefer to add information to the service (e.g., via API, metadata upload, etc.)?

We provide all of the information in CLARIN via an API, hence re3data could harvest this information at any time. However, to keep the information synced, it would be helpful for us to expose our data via a re3data-compliant API. If there is such a description we could even consider providing support for such an API, next to the existing APIs. The main aim of this exercise would be more up-to-date information on the side of re3data.



Show all entries Search:

Centre	Shorthand	Type	Type status	Assessment dates	City	Consortium	CoreTrustSeal/DSA
ASV Leipzig	ASV	B	Certified	2019-02-19 valid till 2022-02-19 for [B]	Leipzig	CLARIN-D (DE)	seal
Austrian Centre for Digital Humanities and Cultural Heritage - A Resource Centre for the HumanitiEs	ACDH-ARCHE	B	Certified	2018-02-05 valid till 2022-03-19 for [B]	Vienna	CLARIAH-AT (AT)	seal
Bayerisches Archiv für Sprachsignale	BAS	B	Certified	2019-05-17 valid till 2022-05-18 for [B]	München	CLARIN-D (DE)	seal
Berlin-Brandenburg Academy of Sciences and Humanities	BBAW	B	Certified	2018-10-25 valid till 2021-10-25 for [B]	Berlin	CLARIN-D (DE)	seal
Center of Estonian Language Resources	CELR-EKK	B	Certified	2018-11-02 valid till 2021-11-02 for [B]	Tartu	CELR (EE)	seal
Centre for Language and Speech Technology	CLST	C	*none*		Nijmegen	CLARIAH (NL)	*none*
Centre for the Digital Foundation of Research in the Humanities, Social, and Educational Sciences	CEDIFOR	C	Aiming for B		Frankfurt, Germany	CLARIN-D (DE)	*none*
CLARIN Centre of Latvian language resources and tools	CLARIN-LV	C	*none*		Riga	CLARIN-LV (LV)	*none*
CLARIN Knowledge Center for Terminology Resources and Translation Corpora (TRTC)	TRTC	K	Certified	2019-01-16 valid till 2022-01-15 for [K]	Vienna	CLARIAH-AT (AT)	*none*
CLARIN Knowledge Centre for Atypical Communication Expertise	ACE	K	Certified	2019-06-19 valid till 2022-06-18 for [K]	Nijmegen	CLARIAH (NL)	*none*
CLARIN Knowledge Centre for Belarusian text and speech processing	K-BLP	C K	Certified K	2020-02-10 valid till 2023-02-09 for [K]	Minsk	*none*	*none*
CLARIN Knowledge Centre for Data Management at NSD	NSD	K	Certified	2018-12-04 valid till 2021-12-04 for [K]	Bergen	*none*	*none*
CLARIN Knowledge Centre for linguistic diversity and language documentation	CKLD	K	Certified (K)	2018-03-28 valid till 2021-03-27 for [K]	Cologne	*none*	*none*
CLARIN Knowledge Centre for South Slavic Languages	CLASSLA	K	Certified	2019-03-19 valid till 2022-03-18 for [K]	Ljubljana	CLARIN.SI (SI)	*none*
CLARIN Knowledge Centre for Speech Analysis	CLARIN-SPEECH	K	Certified	2019-05-17 valid till 2022-05-18 for [B]	Stockholm	SWE-CLARIN (SE)	*none*
CLARIN Knowledge Centre for Swedish in a Multilingual Setting (SMS)	CLARIN-SMS	K	Certified	2019-01-16 valid till 2022-01-15 for [K]	Stockholm	SWE-CLARIN (SE)	*none*

Note: the centre registry is *not* intended to identify depositing services

Depositing Services

One of the fundamental services of the CLARIN infrastructure is making sure that language resources can be archived and made available to the community in a reliable manner. To help researchers to store their resources (e.g. corpora, lexica, audio and video recordings, annotations, grammars, etc.) in a sustainable way, many of the CLARIN centres offer a **depositing service**. They are willing to store the resources in their repository and assist with the **technical and organisational details**. This has a wide range of advantages:

- **Long-term archiving**: a storage guarantee can be given for a long period (up to 50 years in some cases)
- Resources can be cited easily with a **persistent identifier**.
- The resources and their metadata will be integrated into the infrastructure, making it possibe to **search** them efficiently.
- Password-protected resources can be made available via **an institutional login**.
- Once resources are integrated in the CLARIN infrastructure, they can be analyzed and enriched more easily with various **linguistic tools** (e.g. automated part-of-speech tagging, phonetic alignment or audio/video analysis).

The following certified CLARIN centres are offering depositing services:

Centre	Location	Depositing offer
CLARIN Centre Vienna	Austria	Any linguistic and/or NLP data and tools
LINDAT-Clarin	Czech Republic	Any linguistic and/or NLP data and tools: corpora, treebanks, lexica, but also trained language models, parsers, taggers, machine translation systems, web services, etc.
CLARIN-DK-UCPH	Denmark	Danish language resources The focus is on written and spoken language resources. Possible to deposit: text corpora and texts with annotations, imdi-sessions containing audio, video and annotations of these resources, together with lexicons and other data.
CELR	Estonia	Estonian language resources: texts, corpora, audio and video recordings, lexical data, terminologies, tools for NLP, etc.
FIN-CLARIN	Finland	All language resources related to Finnish, Finland Swedish and the Fenno-Ugric languages, as well as other language resources created in Finland.
BAS	Germany	corpora of spoken languages which contain a minimum of at least one measured signal that is based on the physical processes of speech production (e.g. acoustic signals, videos, series of measurements, series of pictures)
BBAW	Germany	German corpora or parallel corpora, historical prints and manuscripts (in German), lexical resources (also in German)

Monitoring

How do you use re3data to monitor the repository landscape?

We mainly monitor CLARIN-related repositories.

How could re3data improve to support your analysis?

The current classification by subject is hard to relate to the information that is exposed by the CLARIN Centre Registry – it is not always on which information this classification is based.
For example, many CLARIN centres provide repositories in the area of the humanities, often related to linguistics and other fields, but they currently do not show up in the subject based visualization and search.

Reference

re3data coref

Do you use re3data to persistently refer to repositories?

No

How could the reference to repositories be improved?

By having the DOIs refer to the actual repository and not to the re3data landing page.

How would you use a PID for repositories?

We mostly use PIDs to refer to data resources rather than to repositories.

Recommendation

re3data coref

What kind of information (i.e., metadata fields) do you use to select recommended repositories?

We ourselves do not actively search for repositories to deposit data ourselves.

If we receive inquiries from the user community, we recommend repositories based on a) subject field b) data type c) trust level/certification status d) country or affiliation of the depositor.

Do you need predefined recommendations (e.g., based on the FAIR Data Principles, journal requirements, community standards)?

We are already using predefined recommendations in suggesting to use CLARIN certified repositories.

How would you like to export these recommendations from re3data?

By the CLARIN certified metadata field.

Other Use Cases

re3data coref

Are there any other use cases or requirements we did not address so far?
Do you have any other concluding remarks?

We noticed that re3data now also includes CLARIN Knowledge Centres, while these do not necessarily offer a repository. It might be good to look into this.

Prioritization of Use Cases

re3data coref

Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.
2.
3.
4.
5.

Prioritization of Service Gaps

re3data coref

Which of the service gaps mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.
2.
3.
4.
5.

Thank You!

Tell us your use case story:

info@re3data.org

www.re3data.org

@re3data

doi.org/10.17616/R3D

Credits

Thanks to Thorsten Trippel from the university of Tübingen & CLARIN-D for co-authoring our feedback!

FAIRsFAIR Recommendations

re3data Use Case Analysis

Ilona von Stein (DANS, The Netherlands), (https://orcid.org/0000-0003-3179-0773), 26 November 2020

representing the thinking of the FAIRsFAIR project

Introduction

supplying **practical solutions** for the use of the FAIR data principles throughout the research data life cycle;

fostering **FAIR data culture** and the uptake of good practices in making data FAIR;

- 10 million euro
- 36 months
- Starting date: March 1 2019
- 22 partners from 8 European countries

https://www.fairsfair.eu/

FAIRSFAIR

Fostering Fair Data Practices in Europe

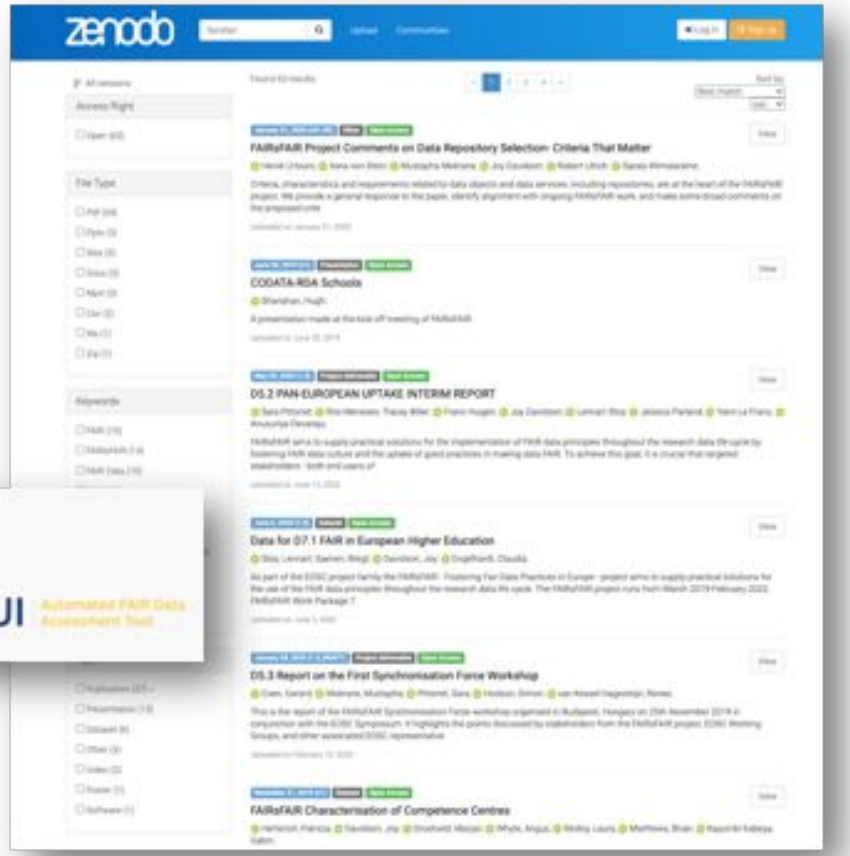
Wp1	Project Management and Sustainability
Wp2	FAIR Practices: Semantics, Interoperability, and Services
Wp3	FAIR Policy and Practice
Wp4	FAIR Certification
Wp5	Engagement, Communication and Uptake
Wp6	FAIR Competence Centre
Wp7	FAIR Data Science and Professionalisation

Software & Tools

FAIR Aware

FAIR Data Object Assessment Metrics

F-UJI Automated FAIR Data Assessment Tool



Stay tuned:
https://www.fairsfair.eu/user/register

https://zenodo.org/communities/fairsfair

re3data coref

FAIR certification

- Support the **FAIR-alignment of certification schemes**
- In-depth FAIR-enabling **repository certification support programme**
- FAIR assessment of **digital data objects**
  - F-UJI: programmatic assessment of data FAIRness
  - FAIR-Aware: tool to assess your knowledge and awareness of FAIR
- Improved registry for **finding and selecting** relevant trustworthy **repositories**

FAIRsFAIR perspective

What do we do?

To increase the availability of FAIR data in FAIR-enabling repositories

- Integrated assessment of CoreTrustSeal+FAIR and FAIR digital objects
- Considering possible badging solutions to FAIR-enabling repositories and FAIR objects

What information do we need?

- Additional components to the metadata schema and tools that enable reuse to align with the FAIR data principles
- Improved description of repository metadata
    - Supported by CoreTrustSeal+FAIR alignment of **repository practices** with **object assessment** against FAIR

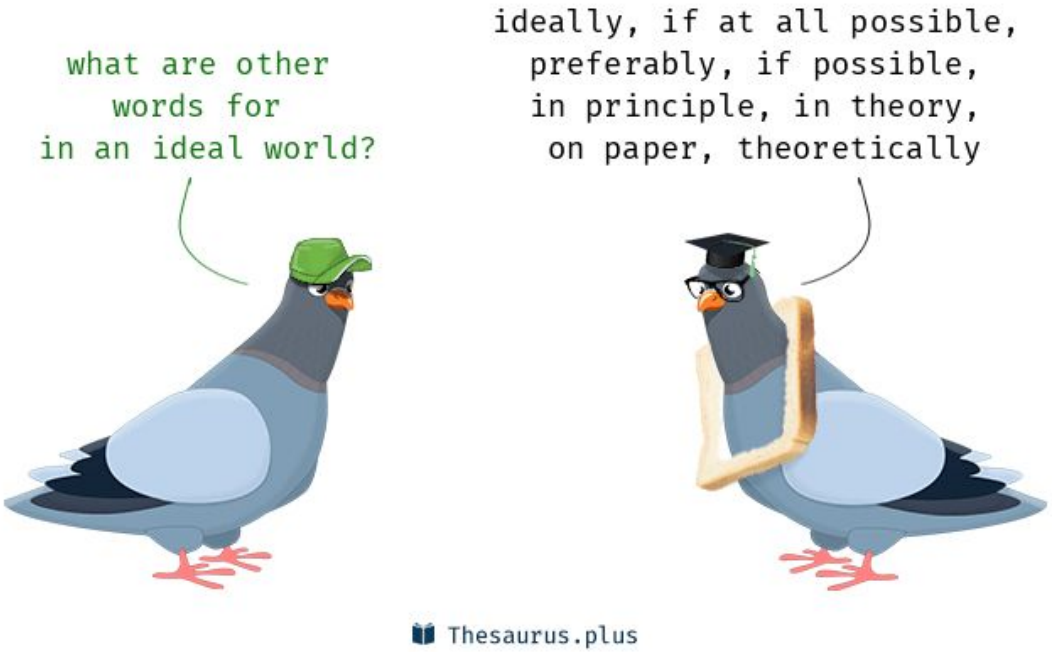
Best case scenario...

FAIRsFAIR deals with **interactions between repository and object** characteristics evaluations

Ideal scenario:

- DataCite badges objects
- DataCite then performs some kind of test with an outcome such as “X% of the collection is being tested as FAIR”
- CoreTrustSeal could then use this outcome in undertaking CoreTrustSeal+FAIR assessment

Best case scenario...



Short term priorities

- **Validation status of re3data records**
  - current status of re3data records have validation status unknown

“self-declared” (e.g. authentication and authorisation of re3data.org property repositoryContact)

vs

“validated” by an external 3rd party (e.g. pulling or pushing values into re3data metadata)

- a solution that can take into account **a change in repository status**
  - If CoreTrustSeal certificate (or other) is withdrawn → update should be possible

Mid term recommendations (1/2)

- **a solution that can take into account a change in repository status**
  - If CoreTrustSeal certificate (or other) is withdrawn → update should be possible
  - Also visa versa: **re3data might also limited temporal metadata** (e.g. Valid To/From)

- a solution that can take into account **a change in object status**
  - FAIRness is a ‘snapshot’ in time
  - What happens after a change in an object áfter it has been tested ‘FAIR’?

Mid term recommendations (2/2)

- re3data to be **aligned with CoreTrustSeal repository types**
  - based on community request for comments in response to demand from repositories beyond domain/subject based curators.
  - highly relevant to our FAIRsFAIR work as well

- re3data to **support automatic monitoring of the landscape**
  - by updating repository metadata to include requirements relating to EOSC participation (currently being defined)

- re3data is in a strong position to offer **PIDs for repositories**
  - to identify repositories as important actors in the research ecosystem
  - include a metadata field to capture existing PIDs for repository host organisations (such as RoR)

Thank you!

Image credit: Ainsley Seago CC BY

FAIRsFAIR seeks solutions to deal with:

- **validation status of re3data records**
- **changes in repository status**
- changes in object status
- alignment with CoreTrustSeal repository types
- automatic monitoring of the landscape
- PIDs for repositories

Twitter: @FAIRsFAIR_EU

fairsfair.eu

ilona.von.stein@dans.knaw.nl

re3data coref

re3data Use Case Analysis:

Supporting researchers in finding a suitable data repository

Sven Vlaeminck, (https://orcid.org/0000-0002-7905-4209), 26 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

DFG

Introduction

re3data coref

[Short introduction of the initiative/institution and its work - What is your current role/position?]

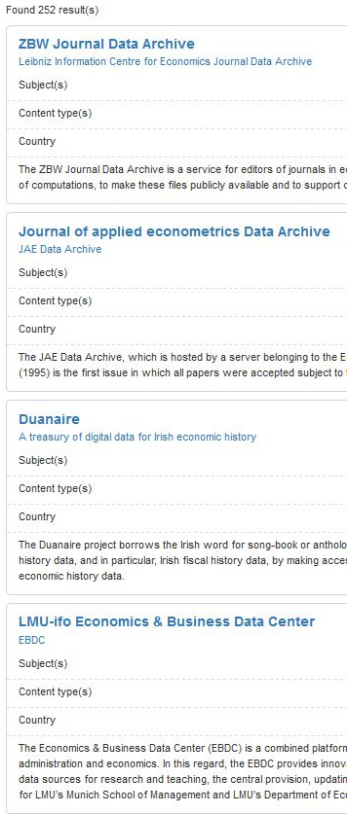
- ZBW: world's largest research infrastructure for economic literature (formerly: German National Library of Economics)
- Part of the Leibniz-Association (non-university research)
- ~270 staff, two locations in Hamburg and Kiel (Germany)
- ZBW is also engaged in RDM and offers some services for economics research
  - ZBW Journal Data Archive (JDA): A data repository for economics journals with the aim to foster reproducible research
  - da|ra (together with GESIS): DOI registration agency for social and economic data (part of DataCite)
  - IREE (journal dedicated to replication studies in economics)
  - in addition: some infrastructure projects (e.g. GerDI); participation in projects and proposals for a national research data infrastructure (NFDI)
- My roles:
  - Product management of the JDA
  - Part of the da|ra team @ ZBW
  - Educational training for doctoral candidates & consulting for institutions

Information

re3data coref

[What information on research data repositories do you generally need at your organization?
What information on research data repositories is currently missing (i.e., in general / for your organization) in re3data?]

- Typical use case: Researchers need advise on where to store their data (from projects or papers)
- Usually, researchers prefer disciplinary data repositories over generalist solutions
- At first glance, most results delivered by re3data are not really satisfactory - in most cases, because the data repos shown in the list do not offer data upload (and that's basically **THE** most important use case for researchers)
- Sure, this information is available. But you have to use the filters on the left side, which might be confusing to use (many options, partially overlapping, too much information...)



Data Export from re3data

*[How do you currently access repository descriptions in re3data?
How would you prefer to access repository descriptions in re3data (e.g., API, dashboards, website, subsets based on (which) criteria, spreadsheet export ...)?]*

- Currently: Access via website (sufficient for guidance and advise)
- Maybe our infrastructure projects might have further demands (access via API should always be sufficient)

Data Import to re3data

[What valuable information could your organization add to repository descriptions in re3data?
How would you prefer to add information to the service (e.g., via API, metadata upload, etc.)?]

- Depends on what is needed. We can provide detailed information (legal, curation, additional services, authority files in use...) on our data repository listed in re3data.
- Upload via website is sufficient. Metadata upload only makes sense when you have many data repositories or when there are frequent updates on the information required by re3data.

Monitoring

[How do you use re3data to monitor the repository landscape? How could re3data improve to support your analysis?]

- For our daily working routines, we do not monitor the repository landscape.
- However, there might be software projects at our institution which might require this information every now and then.

Reference

*[Do you use re3data to persistently refer to repositories?
If yes, how (e.g., repository name, re3data PID, etc.)?
How could the reference to repositories be improved?
How would you use a PID for repositories?]*

- Yes, either by badge or by re3data PID.
- For our purposes, the current reference is sufficient

Recommendation

- Define your most important user groups and their demands and define the most relevant use cases.
- Offer two separate entries for different types of users: Researchers and information professionals

- For researchers: Keep it as simple as possible. Remove (or move to advanced search) most of the filters, as researchers most often will not need these (e.g. syndication, software, database access & restriction, API, Provider type, …)
- Sort list by “Data upload - open” as the most relevant property.

Recommendation (II)

re3data coref

[What kind of information (i.e., metadata fields) do you use to select recommended repositories? Do you need predefined recommendations (e.g., based on the FAIR Data Principles, journal requirements, community standards)?
How would you like to export these recommendations from re3data?]

- Predefined recommendations might be useful (e.g. funder requirements, publisher requirements)
- These predefined recommendations should be easy to select (e.g. publisher or funder requirements, -> drill down navigation?)

- ...and one last thing: Don't try to build an eierlegende Wollmilchsau aka Swiss Army Knife :)

Source: "Wollmilchsau" by Georg Mittenecker.. License: CC-BY-SA 2.5

Thank You!

Contact:

s.vlaeminck@zbw.eu

www.journaldata.zbw.eu

@ZBW_news

re3data coref

re3data Use Case Analysis: Open Access Office Berlin

Maxi Kindling (https://orcid.org/0000-0002-0167-0466), 24 November 2020

@maxi_ki

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

Karlsruher Institut für Technologie

DFG

Introduction

re3data coref

Two relevant positions (part-time)

- Project member of re3data COREF at Humboldt-Universität zu Berlin (IBI)
- Open Access consultant and lead of Open Access Office Berlin (Open-Access-Büro Berlin) located at Freie Universität Berlin

open access büro berlin

@oa_berlin

Open Access Office financed by the City / Federal State of Berlin

Support of the Berlin research institutions (focus on universities and universities of applied sciences)

To implement the Berlin Open Access Strategy (2015)

Introduction

re3data coref

Berlin Open Access Strategy (2015)

Scholarly publications

60% Open Access (green & gold) till 2020 for journal articles

Increase of open access share for monographs and collections

Research Data

Commitment to open access and reuse of research data

Contribution to strategies, coordinated nationally and internationally

Cultural Heritage

Continuation of digitalization projects

Open Access and reuse

Introduction

re3data coref

Berlin Open Research Initiative (work in progress)

Open Access Office Berlin recommends e.g. Berlin Open Research Strategy including:

Strong emphasis on Open and FAIR Research Data provided by Berlin researchers at all institution types to improve visibility and re-use of research outputs

Acknowledgement of Open Research Practices

- Open Access versions publications
- Open Research Data
- Open Software
- Open Protocols

“Open Research Monitoring”

- Visibility of Open Research output
- Services offered by or in cooperation with Berlin research institutions (Mapping)

Information

What information on research data repositories do you generally need at your organization?

- *Landscape of research data repositories on the regional level resp.on the level of the German Bundesländer (federal states)*
- *Berlin state funded research organizations that offer research data repositories or research data services*
    - *types of institutions/organizations, service types, serving disciplines, funding, amount of data sets, standards, data access*

What information on research data repositories is currently missing (i.e., in general / for your organization) in re3data?]

- *Filter for German Bundesländer (and probably some other countries may have similar interests) based on institutions (allocation from institution to Bundesland)*
- *Funding information of the services (Open Initiatives?)*
- *Funding information for the data sets / collections*
- *Service types of software and protocol repositories*

Data Export from re3data

How do you currently access repository descriptions in re3data?

Web interface and API

How would you prefer to access repository descriptions in re3data (e.g., API, dashboards, website, subsets based on (which) criteria, spreadsheet export ...)?

Dashboards with csv export

Data Import to re3data

What valuable information could your organization add to repository descriptions in re3data?
How would you prefer to add information to the service (e.g., via API, metadata upload, etc.)?

Some kind of metadata curation for Berlin research institutions

Monitoring

How do you use re3data to monitor the repository landscape?

Search for Berlin located research institutions

How could re3data improve to support your analysis?

ROR metadata (Metadata could be curated by the office)

Funding metadata etc. (see above)

Metadata search and analysis on data level (dream big ;-))

Reference

re3data coref

Do you use re3data to persistently refer to repositories?

Planned for a mapping (accompanying study of the Berlin Open Research Initiative)

If yes, how (e.g., repository name, re3data PID, etc.)?

Name, institution, re3data PID

How could the reference to repositories be improved?
How would you use a PID for repositories?

If the re3data data PID were an established, standardized repository reference it would be helpful e.g. to achieve comparability (Open Research Monitoring on regional, state, country level)

Recommendation

re3data coref

What kind of information (i.e., metadata fields) do you use to select recommended repositories?
Do you need predefined recommendations (e.g., based on the FAIR Data Principles, journal requirements, community standards)?
How would you like to export these recommendations from re3data?

See above: all that is helpful to be a sound database to sketch the Berlin Open Research landscape

Other Use Cases

re3data coref

[Are there any other use cases or requirements we did not address so far?
Do you have any other concluding remarks?]

Prioritization of Use Cases

re3data coref

Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1 Information
2 Monitoring
3. Reference
4.
5.

Prioritization of Service Gaps

re3data coref

Which of the service gaps mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.
2.
3.
4.
5.

Thank You!

Tell us your use case story:

info@re3data.org

www.re3data.org

@re3data

doi.org/10.17616/R3D

re3data coref

re3data Use Case Analysis: OpenAIRE

Jochen Schirrwagen, (https://orcid.org/0000-0002-0458-1004), 24 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

Karlsruher Institut für Technologie

DFG

Introduction - Briefly about OpenAIRE

re3data coref

"Bridging the worlds where science is **performed** and science is **published**"

https://twitter.com/dasaptaerwin

OpenAIRE

SUPPORT	OS Helpdesk /RDM to researchers & scientists
ACCELERATE	Interoperability / Exchange to content providers, RIs
MONITOR	Open Science / Research outcomes To funders, institutions, RIs, research communities

pan-european network of 34 national open access desks
participatory, scholarly communication infrastructure

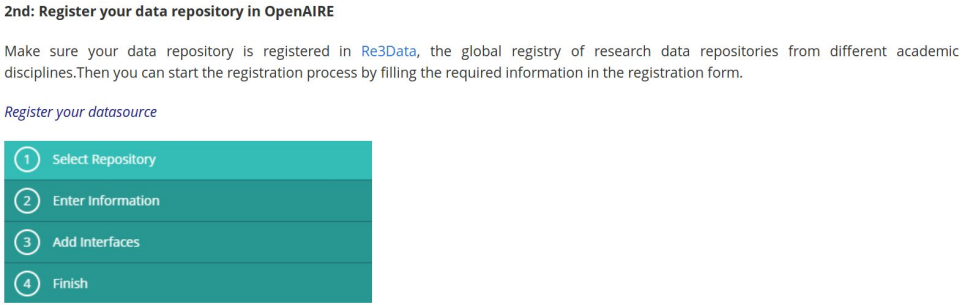
On a Side Note

re3data coref

MoU between re3data and OpenAIRE, Oct 2013

- “to work jointly to facilitate research data registration, discovery, access and re-use”
- “metadata exchange to enable easy identification of appropriate research data repositories”
- “OpenAIRE will integrate data repositories indexed in the re3data.org registry … return ... usage statistics for datasets … inferred links between data and publications”
- “intensify dialogue on best-practices and possible standards and guidelines for the emerging research data repository landscape”

https://www.openaire.eu/re3data-and-openaire-sign-memorandum-of-understanding

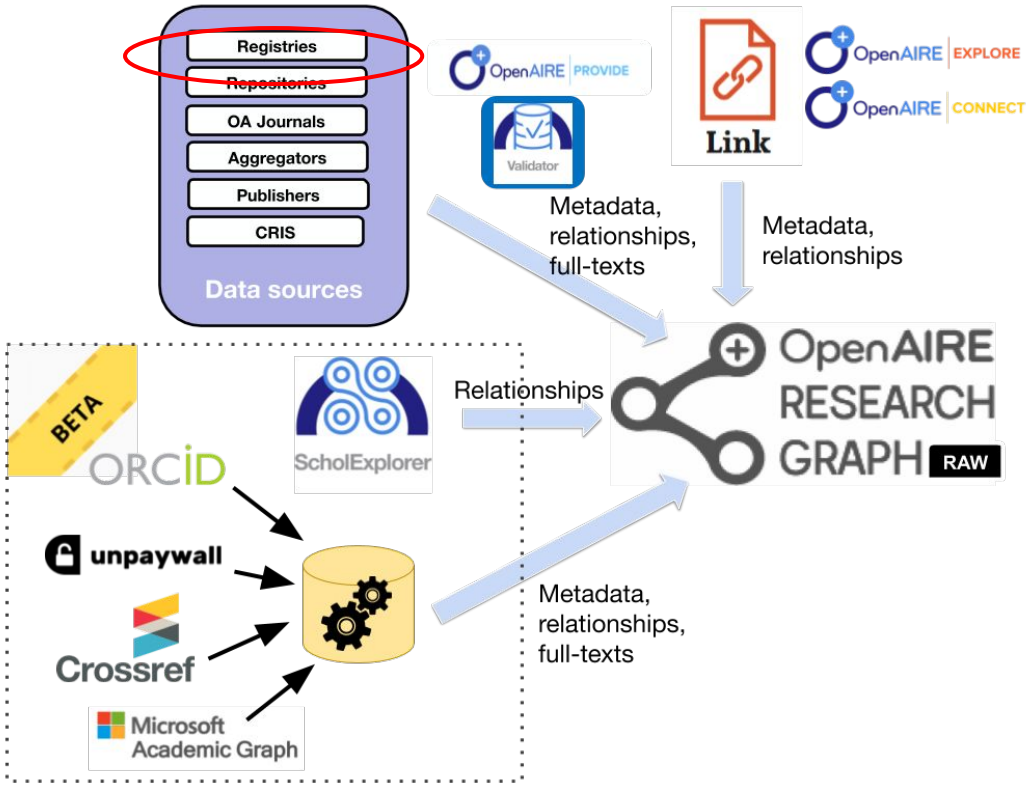


https://www.openaire.eu/register-your-data-repository-in-openaire

Participant Portal H2020 online manual

Introduction - Briefly about OpenAIRE

re3data coref



authoritative lists of entities used by OpenAIRE

Registries of researchers

ORCID

Registries of organizations

GRID

CORDIS

and other funders

Registries of projects

CORDIS

FCT Fundação para a Ciência e a Tecnologia

ACADEMY OF FINLAND

NWO

Science Foundation Ireland sfi For what's next

Australian Government Australian Research Council

and other funders

Registries of data sources

re3data.org REGISTRY OF RESEARCH DATA REPOSITORIES

OpenDOAR

DOAJ DIRECTORY OF OPEN ACCESS JOURNALS

+ journal lists from publishers

Information

re3data coref

What we need and use from re3data

- unique identifier of the repository, e.g. service-DOI of the repository
- URL of the repository
- repository name
- institution or responsible organization
- geo-location
- APIs
- repository contact
- repository type

What we miss from re3data

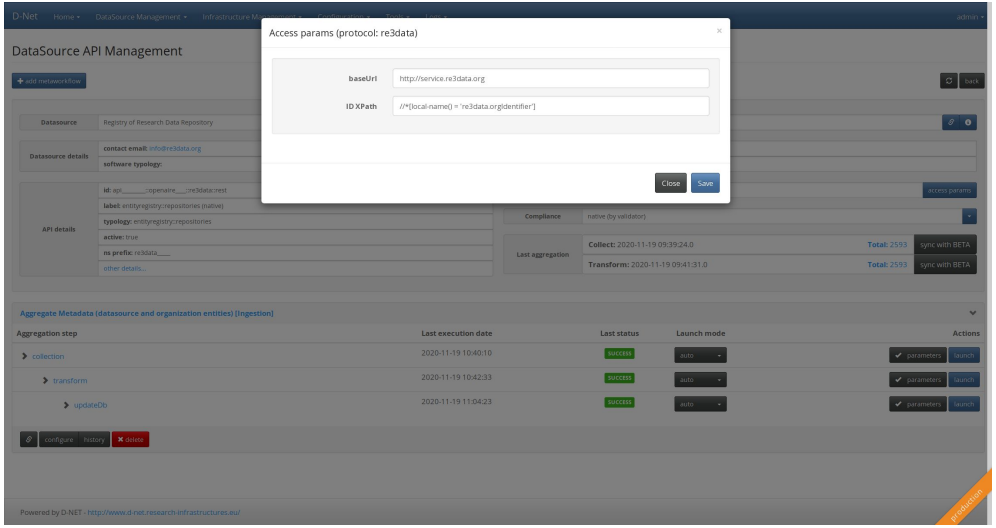
- geo-location of the repository
- extended vocabulary for content types (ie. taking into account data repositories which store not only datasets but also literature
- (lack of re3data repository PID in DataCite bibliographic metadata records)

Data Export from re3data

re3data coref

OpenAIRE is exploring and using the re3data API

- collecting all information about all repositories



DataSource Aggregation backend

```
<?xml version="1.0" encoding="UTF-8"?>
<oai:record xmlns:dri="http://www.driver-repository.eu/namespace/dri"
  xmlns:oai="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <oai:header>
    <dri:objIdentifier>re3data_____::000b0e9fe52ff4d9aefaecd46fc39934</dri:objIdentifier>
    <dri:recordIdentifier>r3d100013215</dri:recordIdentifier>
    <dri:dateOfCollection>2020-11-19T09:38:56.543Z</dri:dateOfCollection>
    <dri:repositoryId>2d627c6e-10c1-43a0-89da-ace43257acc1_UmVwb3NpdG9yeVNlcnZpY2VSZXNvdXJjZXMvUmVwb3NpdG9yeVNlcnZpY2VSZXNvdXJjZVRScGU=
</dri:repositoryId>
    <dri:datasourceprefix>re3data_____</dri:datasourceprefix>
  </oai:header>
  <oai:metadata>
    <!--re3data.org Schema for the Description of Research Data Repositories. Version 2.2, December 2014. doi:10.2312/re3.006-->
    <r3d:re3data xmlns:r3d="http://www.re3data.org/schema/2-2" xsi:schemaLocation="http://www.re3data.org/schema/2-2 http://schema.re3data.org
/2-2/re3dataV2-2.xsd">
      <r3d:repository>
        <r3d:re3data.orgIdentifier>r3d100013215</r3d:re3data.orgIdentifier>
        <r3d:repositoryName language="eng">eData: the STFC Research Data Repository</r3d:repositoryName>
        <r3d:repositoryURL>https://edata.stfc.ac.uk/</r3d:repositoryURL>
        <r3d:description language="eng">eData is an institutional repository where STFC staff can deposit data and software that underpin journal
articles and other published research.</r3d:description>
        <r3d:repositoryContact>edata@stfc.ac.uk</r3d:repositoryContact>
        <r3d:type>institutional</r3d:type>
        <r3d:size updated="2020-01-31">706 datasets</r3d:size>
        <r3d:startDate/>
        <r3d:endDate/>
        <r3d:repositoryLanguage>eng</r3d:repositoryLanguage>
        <r3d:subject subjectScheme="DFG">3 Natural Sciences</r3d:subject>
        <r3d:subject subjectScheme="DFG">308 Optics, Quantum Optics and Physics of Atoms, Molecules and Plasmas</r3d:subject>
        <r3d:subject subjectScheme="DFG">309 Particles, Nuclei and Fields</r3d:subject>
        <r3d:subject subjectScheme="DFG">311 Astrophysics and Astronomy</r3d:subject>
        <r3d:subject subjectScheme="DFG">32 Physics</r3d:subject>
```

stored metadata record about a repository

Data Import to re3data

[What valuable information could your organization add to repository descriptions in re3data?
How would you prefer to add information to the service (e.g., via API, metadata upload, etc.)?]

There is a dedicated task in OpenAIRE-Advance

- Interfacing with scholarly communication services

Feedback from OpenAIRE users and via regular harvesting activities

- Suggested corrections or additions e.g.
  - on repository institution
  - repository manager contact information
  - information on APIs (OAI-PMH, REST, ...)

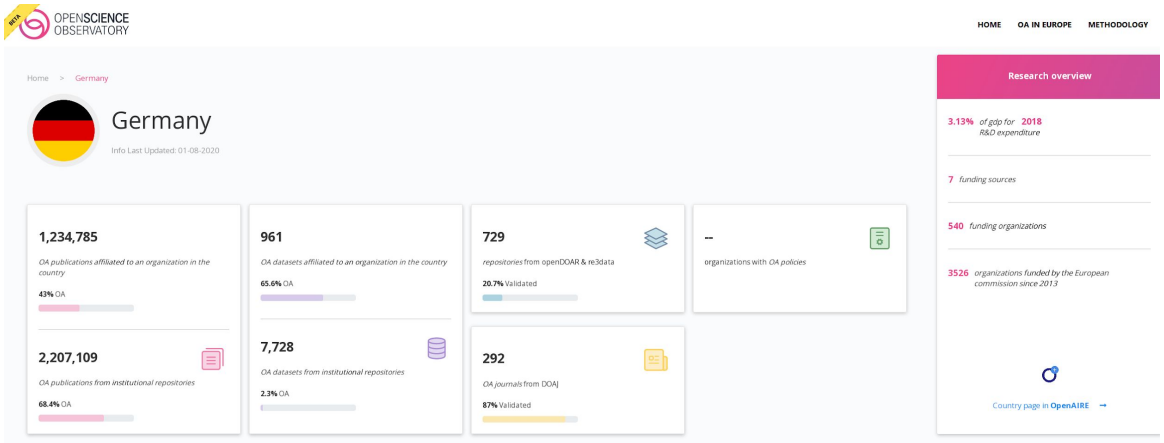
The preferred way to add such information is via API.

Monitoring

[How do you use re3data to monitor the repository landscape?
How could re3data improve to support your analysis?]

OpenAIRE is developing an Open Science Observatory, https://osmonitor.openaire.eu/home

- providing different views on
  - region, country
  - OA datasets by repository, organization, funder

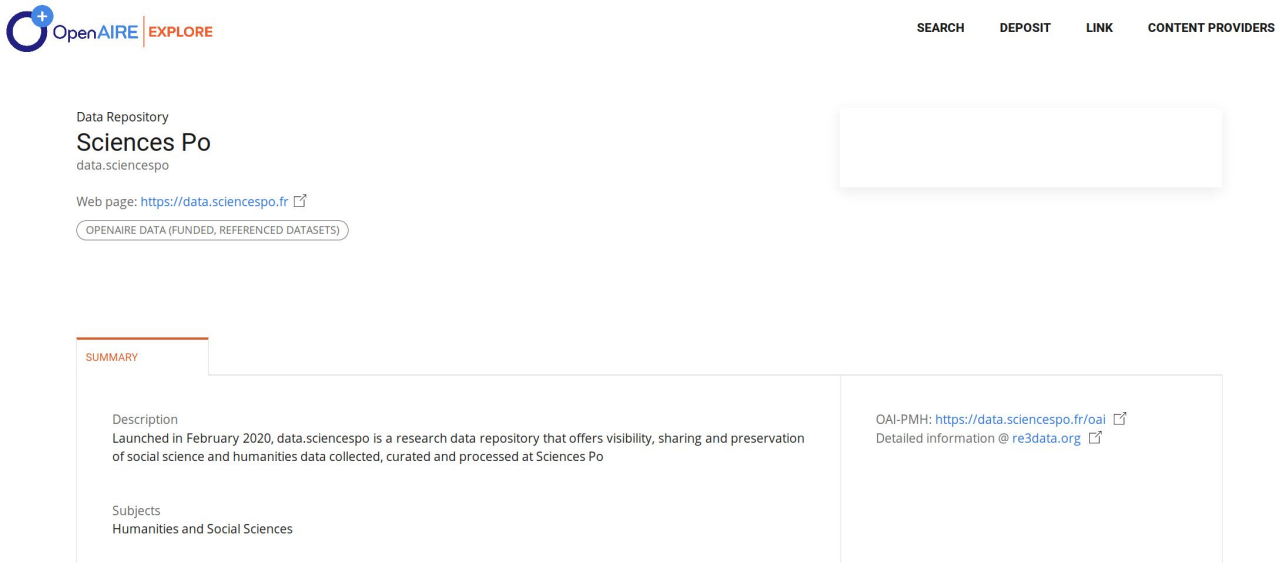


re3data reference in the OpenAIRE Explore portal

re3data coref

Do you use re3data to persistently refer to repositories?
If yes, how (e.g., repository name, re3data PID, etc.)?

- In https://explore.openaire.eu by referring to the URL and using the re3data repository PID, e.g. http://service.re3data.org/repository/r3d100011726

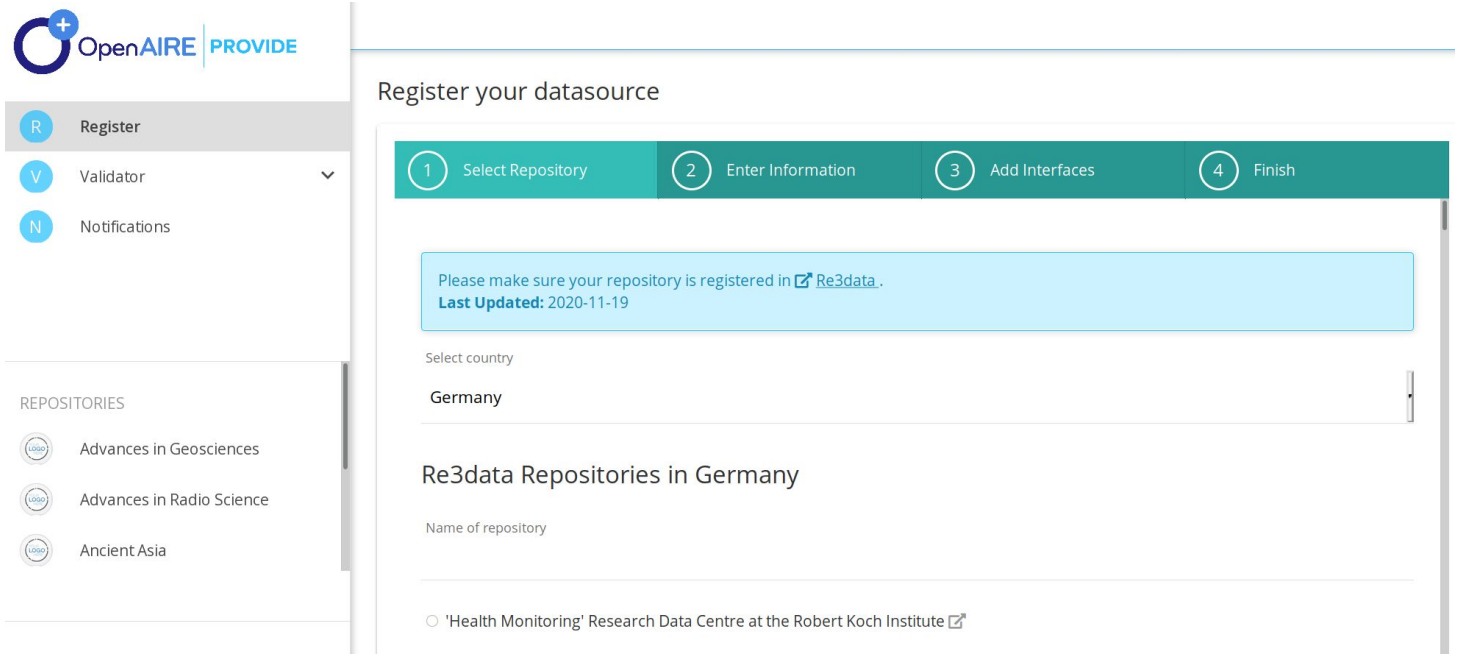


re3data reference in the OpenAIRE Provide Dashboard

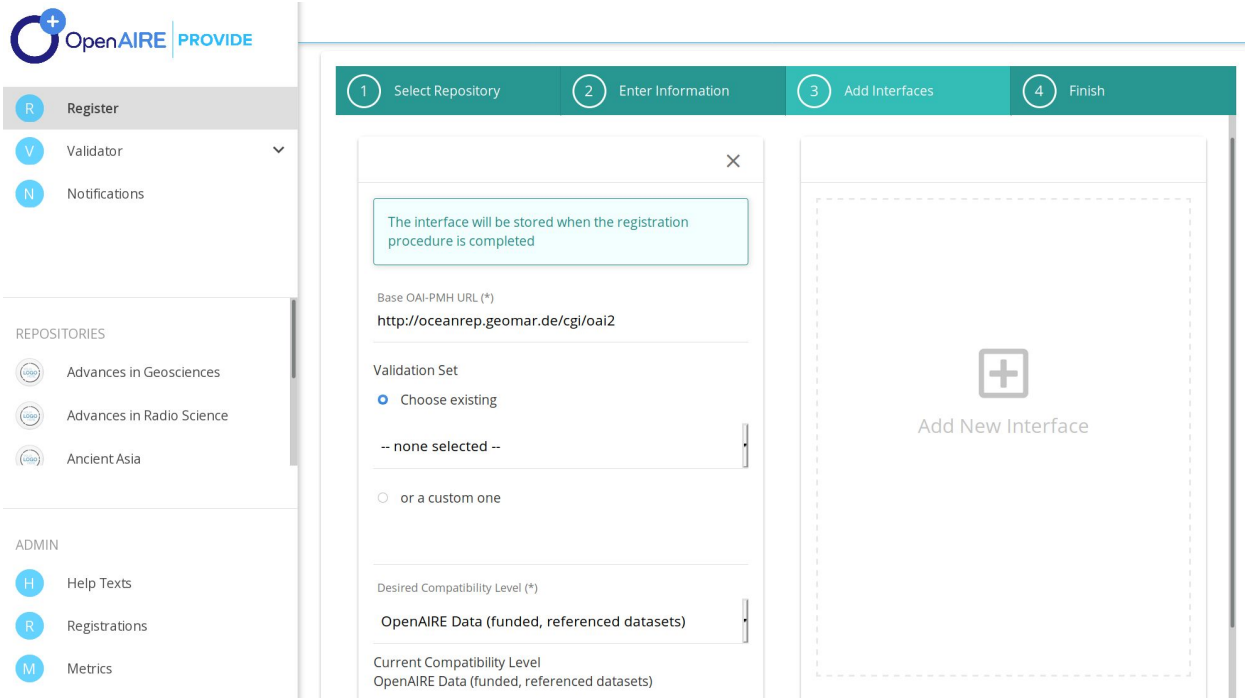
Do you use re3data to persistently refer to repositories?
If yes, how (e.g., repository name, re3data PID, etc.)?

In https://provide.openaire.eu in the process of data repository registration in OpenAIRE by repository name and repository URL

OpenAIRE Registration Interface for Repositories



OpenAIRE Registration Interface for Repositories



Reference

How could the reference to repositories be improved?

We don’t see any issue with referencing a repository in re3data but with unique and persistent identification of the same repository which is listed in various repository registries like re3data, OpenDOAR and FAIRsharing (example zenodo, institutional repositories)

registered in OpenDOAR →

Institutional Repository
Publications at Bielefeld University
OPENAIRE 3.0 (OA, FUNDING)
Website URL: http://pub.uni-bielefeld.de/
OAI-PMH URL: http://pub.uni-bielefeld.de/oai
Subject: Multidisciplinary
This site provides access to the research output of the institution. Some content is not available as full-text. Academics may also make use of publication lists. The interface is available in German and English.

registered in re3data →

Data Repository
PUB Data Publications
OPENAIRE DATA (FUNDED, REFERENCED DATASETS)
Website URL: https://pub.uni-bielefeld.de/data?lang=en
OAI-PMH URL: https://pub.uni-bielefeld.de/oai
Subject: Humanities and Social Sciences, Life Sciences, Natural Sciences

How would you use a PID for repositories?

- using the PID to refer to the repository entry in re3data
- using the PID to resolve to the repository landing page itself.

Recommendation

re3data coref

What kind of information (i.e., metadata fields) do you use to select recommended repositories?

- used metadata fields: subjects, name, description, organisation, country

Do you need predefined recommendations (e.g., based on the FAIR Data Principles, journal requirements, community standards)?

- recommendations or indicators based on FAIR Data Principles and / or FAIR Data Maturity;
- community and cross-domain (metadata) standards

How would you like to export these recommendations from re3data?

- via API

Other Use Cases

Are there any other use cases or requirements we did not address so far?
Do you have any other concluding remarks?

Enrichment of institutions & organisations responsible to a repository with organisation-Ids like ROR

Prioritization of Use Cases

re3data coref

Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.
2.
3.
4.
5.

Prioritization of Service Gaps

re3data coref

Which of the service gaps mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.Lack of FAIR Data indicators
2.Unique identification of a repository across several registries
3.Org-IDs for institutions / organisations
4.Additional metadata fields
5.

Thank You!

jochen.schirrwagen@uni-bielefeld.de

info@openaire.eu

http://www.twitter.com/OpenAIRE_eu

Credits

The OpenAIRE NOADs,
https://www.openaire.eu/contact-noads

The OpenAIRE Research Graph team,
https://www.openaire.eu/research-graph-team

re3data coref

re3data Use Case Analysis: ORCID

Paloma Marín-Arraiza, (https://orcid.org/0000-0001-7460-7794)

Tom Demeranville, (https://orcid.org/0000-0003-0902-4386), 26 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

DFG

Introduction

re3data coref

Paloma Marín-Arraiza
@pmarrai
Engagement Lead Europe

Tom Demeranville
@tomdemeranville
Product Director

ORCID's Mission and Vision

re3data coref

ORCID's vision is a world where all who participate in research, scholarship, and innovation are uniquely identified and connected to their contributions across disciplines, borders, and time.

ORCID is part of the wider digital infrastructure needed for researchers to share information on a global scale. We enable transparent and trustworthy connections between researchers, their contributions, and affiliations by providing an identifier for individuals to use with their name as they engage in research, scholarship, and innovation activities.

ORCID iD

https://orcid.org/0000-0001-7460-7794

- Persistent and unique identifier for people linked to research (researchers and contributors)
- 16-character alphanumeric code
- Norm ISO 27729 (International standard name identifier - ISNI)
- Registration https://orcid.org/signin

A PID-powered vision

re3data coref

PIDs are only powerful when they are combined!

- PIDS for people — close to 8.9M ORCID iDs
- PIDs for places — over 91k ROR identifiers
- PIDs for "things" — over 115M Crossref DOIs and 30M DataCite DOIs

10M ORCID iDs already!

"As PID adoption has grown, there is increased awareness of the need for organizations, countries, and even regions to develop an overarching PID strategy."

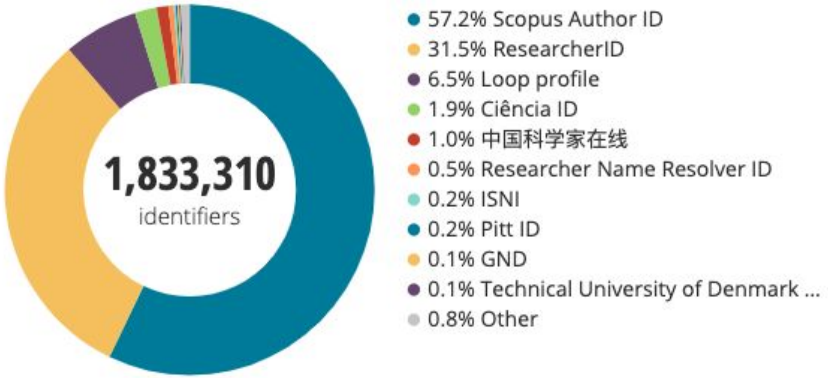
Alice Meadows

Source: https://scholarlykitchen.sspnet.org/2020/06/29/the-uk-national-pid-consortium-a-pathway-to-increased-adoption/
https://orcid.org/blog/2020/11/20/10m-orcid-ids

A “hub” of identifiers

Identifiers

Person identifier types



Records with an identifier

1,345,854

Records with an identifier %

13.0%

Work identifiers

Works

Works

57,448,023

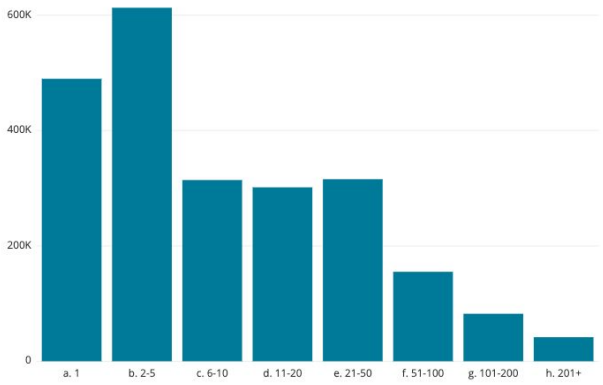
Records with a work

2,312,836

% Records with a work

22.5%

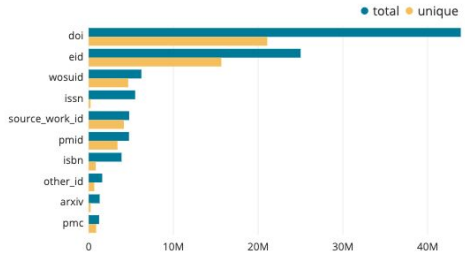
Works per record



work type	count of all work type	percentage
JOURNAL_ARTICLE	46,740,361	81.36
CONFERENCE_PAPER	4,763,039	8.29
BOOK_CHAPTER	1,512,046	2.63
BOOK	1,354,504	2.36
OTHER	1,180,979	2.06
CONFERENCE_ABSTRACT	319,828	0.56

View all 43 rows

Popular Work Identifiers



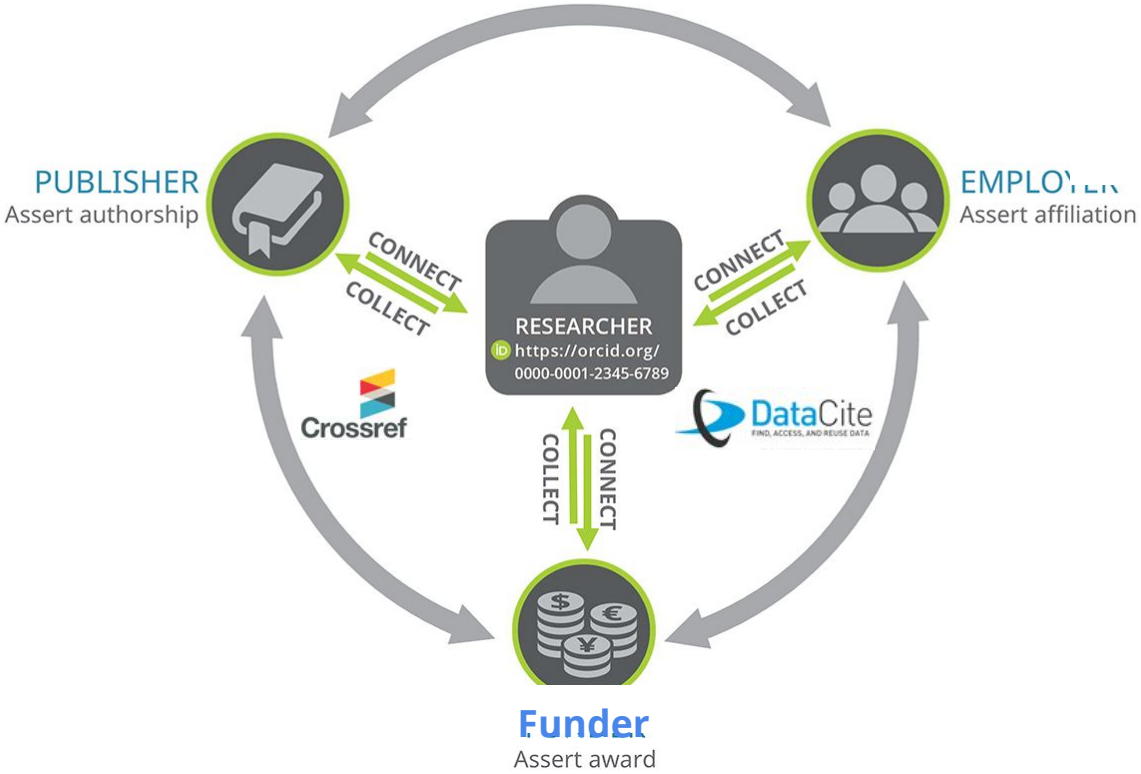
iDs with work identifiers

type	records with type
doi	2,066,790
eid	827,634
issn	598,639
isbn	284,110
pmid	263,747

View all 44 rows

INTEROPERABILITY

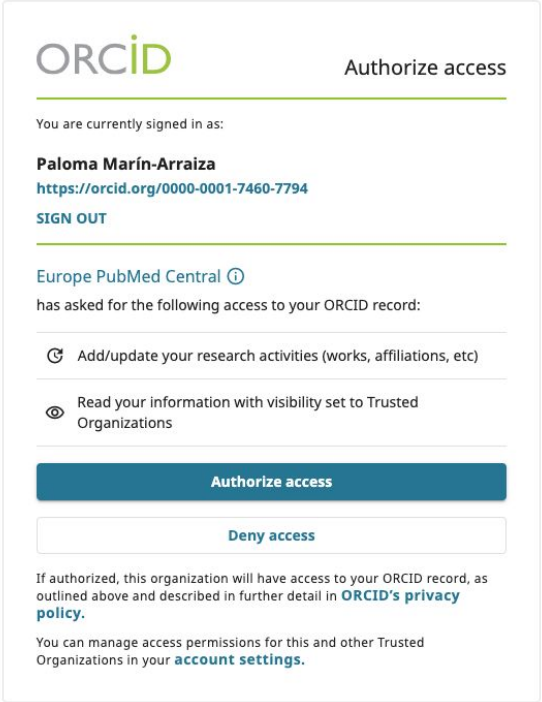
ENTER ONCE
REUSE OFTEN



Enter once, reuse often

re3data coref

- Anyone can read information marked “Everyone” on ORCID records
  - Only ORCID members can read information marked “Trusted Parties”
- Only ORCID members can add information to records
  - They can only edit information they’ve added
- Researcher can revoke permissions at any time

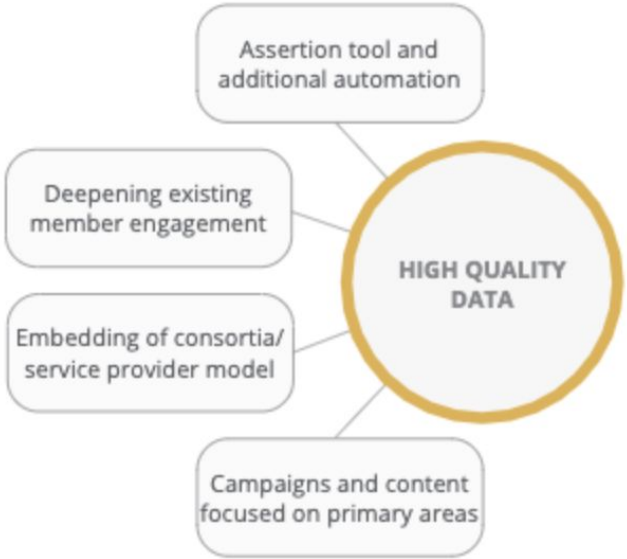


Goal: High Quality Data

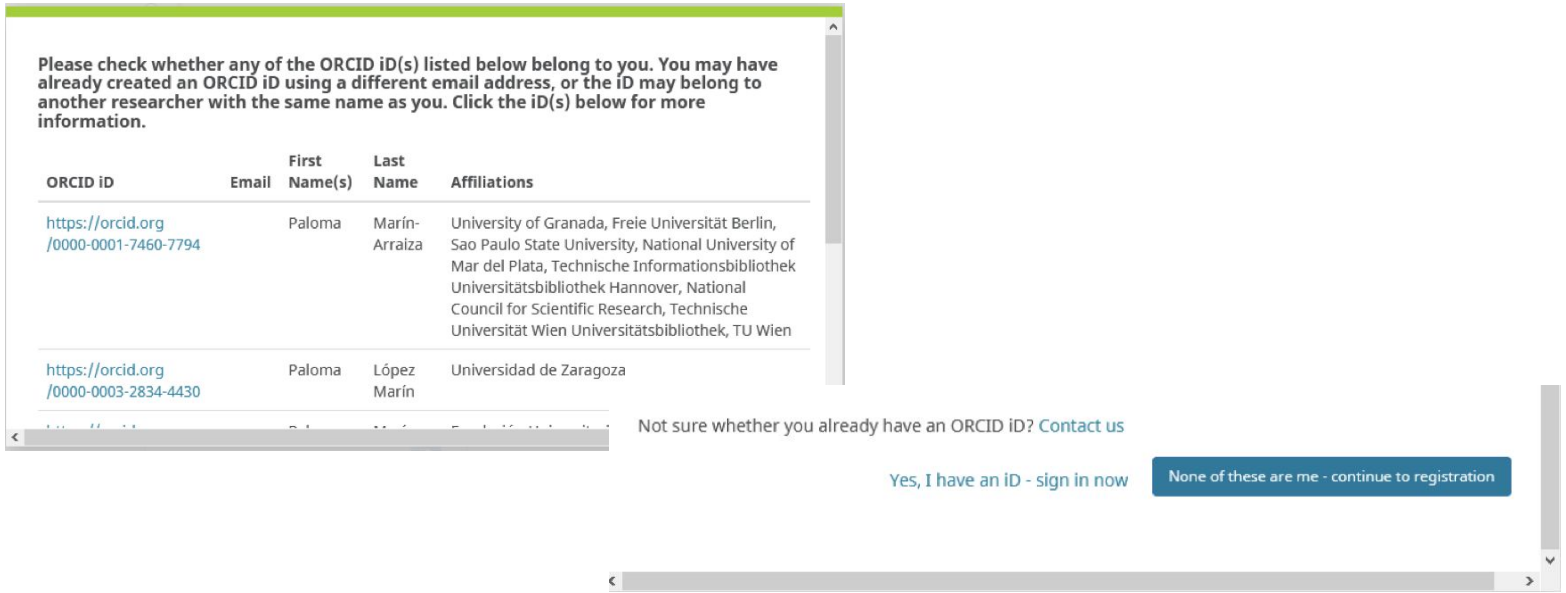
By 2025, we would like researchers and organizations to consider the ORCID record as a reliable and easily accessible data source and to have implemented processes that enable researchers to share their ORCID information.

ACTIVITIES

- Test and launch new products to enable low-tech use of ORCID services
- Deepen the sophistication of member API usage
- **Create channels for deeper engagement with Service Providers and Consortia**
- **Improve communications and maturity modeling of our Consortia and Service**
- **Providers through improved API usage insights reporting**



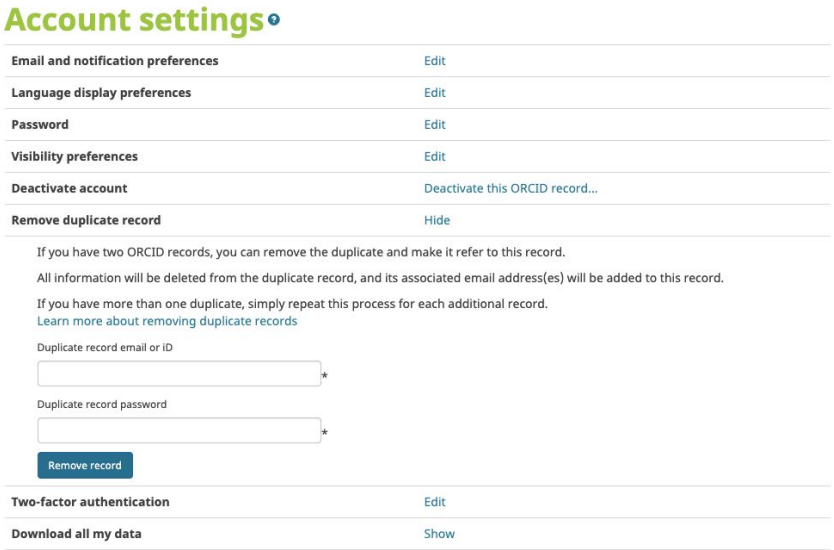
Mechanisms to avoid iD duplicates (I)



- When registering, ORCID warns of possible duplicates.
- It is important to add multiple email addresses (institutional and personal, present and past) to avoid losing access to the account.

Mechanisms to avoid iD duplicates (II)

re3data coref



- Only possible by the user (following the ORCID user-driven principle).
- The information contained is not combined (only the email), only that of the main iD is kept.
- Institutions can collect the main iD through authentication

Removing your additional or duplicate ORCID iD https://support.orcid.org/hc/en-us/articles/360006896634-Removing-your-additional-or-duplicate-ORCID-iD

Product Interest Group

re3data coref

- When? → 7th of December, at 16:00 CET
- What? → Update on our outstanding 2020 projects and a walk through of our 2021 priorities
- Who? → Open for everyone
- Where? → Register here https://register.gotowebinar.com/register/6666276736708961808

Follow our Product Roadmap:
https://trello.com/b/CTB1InDi/orcid-product-roadmap

Thank you!

Contact us:

p.arraiza@orcid.org/ t.demeranville@orcid.org/ support@orcid.org

www.orcid.org

@pmarrai, @tomdemeranville @ORCID_Org

re3data coref

re3data Use Case Analysis: RDMO

Jochen Klar, (https://orcid.org/0000-0002-5883-4273), 24 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

DFG

Introduction

re3data coref

RDMO

The Research Data Management Organiser (RDMO) is a web based open source software tool, which:

- enables institutions as well as researchers to plan and carry out their data management,
- assembles all relevant planning information and data management tasks across the whole life cycle of the research data,
- creates data management plans (DMP) for funders and other stakeholders,
- can be easily installed and operated by an institution or a consortium,
- can be integrated into local infrastructures,
- can be adapted to institutional and disciplinary aspects.

More information: https://rdmorganiser.github.io

Introduction

re3data coref

RDMO started as a project funded by the German Research Foundation (DFG), and is now transitioning to a community-driven working group.

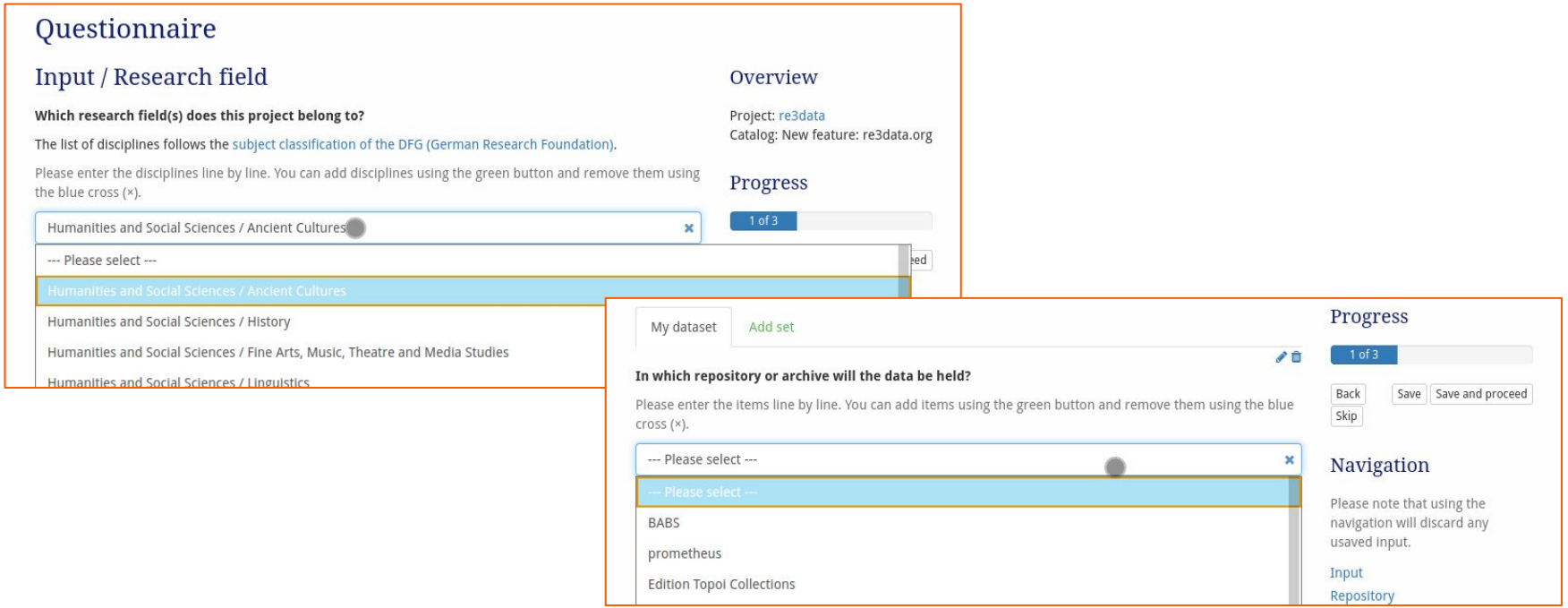
There are ~20 RDMO instances in production (another ~20 testing/evaluation).

An important part of RDMO is the creation and exchange of content within the community. This includes questionnaires, controlled vocabularies, template for output views and related software and extensions.

Data Export from re3data

re3data coref

In 2020 we added a feature to integrate re3data into a RDMO questionnaire.



Data Export from re3data

re3data coref

How does it work?

- Instead of a fixed set of options, questions are assigned a dynamic optionset pointing to a *Provider*.
- The Provider is a Python class which maps the internal vocabulary of RDMO (Attributes) to the *subjects* in re3data, sends a request to the beta API, and parses the returned XML.
- The provider is maintained independent of RDMO and can be customized.
- This is supposed to be a general workflow for RDMO and can be extended to use more RDMO information, as well as other API endpoints.

re3data optionset provider: https://github.com/rdmorganiser/rdmo-plugins/blob/master/rdmo_plugins/optionsets/re3data.py

Reference

re3data coref

Do you use re3data to persistently refer to repositories?

- Usually, options are stored in RDMO as foreign keys in the database.
- For exports RDMO uses its own system of URI.
- For dynamic option sets (like the re3data provider), RDMO stores the *name* value (as free text) and the *id* as external_id (but never uses it).
- Different from other option sets, dynamic option sets do not support multiple languages.
- If re3data would use (other) PID systems, RDMO would adjust.

Recommendation

re3data coref

What kind of information do you use to select recommended repositories?

- Luckily, we use the same DFG Classification re3data uses. We could use more information from the RDMO user input to query re3data.

Do you need predefined recommendations?

- Yes, we could use these directly in RDMO (easy) or could map RDMO input to them (hard).

How would you like to export these recommendations from re3data?

- On a technical level, we found no issues with the way re3data provides its information.

Prioritization of Use Cases

re3data coref

Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.Data Export from re3data
2.Reference
3.Recommendation
4.
5.

Thank You!

Tell us your use case story:

info@re3data.org

www.re3data.org

@re3data

doi.org/10.17616/R3D

re3data coref

re3data Use Case Analysis: RRID

Anita Bandrowski, PhD, (https://orcid.org/0000-0002-5497-0243), 24 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

HUMBOLDT-UNIVERSITÄT ZU BERLIN

KIT Karlsruher Institut für Technologie

DFG

Introduction to RRIDs & the SciCrunch Registry

re3data coref

- RRIDs (global persistent unique identifiers) exist to tell readers what an author used in the study (antibodies, cell lines, organisms, plasmids, tools)

- SciCrunch Registry is a listing of “tools” that authors may use in their study which includes repositories
  - Tools (18K) include: Software (7K), Instruments (>1K*), Core Facilities (2K), Databases (3K), Stock centers (1.5K)

*Instruments are being added by the UsedIT project from FSU Libraries

Download PDF Export Search ScienceDirect Advanced search

NMHC-IIA, binds CM22 in vitro.[7] N2 blocks IgM antibody deposition and inhibits injury in rodent models of intestinal, skeletal muscle, and myocardial I/R injury as well as burn injury and hemorrhagic shock.[8, 9, 10, 11, 12 and 13] Thus, NMHC-IIA appears to be a conserved injury/ischemia antigen in multiple rodent tissues. A critical, unanswered question is whether human antibodies that target NMHC-IIA exist and if so, whether NMHC-IIA is a conserved injury antigen in humans.

To address these questions, we engrafted NOD.SCID.IL2rg knockout mice, which lack endogenous murine NK cells and T and B cells, with human peripheral blood lymphocytes to generate humanized mice.[14 and 15] P
mice stably engraft with human T cells and generate
isotypes.[16 and 17] In this model, we characterize B ce
the phenotype of intestinal I/R injury. We then test th
using this novel humanized mouse model.

Methods

Generation of humanized PBL-SCID mice

NOD.PkSCID.IL2 receptor gamma chain null mice v
Laboratories (Bar Harbor, ME). Human lympho
density centrifugation from peripheral blood obt
Three- to 4-week-old male mice were injected with 2

Why do RRIDs exist?

The Jackson Laboratory

JAX® MICE SEARCH

Search by Stock # | Keyword Search

Enter Search Term

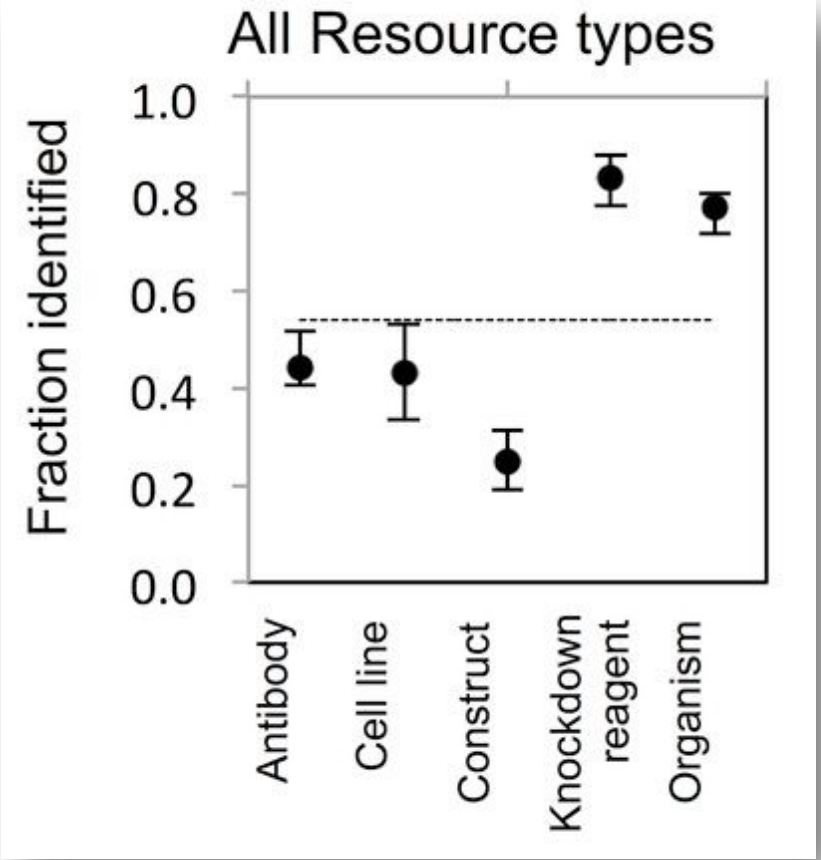
RESULTS FOR: NOD.PkSCID.IL2 ×

Sorry, your search yielded no results

CLEAR ALL

1. **Text Search**
A text search looks for the entire text st
(albumin cre),
search for the first (albumin), then sear
set.

How common is this?



Papers are poor at identifying the simplest part of the paper, the materials used

Vasilevsky 2013

WHAT DOES A SOLUTION LOOK LIKE?

CANCER RESEARCH

Tumor Biology and Immunology

Proteomic Profiling of the ECM of Xenograft Breast Cancer Metastases in Different Organs Reveals Distinct Metastatic Niches

Jess D. Hebert, Samuel A. Myers, Alexandra Naba, Genevieve Abbruzzese, John M. Lamar, Steven A. Carr, and Richard O. Hynes

DOI: 1

Methods

Vendor + Catalog Number + Persistent Unique Identifier (RRID)!

Cells and vectors

The human MDA-MB-231 (ATCC Cat# CRL-12532, RRID:CVCL_0062) mamm
(24) expressing firefly luciferase was a kind gift of Joan Massagué (Memorial Sl
Center, New York, NY). These cells were further retrovirally infected to express

re3data coref

re3 asks: What information on research data repositories do you generally need at your organization?

What information on research data repositories is currently missing (i.e., in general / for your organization) in re3data?

Are there ways of citing repositories?

also

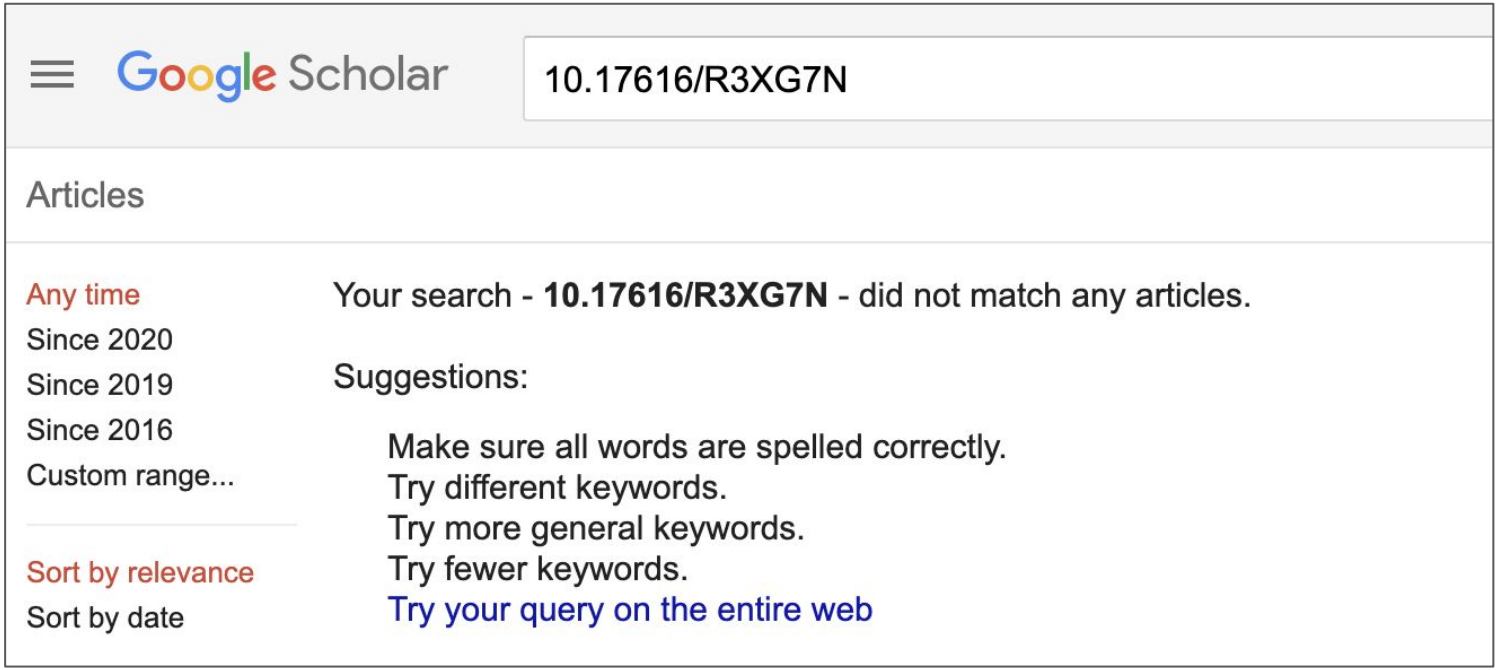
DataCite

Cite this re3data.org record:

re3data.org: The Antibody Registry; editing status 2020-04-02; re3data.org - Registry of Research Data Repositories. http://doi.org/10.17616/R3XG7N last accessed: 2020-11-03

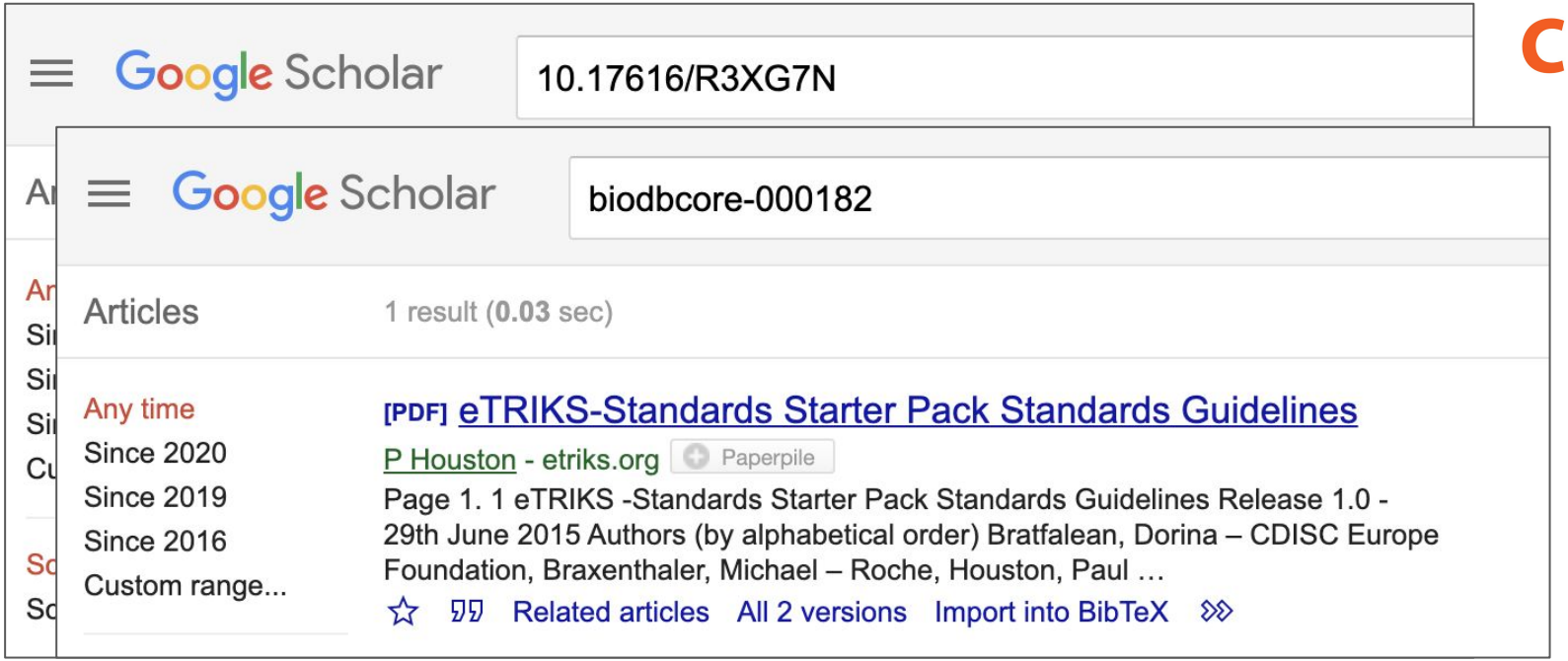
Information: Who cited my repository?

re3data coref



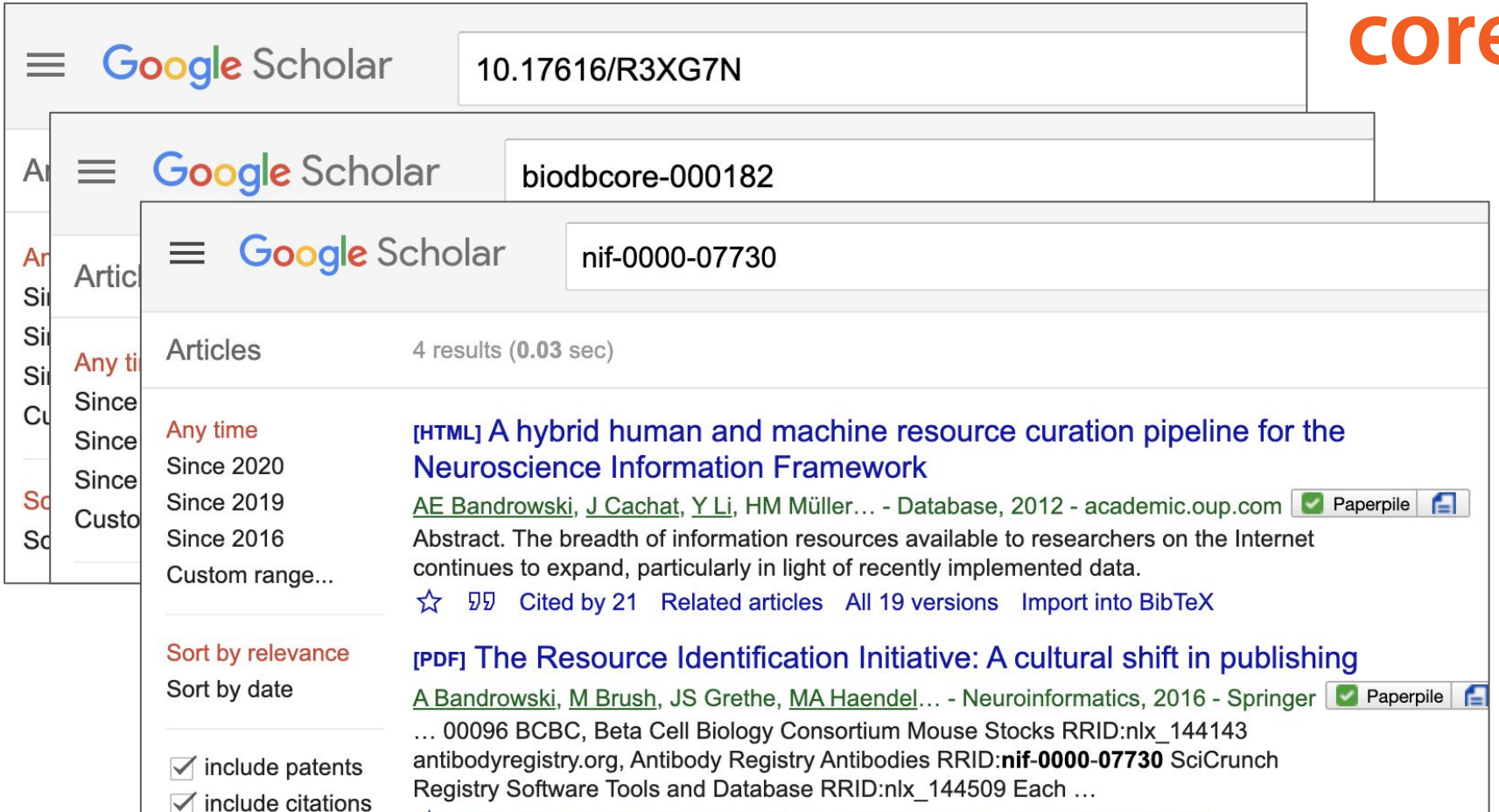
Information: Who cited my repository?

re3data coref



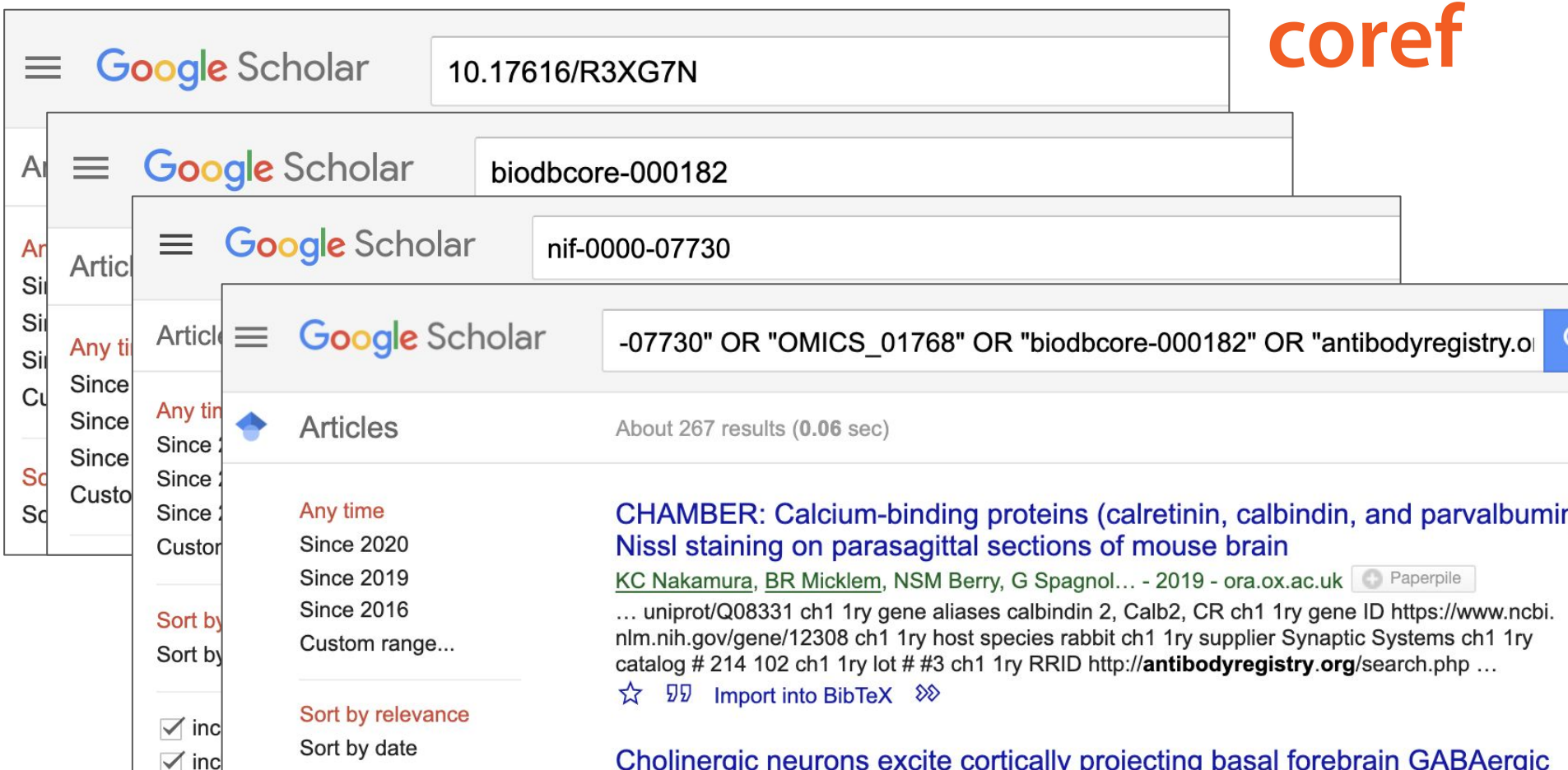
Information: Who cited my repository?

re3data coref



Information: Who cited my repository?

re3data coref



Who cited my repository?

identifiers.org/RRID/RRID:SCR_006397 or n2t.net/RRID:SCR_006397

re3data coref

Resource Name

Antibody Registry

RRID:SCR_006397 Login to claim ownership

Alternate IDs: nif-0000-07730, OMICS_01768, biodbcore-000182

Resource Information

URL: http://antibodyregistry.org/

Proper Citation: Antibody Registry (RRID:SCR_006397)

Description: Public registry of antibodies with unique identifiers for commercial and non-comm
The registry contains antibody product information organized according to genes, species, reag
formats so that authors of biological papers, text mining tools and funding agencies can quickly
allows any user to submit a new antibody or set of antibodies to the registry via a web form, or

Resource Type: Resource, service resource, data or information resource, data repository, sto

Keywords: antibody, reagent

Parent Organization: Neuroscience Information Framework

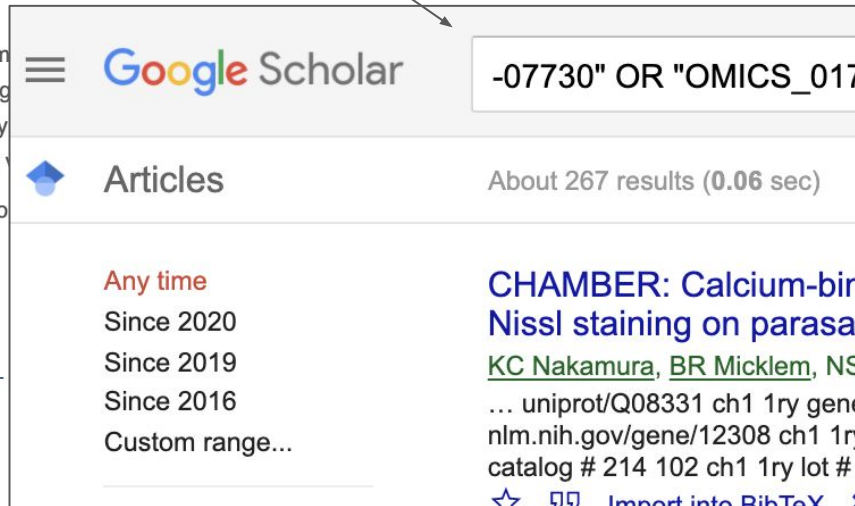
Expand All

Usage and Citation Metrics

We found 69 mentions in open access literature.

View full usage report

Check Google Scholar for all resource mentions.



How to map all repositories?
Data Export from re3data
Alignment of re3data

re3data coref

How would you prefer to access repository descriptions in re3data (e.g., API, dashboards, website, subsets based on (which) criteria, spreadsheet export ...)?]

- We would like to reconcile the entries in re3data with SciCrunch Registry
- We are currently doing this with NITRC (done), Elixr bio.tools (not done) & Debian for software tools (not done)
- We are also aligning to ror.org for institutions (not done)
- Is there some method (open refine?) that can be used to help us align the data? Is this of interest to re3Data?

Keeping in Synch
Data Import to re3data

re3data coref

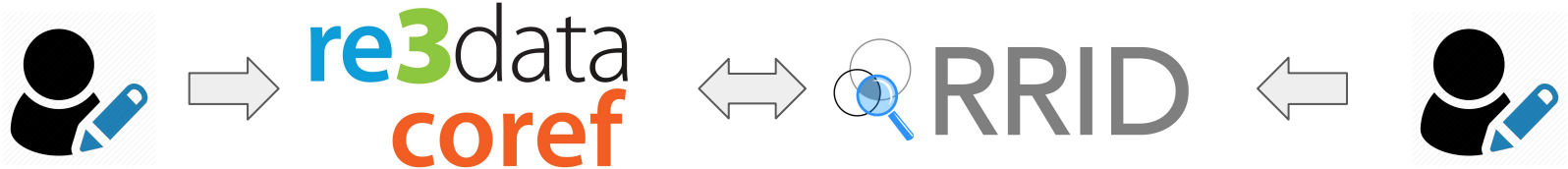
- **Once a mapping is completed, how do we keep systems in synch?**

Monitoring

How do you use re3data to monitor the repository landscape?
How could re3data improve to support your analysis?

- SciCrunch registry entries are created primarily by authors who need to cite a tool or by special interest projects.
- If an entry exists, authors can find it and cite it, registration is a more difficult step.
- ~3 RRIDs are added per day, this is not all for databases (many are software/cores)

Proposed workflow



submit: metadata package
return: identifier

Reference

re3data coref

Do you use re3data to persistently refer to repositories?

Yes, we use RRIDs

How could the reference to repositories be improved?

RRIDs could be listed for all repositories

How would you use a PID for repositories?

https://www.rrids.org/journals

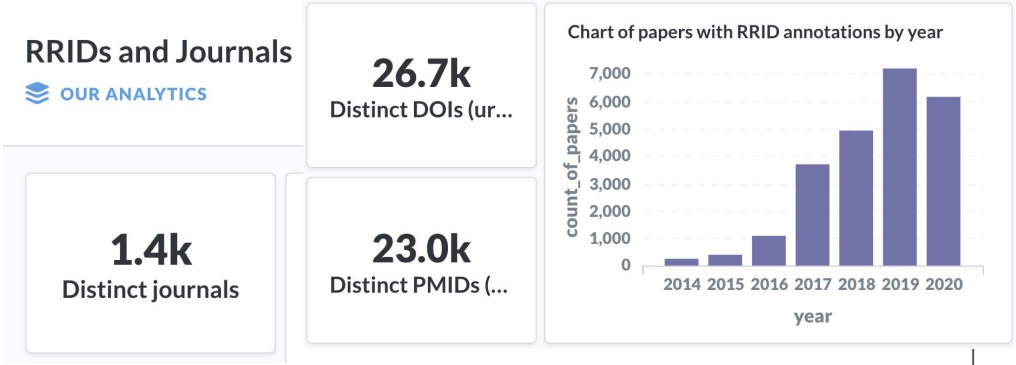
Research Resource IDentification

Endorsements

JOURNALS

Journals that ask for RRIDs

- Cell Reports
- eLife
- Cell
- Journal of Comparative Neurology
- Immunity
- Neuron
- The Journal of Neuroscience
- Molecular Cell
- Developmental Cell
- Cell Stem Cell
- Cancer Cell
- Cell Metabolism



How do you get compliance from authors?

re3data coref

Sentences that require RRIDs are detected
* RRIDs are suggested
* RRIDs are checked

Rigor criteria are checked

STAR Methods MDAR* created

MDAR asks for RRIDs & data

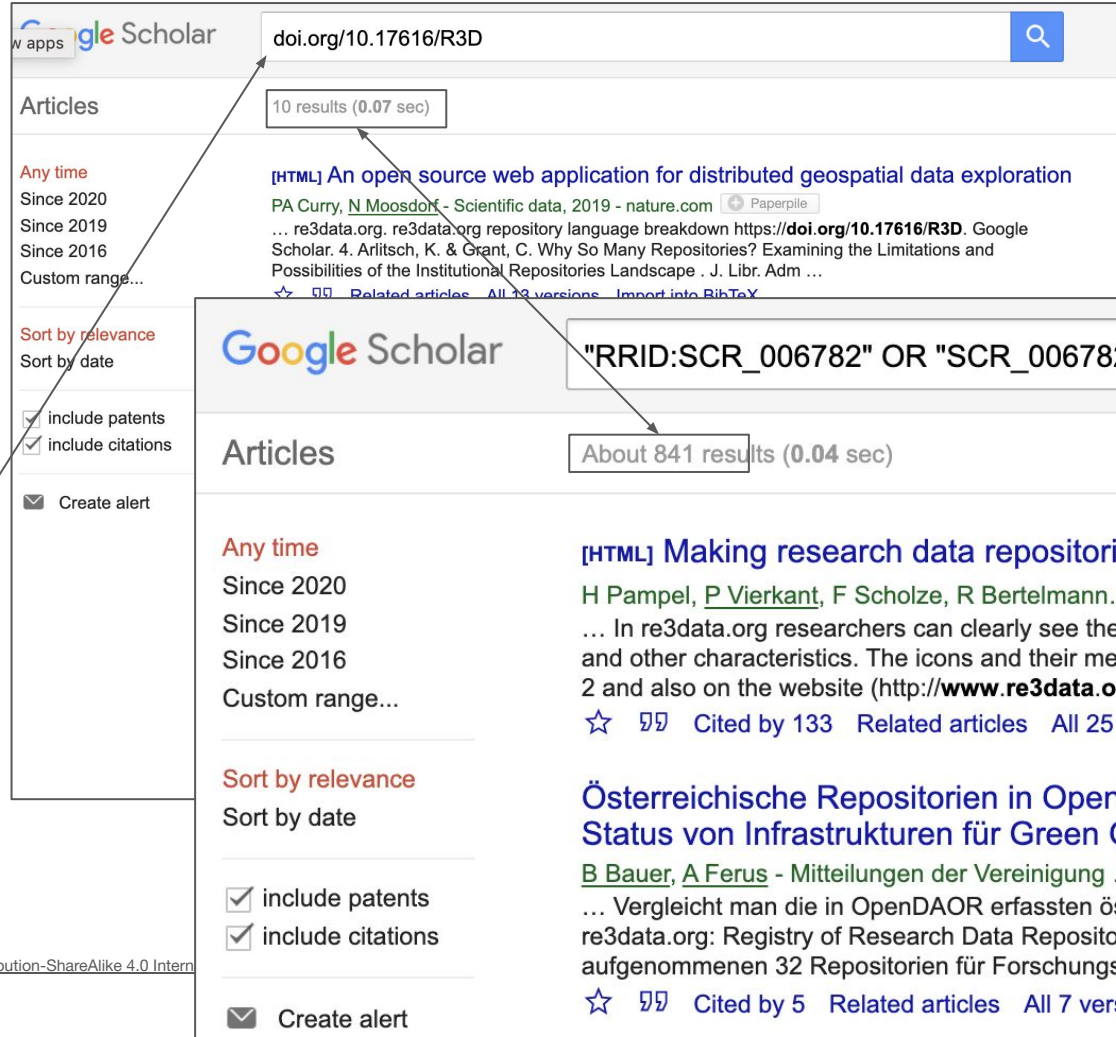
Thank You!

Tell us your use case story:
info@re3data.org

www.re3data.org

@re3data

doi.org/10.17616/R3D



re3data coref

re3data Use Case Analysis: Springer Nature

Varsha Khodiyar, (https://orcid.org/0000-0002-2743-6918), 24 November 2020

DataCite FIND, ACCESS, AND REUSE DATA

HELMHOLTZ Open Science

KIT Karlsruher Institut für Technologie

DFG

Introduction

re3data coref

For over 175 years Springer Nature has been advancing discovery by providing the best possible service to the whole research community. We help researchers uncover new ideas, support librarians and institutions with innovations in technology and data, and provide quality publishing support to societies.

Varsha Khodiyar is responsible for curating and maintaining the Springer Nature recommended repository list, and leading the team of data experts delivering Springer Nature's Research Data Support service. Varsha also contributes to the design, development and delivery of the Nature Research Academy data training workshops. She is an Executive Advisor of FAIRsharing.org, a member of CODATA's International Data Policy committee, programme chair for the Better Research through Better Data conference series and co-author of the TRUST principles.

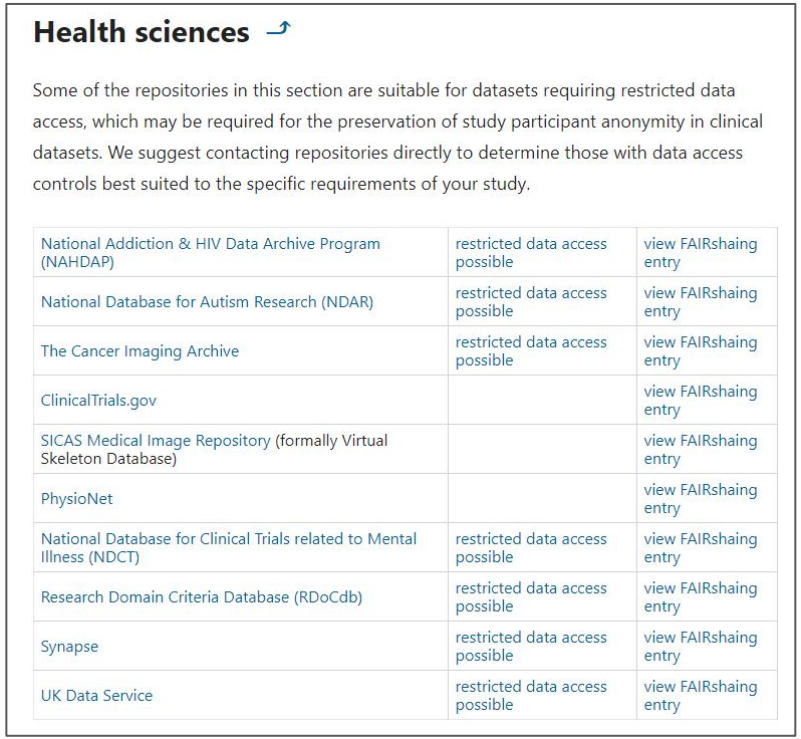
Information

re3data coref

What information on research data repositories do you generally need at your organization?

What information on research data repositories is currently missing (i.e., in general / for your organization) in re3data?

Requirements:
Curated index of repositories
Key information about each repository to be verified



Health sciences

Some of the repositories in this section are suitable for datasets requiring restricted data access, which may be required for the preservation of study participant anonymity in clinical datasets. We suggest contacting repositories directly to determine those with data access controls best suited to the specific requirements of your study.

National Addiction & HIV Data Archive Program (NAHDAP)	restricted data access possible	view FAIRshaing entry
National Database for Autism Research (NDAR)	restricted data access possible	view FAIRshaing entry
The Cancer Imaging Archive	restricted data access possible	view FAIRshaing entry
ClinicalTrials.gov		view FAIRshaing entry
SICAS Medical Image Repository (formally Virtual Skeleton Database)		view FAIRshaing entry
PhysioNet		view FAIRshaing entry
National Database for Clinical Trials related to Mental Illness (NDCT)	restricted data access possible	view FAIRshaing entry
Research Domain Criteria Database (RDoCdb)	restricted data access possible	view FAIRshaing entry
Synapse	restricted data access possible	view FAIRshaing entry
UK Data Service	restricted data access possible	view FAIRshaing entry

Data Export from re3data

re3data coref

We currently link out to re3data from our online repository lists

Archaeology Data Service	view re3data entry
Harvard Dataverse	view re3data entry
openICPSR	view re3data entry
Open Science Framework	view FAIRsharing entry
Qualitative Data Repository	view FAIRsharing entry
UK Data Service	view re3data entry

Recommended Repositories

In general, data should be submitted to discipline-specific, community-recognised repository where possible, or to generalist repositories if no suitable community resource is available.

This list is derived from *Scientific Data*'s recommended repository list. The list is also available in figshare for free reuse, with attribution, by others. *Scientific Data* is an open access data journal published by Springer Nature.

Authors should consult journal information for authors in case of more specific repository recommendations. Some Springer Nature journals, such as *Human Genome Variation*, maintain their own research data repositories.

If an author's preferred repository is not listed below we encourage repository managers to investigate listing their repositories with re3data.org and/or FAIRsharing.org in the first instance and, for those that meet the criteria, applying for listing with *Scientific Data*.

Verification of repository names would be helpful

Open Access | Published: 22 May 2018

What's in a name?

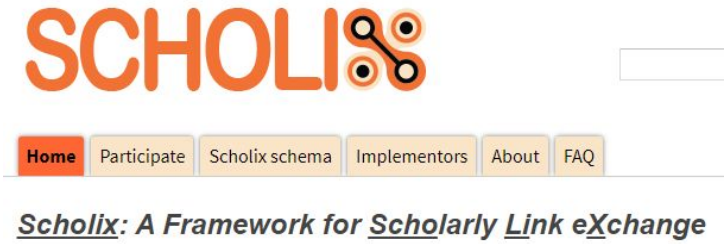
Scientific Data **5**, Article number: 180092 (2018) | Cite this article

581 Accesses | **1** Citations | **25** Altmetric | Metrics

Data Import to re3data

re3data coref

Citations to datasets as referenced in the peer-reviewed literature



Monitoring

We do not use re3data to wholescale monitor the repository landscape, as we feel our role is primarily to support our authors in identifying the most suitable repository for their specific needs based on journal policies.

It would however be useful to have certainty that if a repository appears in re3data, it can be considered to be a bona fide data repository and that this has been verified. It would be useful to understand how often re3data records are checked for deprecated repositories.

Reference

[Do you use re3data to persistently refer to repositories?
If yes, how (e.g., repository name, re3data PID, etc.)?
How could the reference to repositories be improved?
How would you use a PID for repositories?]

We link out to re3data records within the repository lists. As long as these URLs are stable, we do not have any requirements for a PID to repository indexing records.

Recommendation

re3data coref

Repository features that Springer Nature would like to know about a repository:

- Does the repository use standard open licences (e.g. Creative Commons)?
- Does it have terms of use which meet commonly understood definitions of open data (i.e. unrestricted reuse and redistribution in both academic and commercial settings)?
- Has it implemented the use of standard PIDs, e.g. DataCite/JALC/ISTIC DOIs or accession IDs?
- What are the conditions for data deposition?
- What are the conditions for data access? Do these apply equally to all researchers regardless of geographical location, or nationality?
- Does the repository have a scalable peer reviewer access mechanism for embargoed data?
- Does it have the capability to mediate controlled data access for sensitive data?
- Is there a suitable data preservation plan in place, for when the repository is no longer active?
- Does the repository have a data preservation plan in place, including implementation of a tombstone page for data which are no longer available?

Other Use Cases

re3data coref

Having two competing repository indexing services (re3data and FAIRsharing) is detrimental to the shared ideal of widespread data sharing. It would be best for the wider research community if repository records were mirrored across repository indices to avoid conflicting information being presented to researchers, authors and editors.

Ideally efforts would be focused on the verification of the information contained in the indices. We would prefer to be able to access reliable information about each repository, which would then negate the need for publishers to carry out their own evaluation processes.

Prioritization of Use Cases

re3data coref

Which of the use cases mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.Verification of repository attributes, especially repository names
2.The ability for publishers to select a subset of repositories which meet their particular policies
3.
4.
5.

Prioritization of Service Gaps

re3data coref

Which of the service gaps mentioned above are among the five most important to your organization (roughly ranked; 1 being most important to 5 being least important)?
1.Competing indexing services
2.
3.
4.
5.

Thank You!

Tell us your use case story:

info@re3data.org

www.re3data.org

@re3data

doi.org/10.17616/R3D

re3data Use Case Analysis

ZB MED's Repository Finder

Justine Vandendorpe (https://orcid.org/0000-0002-9421-8582), 24 November 2020

ZB MED Informationszentrum Lebenswissenschaften

Introduction

- **Institution:** ZB MED – Information Centre for Life Sciences (Germany)

An infrastructure and research centre for life sciences information and data.

- **Team:** Research Data Management

A team that offers dedicated services and tools along the research data life cycle, such as the Repository Finder.

- **Role:** Research Assistant

A role that includes various tasks such as the automation of the actualisation of the Repository Finder.

Information on research data repositories

- **What information we generally need at ZB MED:**
  - **Subject:** Life Sciences
  - **Type of access to data:** open
  - **Repository size:** > 50 datasets
  - **Repository types:** not institutional or national
  - **Data upload restriction type(s):** registration at most
- **What information is currently missing for ZB MED in re3data:**
  - **Date** on which the repository was uploaded or updated on re3data
  - Whether it is a **proper repository**, or just a database where authors cannot publish their data

Data Export from re3data

- **How we currently access repository descriptions in re3data:**
  - Manually
  - Web scraping
- **How we would prefer to access repository descriptions in re3data:**
  - API

Reference

We use re3data to persistently refer to repositories in ZB MED‘s Repository Finder:

Last updated: 12/21/2018

	Select category	Select category	Select category	Select category	Select category
Name	*Subject area focus in the life sciences*	*Further subject area*	*Regionally limited content*	*Operator’s location*	*Data publication only in connection with an article?*
1000 Functional Connectomes Project	Neurosciences	Medicine	not limited	USA	no
AceView	Biology	none	not limited	USA	no

Recommendation

- **What kind of information (i.e., metadata fields) we use to select recommended repositories:**
  - Subject
  - Type of access to data
  - Repository size
  - Repository types
  - Data upload restriction type(s)

For further information I am at your disposal

Justine Vandendorpe

Research Assistant

ZB MED

Gleueler Straße 60
50931 Köln

vandendorpe@zbmed.de

+49 (0) 221 478-7120

Thank you!

Informationszentrum Lebenswissenschaften

ZB MED

www.zbmed.de