



**FAIRSFair**  
Fostering Fair Data Practices in Europe

Project Title	Fostering FAIR Data Practices in Europe
Project Acronym	FAIRsFAIR
Grant Agreement No	831558
Instrument	H2020-INFRAEOSC-2018-4
Topic	INFRAEOSC-05-2018-2019 Support to the EOSC Governance
Start Date of Project	1st March 2019
Duration of Project	36 months
Project Website	<a href="http://www.fairsfair.eu">www.fairsfair.eu</a>

## D4.7 Tools for finding and selecting certified repositories for researchers and other stakeholders

Work Package	WP4 - Tools for finding and selecting certified repositories for researchers and other stakeholders
Lead Author (Org)	Sarala Wimalaratne (DataCite), Robert Ulrich (Karlsruhe Institute of Technology)
Contributing Author(s) (Org)	Kristian Garza (DataCite), Patricia Herterich (DCC), Benjamin Mathers (UKDS), Joseph Rhoads (DataCite), Margarita Trofimenko (Karlsruhe Institute of Technology), Maaïke Verburg (DANS)
Due Date	28.02.2022

Date	21.02.2022
Version	1.0
DOI	<a href="https://doi.org/10.5281/zenodo.6090418">https://doi.org/10.5281/zenodo.6090418</a>

#### Dissemination Level

<input checked="" type="checkbox"/>	PU: Public
<input type="checkbox"/>	PP: Restricted to other programme participants (including the Commission)
<input type="checkbox"/>	RE: Restricted to a group specified by the consortium (including the Commission)
<input type="checkbox"/>	CO: Confidential, only for members of the consortium (including the Commission)

### Versioning and contribution history

Version	Date	Authors	Notes
0.5	25.01.2022	Sarala Wimalaratne, Robert Ulrich, Joseph Rhoads, Kristianz Garza, Margarita Trofimenko	First draft to share with WP4
0.6	29.01.2022	Patricia Herterich, Benjamin Mathers, Maaïke Verburg	Remarks from WP4
0.8	30.01.2022	Robert Ulrich, Sarala Wimalaratne	Weaved feedback from WP4 into the document / Draft for internal review
1.0	21.02.2022	Sarala Wimalaratne, Robert Ulrich, Maaïke Verburg	Weaved internal feedback into the document / Content ready

### Disclaimer

FAIRSFAR has received funding from the European Commission's Horizon 2020 research and innovation programme under the Grant Agreement no. 831558. The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.

## Abbreviations and Acronyms

CTS	CoreTrustSeal
FAIR	Findable, Accessible, Interoperable, Reusable
IRI	Internationalized Resource Identifier
PID	Persistent Identifier
RDR	Research Data Repository

## Executive Summary

With the goal to advance the discovery of FAIR-enabling repositories, the Repository Finder had been the designated tool to improve within FAIRSF AIR. Being a limited simple view on the re3data data source, it has been reshaped and merged into DataCite Commons to embed repository metadata within the PID-Graph. Thus the joint services address researchers looking for repositories and research data as well infrastructure and service providers that require joint information on research data and its context. So the DataCite Commons integration not only makes FAIR enabling repositories discoverable, but also serves as basis for to develop services, e.g. F-UJI.re3data had been updating its metadata schema and interfaces to accompany the transition and allow to group FAIR-enabling repositories as it has been codified implicitly in the Repository Finder. DataCite Commons allows users to view and explore the context of repositories, e.g. links to the data, the researchers, the organizations etc. It also provides a foundation for future services building upon the linked content. Those services are accompanied by outreach and sustainability measures to keep and advance the work beyond the project period of FAIRSF AIR.

## Table of contents

<b>Executive Summary</b>	<b>4</b>
<b>1. Background</b>	<b>6</b>
1.1. Re3data	6
1.2. Repository Finder	6
<b>1.3. DataCite Commons</b>	<b>7</b>
1.4. Iterative development of Repository Search features	8
<b>2. Repositories Search in DataCite Commons</b>	<b>8</b>
2.1. Repository Search	9
2.2. Repository Search Results	9
2.3. Repository Page	10
2.4. Connections	11
2.5. Access to additional features	12
<b>3. re3data integration</b>	<b>13</b>
3.1. Datasource for repository descriptions	13
3.2. Metadata schema updates	13
3.2.1. Certification information	14
3.2.2. API description	15
3.2.3. Community profile property	16
3.3. API	17
3.4. Editorial Board support	18
<b>4. Future work and sustainability</b>	<b>18</b>
4.1. Next steps for DataCite Commons and Re3data Integrations	18
4.2. FAIR Data Point	18
4.3. Sustainability commitment - DataCite and re3data partnership	19
<b>5. Bibliography</b>	<b>20</b>

## 1. Background

### 1.1. Re3data

re3data<sup>1</sup> is a global registry service that indexes data repositories from all academic disciplines (over 2780 in January 2022); it offers a proven solution for identifying Research Data Repositories that provide a number of different repository and/or data-related services. At present, re3data allows users to identify RDRs available for the storage of, and access to data sets for researchers; funding bodies; publishers; and scholarly institutions. It also promotes a culture of sharing, increased access to, and better visibility of research data. re3data is a partner service of DataCite. The provision of detailed information on RDRs is a growing need within the scientific/research community which fuels the need for the further enhancement of the respective metadata descriptions and services provided by re3data in order to reflect the progress and developments within the research data repository landscape, for example reference metadata standards used in a repository, link to supported data types described in other registers etc. As part of the larger efforts to increase adoption of the FAIR principles, DataCite and re3data contribute by providing expertise, alongside their continued exploration of possible options to enhance their services with the aim of addressing these evolving needs.

### 1.2. Repository Finder

Repository Finder<sup>2</sup> was developed to support researchers by providing a search tool that would allow researchers to find repositories that they can deposit their data in. The tool was first developed in the *Enabling FAIR Data project*<sup>3</sup>, led by the American Geographical Union, which was subsequently extended as part of the first iteration of the *FAIRsFAIR* work<sup>4</sup>. It relied on re3data<sup>5</sup> as a data source and provided a simplified interface with predefined recommended filters to look up repositories (Figure 1).

---

<sup>1</sup> Registry for Research Data Repositories, <https://re3data.org>

<sup>2</sup> <https://repositoryfinder.datacite.org/>

<sup>3</sup> <http://www.copdess.org/enabling-fair-data-project/>

<sup>4</sup> [https://datacite.org/assets/RepositoryFinder\\_OpenHours\\_May2020public.pdf](https://datacite.org/assets/RepositoryFinder_OpenHours_May2020public.pdf)

<sup>5</sup> <https://www.re3data.org/>

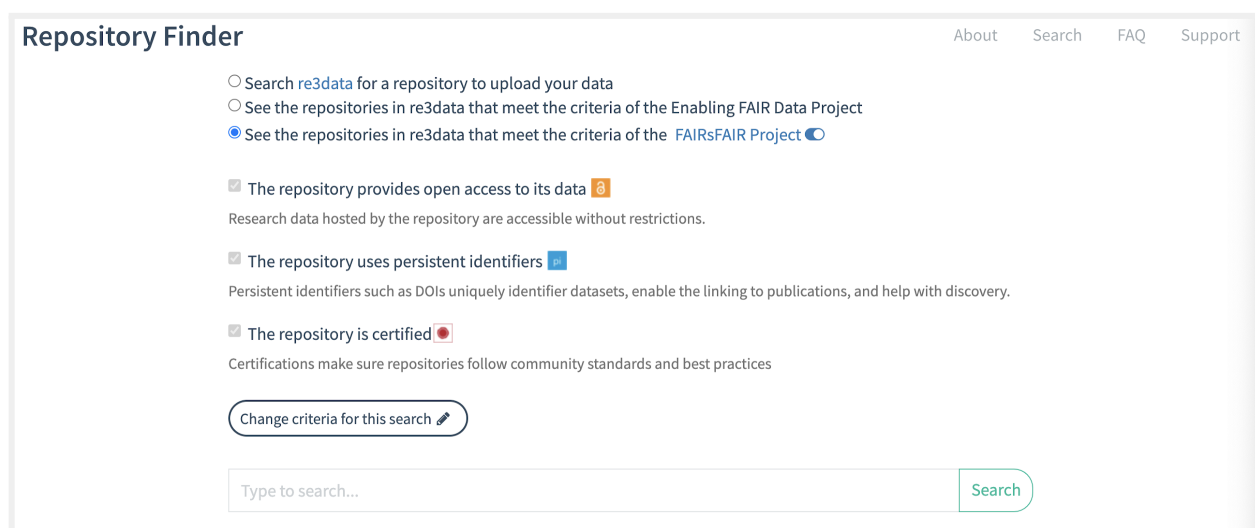


Figure 1: Repository Finder Tool

After the first iteration of development, user feedback and discussions were carried out with *Enabling FAIR Data* and *FAIRSFAR* participants to establish the strategy for the future of the Repository Finder tool. The outcome of this consultation indicated that the *Repository Finder tool* needed to provide a more integrated system, with links to relevant research outputs (e.g. publications, datasets, software), people and organizations in order to be of benefit to users. This led to the transition of the *Repository Finder* tool into *DataCite Commons*<sup>6</sup>, which is DataCite’s integrated discovery service for persistent identifiers (PIDs).

### 1.3. DataCite Commons

*DataCite Commons* is a discovery service that enables simple searches while giving users a comprehensive overview of connections between entities (through the use of PIDs) in the research landscape. Users can explore connections between research outputs (DOI<sup>7</sup>), people (ORCID<sup>8</sup>) and organizations (ROR<sup>9</sup>). This connected graph is referred to as a PID graph<sup>10</sup>, which was first developed as part of the FREYA project<sup>11</sup>, funded by the European Commission. At present, DataCite commons

<sup>6</sup> <https://commons.datacite.org/>

<sup>7</sup> <https://www.doi.org/>

<sup>8</sup> <https://orcid.org/>

<sup>9</sup> <https://ror.org/>

<sup>10</sup> <https://www.project-freya.eu/en/blogs/blogs/the-pid-graph>

<sup>11</sup> <https://www.project-freya.eu/en/front-page>

contains all DataCite DOIs, ORCIDs and RORs, and a subsection of Crossref<sup>12</sup> DOIs (which were imported as part of the FREYA grant).

*GraphQL*<sup>13</sup> is used as the underlying technology to support the DataCite PID graph. The DataCite GraphQL API<sup>14</sup> supports queries of the *DataCite API*<sup>15</sup> using the GraphQL query language<sup>16</sup>. The frontend web application of DataCite Commons is developed in React<sup>17</sup>.

## 1.4. Iterative development of Repository Search features

Development of the ‘Repository Search’ function in Commons followed the typical DataCite Product Development journey<sup>18</sup>. Following the first iteration of the development of *Repository Finder*, feedback was gathered from the various stakeholders to identify use cases that could be taken forward based on the user stories provided<sup>19</sup>. DataCite put together a set of wireframes to illustrate the user interfaces, which were validated during the bi-weekly WP4 meetings. re3data gathered feedback from its stakeholder survey and workshop<sup>20</sup> and request for comments on the upcoming changes resulting in a schema update better fitting the needs of the research community. With the decision to deprecate the Repository finder, but extend DataCite Commons to query and display repository metadata from re3data, the underlying knowledge graph enables the integration and discovery of further connections between research artifacts and directly contributes to findability as well as interoperability. Thus it is providing additional value on top of the discovery function of the original finder service.

## 2. Repositories Search in DataCite Commons

DataCite members create DOIs within their managed repositories. DataCite currently stores over 2,500 repositories with links to *re3data* where they exist. As part of *FAIRSFAR*, we have extended DataCite Commons to support the features listed below.

---

<sup>12</sup> <https://www.crossref.org/>

<sup>13</sup> <https://graphql.org/>

<sup>14</sup> <https://support.datacite.org/docs/datacite-graphql-api-guide>

<sup>15</sup> <https://support.datacite.org/docs/api>

<sup>16</sup> <https://spec.graphql.org/October2021/>

<sup>17</sup> <https://reactjs.org/>

<sup>18</sup> <https://blog.datacite.org/our-community-driven-approach-and-recent-team-changes/>

<sup>19</sup> <https://doi.org/10.5281/zenodo.4590298>

<sup>20</sup> <https://doi.org/10.48440/re3.013>



## 2.1. Repository Search

DataCite Commons was extended to include a ‘Repositories’ tab to enable users to search for repositories (Figure 2). Users are now able to search for repositories that are DataCite members.

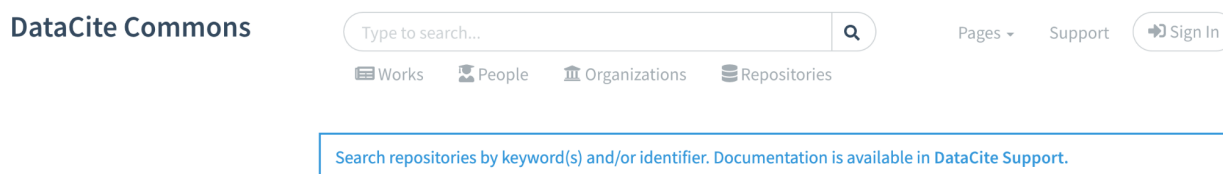


Figure 2: DataCite Commons search landing page<sup>21</sup>

## 2.2. Repository Search Results

The ‘Repository Search’ results page lists the repository names. In addition to the repository name, if a repository is linked to re3data profile, the results page will list the description(s), language(s) and the link(s) to the repository’s re3data landing page, alongside a link to repository’s website (Figure 3).

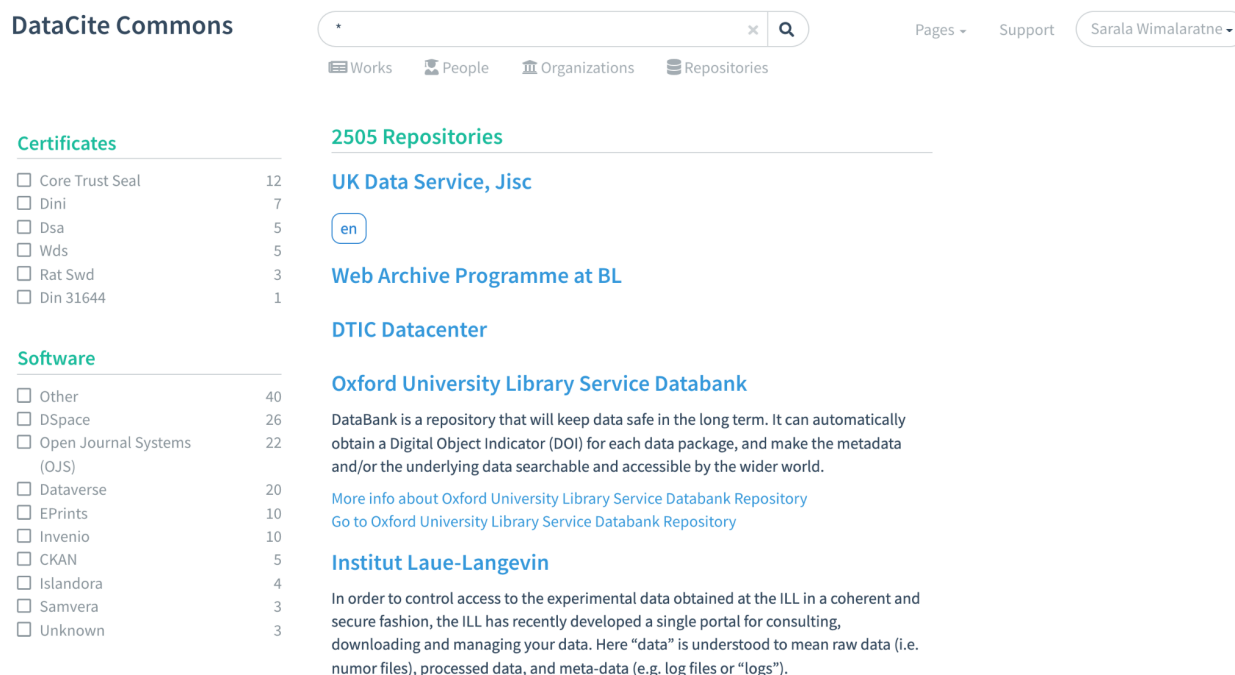


Figure 3: Repository search results<sup>22</sup>

<sup>21</sup> <https://commons.datacite.org/>

<sup>22</sup> [https://commons.datacite.org/repositories?query=\\*](https://commons.datacite.org/repositories?query=*)

The filters on the search results page allows users to filter results by Certification and Software. Information on certification and the software used for the repository are indicators indexed to identify trustworthy repositories. The data source for the filter information is re3data metadata. For example: researchers can filter by 'Core Trust Seal' to find certified repositories to deposit their data. Other certificates can also be selected, and the system is flexible enough to add potential new certificates that the community defines in the future by adding them to the re3data metadata.

### 2.3. Repository Page

Selecting a specific Repository from the search results takes the user to the Repository's page in DataCite Commons, which provides the user with further information about the repository including repository content. This information is pulled from the DataCite metadata store<sup>23</sup> and *re3data*. A repository's page on DataCite Commons also provides aggregated statistics concerning the Datasets stored within the repository that have DataCite DOI's, including the number of registered DataCite DOIs, citations<sup>24</sup>, views and downloads<sup>25</sup> where the information is available. Further aggregated statistics are displayed using charts which provide an overview of the following: year of publication, deposit type; top depositors; field of science, language; and license type(s). A link is also provided on the page to access the source repository homepage. The 'Find related works' button enables users to drill down into the DOIs registered in the DataCite metadata store (see section 2.4 for more information).

---

<sup>23</sup> <https://support.datacite.org/docs/mds-api-guide>

<sup>24</sup> <https://support.datacite.org/docs/contributing-data-citations>

<sup>25</sup> <https://support.datacite.org/docs/views-and-downloads>

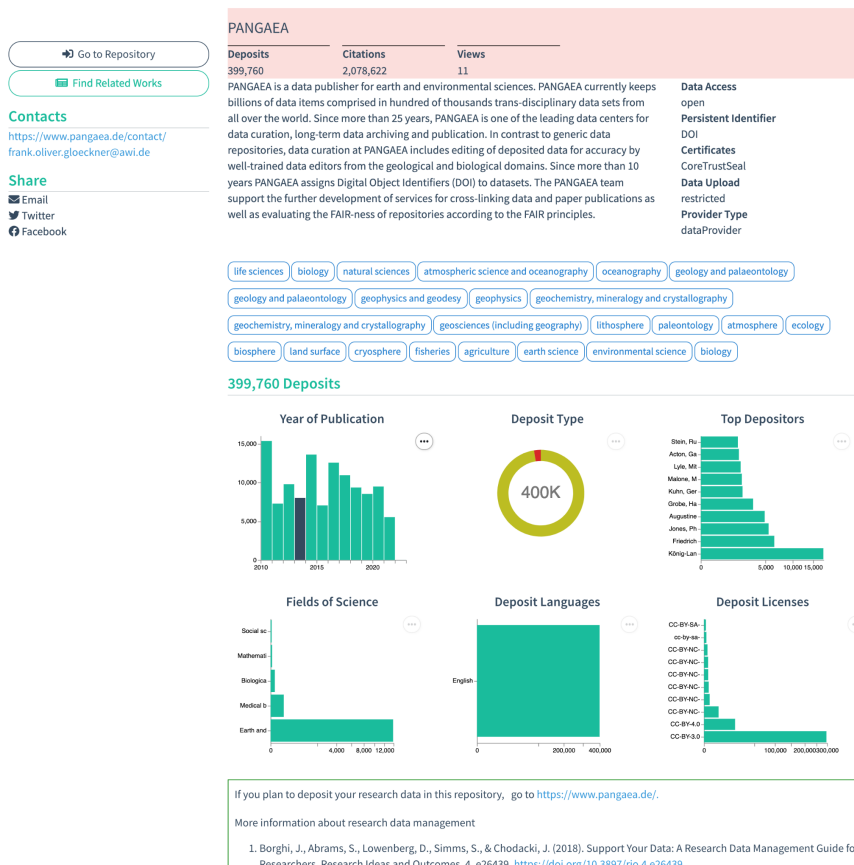


Figure 4: Repository search page<sup>26</sup>

## 2.4. Connections

One of the main benefits of transitioning Repository Finder functionality to DataCite Commons is that it provides navigational support for the PID graph in identifying connected resources. The ‘Related works’ page for a specific repository lists all the content registered within DataCite. The filters on the ‘Related works’ page enables users to search for content by: Publication Year; Work Type; License; Language, Field of Science; Authors etc. This provides researchers with the option to obtain further information about the existing content stored within a specific repository, which can help inform their decision as to whether they wish to deposit their own data within said repository.

<sup>26</sup> <https://commons.datacite.org/repositories/pangaea.repository>

142,428 Works

### Publication Year

<input type="checkbox"/> 2022	61
<input type="checkbox"/> 2021	5,509
<input type="checkbox"/> 2020	7,321
<input type="checkbox"/> 2019	21,384
<input type="checkbox"/> 2018	21,662
<input type="checkbox"/> 2017	20,355
<input type="checkbox"/> 2016	18,891
<input type="checkbox"/> 2015	17,735
<input type="checkbox"/> 2014	11,346
<input type="checkbox"/> 2013	8,724
<input type="checkbox"/> 2012	5,059
<input type="checkbox"/> 2011	2,707
<input type="checkbox"/> 2010	720

### Work Type

<input type="checkbox"/> Dataset	43,727
<input type="checkbox"/> Text	5
<input type="checkbox"/> Collection	4
<input type="checkbox"/> Other	3
<input type="checkbox"/> Image	2
<input type="checkbox"/> Software	1

### License

<input type="checkbox"/> CC0-1.0	43,741
<input type="checkbox"/> CC-BY-3.0	1

### Language

<input type="checkbox"/> English	43,803
----------------------------------	--------

### Field of Science

<input type="checkbox"/> Biological sciences	2,930
<input type="checkbox"/> Natural sciences	335
<input type="checkbox"/> Earth and related environmental sciences	292
<input type="checkbox"/> Sociology	236
<input type="checkbox"/> Computer and information sciences	140

## Data from: Effects of inversions on within- and between-species recombination and divergence

Laurie S. Stevison, Mohamed A. F. Noor & Kenneth B. Hoehn  
Version 1 of Dataset published 2011 in DRYAD

Chromosomal inversions disrupt recombination in heterozygotes by both reducing crossing over within inverted regions and increasing it elsewhere in the genome. The reduction of recombination in inverted regions facilitates the maintenance of hybridizing species, as outlined by various models of chromosomal speciation. We present a comprehensive comparison of the effects of inversions on recombination rates and on nucleotide divergence. Within an inversion differentiating *Drosophila pseudoobscura* and *D. persimilis*, we detected one double-recombinant among 9739 progeny from F1 hybrids screened, consistent with published double crossover frequencies observed within species. Despite similar rates of exchange within and between species, we found no sequence-based evidence of ongoing gene exchange between species within this inversion, but significant exchange was inferred within species. We also observed greater differentiation at regions near inversion breakpoints between species vs. within species. Moreover, we observed strong 'interchromosomal effect' (higher recombination in inversion heterozygotes between species) with up to 9-fold higher recombination rates along collinear segments of chromosome two in hybrids. Further, we observed that regions most susceptible to changes in recombination rates corresponded to regions with lower recombination rates in homokaryotypes. Finally, we showed that interspecies nucleotide divergence is lower in regions with greater increases in recombination rate, potentially resulting from greater interspecies exchange. Overall, we have identified several similarities and differences between inversions segregating within vs. between species in their effects on recombination and divergence. We conclude that these differences are most likely due to lower frequency of heterokaryotypes and to fitness consequences from the accumulation of various incompatibilities between species. Additionally, we have identified possible effects of inversions on interspecies gene exchange that had not been considered previously.

DOI registered July 5, 2011 via DataCite.



1 Citation 178 Views 79 Downloads

[Dataset](#) [English](#)

<https://doi.org/10.5061/dryad.7q0nq>

### Scripts

Kenneth B. Hoehn & Kenneth B. Hoehn  
Data File published 2011 in

Scripts used to calculate average pairwise differences between all 6 pairwise combinations along the XR chromosome between *Drosophila pseudoobscura*, *Drosophila persimilis*, Sex-Ratio *Drosophila persimilis*, and *Drosophila miranda*. These output file can then be used to calculate RND between *D pseudoobscura*, *D persimilis*, and *SR D persimilis* in another program such as Microsoft Excel. All scripts are written in Perl. Author: Kenneth Hoehn. Date: May-5-2011. Sections: 1. Constructing the combined, raw input file from PILEUP outputs (skip if combined input file is provided); 2. Necessary input files for final analysis; 3. Editing combined input file and calculating average pairwise differences; 4. Calculating RND; 5. Double checking output; 6. Simple window writing script. See README.txt for more details.

DOI registered July 5, 2011 via DataCite.

[Dataset](#)

<https://doi.org/10.5061/dryad.7q0nq/1>

Figure 5: Repository search page<sup>27</sup>

## 2.5. Access to additional features

By embedding the 'Repository Search' feature in DataCite Commons, it enable users to access other features that are supported by DataCite Commons<sup>28</sup>, such as DataCite Researcher Profiles, ORCID claiming and Accessibility Achievements.

<sup>27</sup> <https://commons.datacite.org/doi.org?query=client.uid:dryad.dryad>

<sup>28</sup> <https://support.datacite.org/docs/datacite-commons>



### 3. re3data integration

#### 3.1. Datasource for repository descriptions

DataCite Commons grants access to the PID-Graph and dynamic exploration to the relations of the metadata records supplied by ORCID, ROR etc. Repository descriptions are served via re3data and embedded into the Common UI. Data types, licenses, contributors are aggregated and visualized, as well as users can follow links to re3data for the complete repository descriptions.

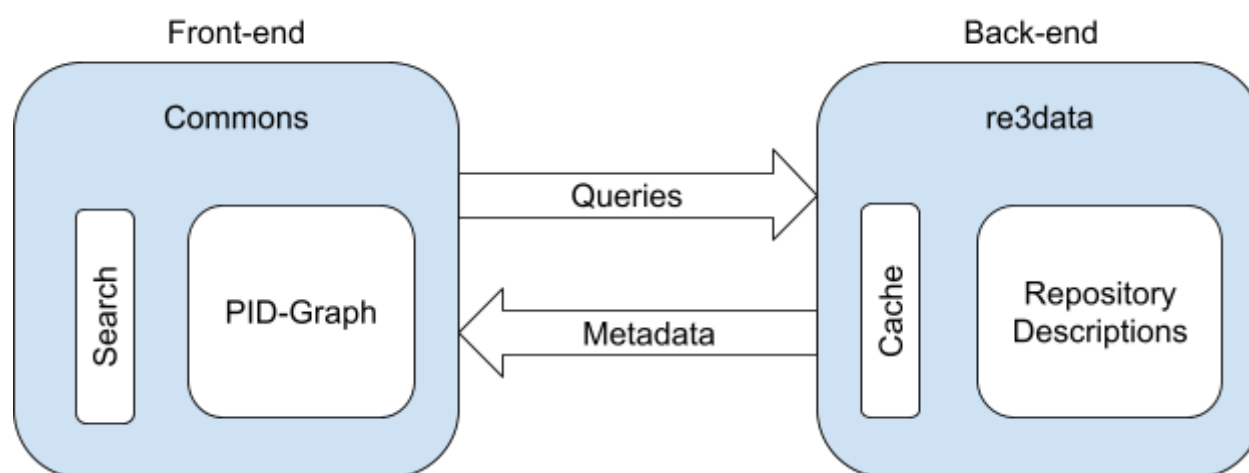


Figure 6: DataCite Commons integration of re3data

#### 3.2. Metadata schema updates

DataCite Commons integrated the repository search feature in its latest release, which enables the linking of a given entity with other related entities<sup>29</sup> in the PID graph and displays this information together with additional data from re3data where available. To integrate the Repository Finder in the resource discovery, changes in re3data had to be made in the backend and required further changes and adoptions regarding its metadata schema. The changes proposed by Milestone 4.8<sup>30</sup> have been weaved into the release of Version 3.1 of the re3data metadata schema<sup>31</sup> following discussion within FAIRsFAIR. Following these discussions, Version 3.1 was subsequently published as a Request for Comment<sup>32</sup> within the open process of schema updates in re3data. Over 150

<sup>29</sup> Research outputs, people, and organizations

<sup>30</sup> <https://doi.org/10.5281/zenodo.4590336>

<sup>31</sup> <https://www.re3data.org/schema>

<sup>32</sup> <https://coref.project.re3data.org/blog/request-for-comments-version-3-1-of-the-re3data-metadata-schema>

comments had been received so far and either been taken as input for later revisions or incorporated already.

### 3.2.1. Certification information

Based on the input and suggestions<sup>33</sup> made by Wim Hugo on behalf of CoreTrustSeal, the current description of repository certification in the re3data metadata schema has been updated with only minor changes. These changes include expiration date (single property) being replaced with the validation period (two properties) and the certificate name now being modeled with a controlled vocabulary codified into the schema (See Table 1). This allows us to add future certificates (e.g. CoreTrustSeal+FAIR certification, if implemented) and/or revise existing certificates without the need to publish a new version of the schema to make each change.

*Table 1. Certificate field in the re3data metadata schema*

#	Property	Description	O/C	Vocabulary / Values
<b>36</b>	certificate	The certificate, accreditation or standard the RDR complies with (wrapper element).	0-n	
<b>36.1</b>	certificateName	The name of the certificate.	1	Example: CoreTrustSeal
<b>36.2</b>	certificateStartDate	The start date of the certification period or the date the certificate was issued, in case certification does not expire.	0-1	YYYY or YYYY-MM-DD or any other format described in W3C-DTF (ISO-8601)
<b>36.3</b>	certificateEndDate	The end date of the certification period.	0-1	YYYY or YYYY-MM-DD or any other format described in W3C-DTF (ISO-8601)
<b>36.4</b>	certificateUrl	The URL to a source outlining details of the certification.	1	URL Example: <a href="https://doi.org/10.5281/zenodo.3638211">https://doi.org/10.5281/zenodo.3638211</a>
<b>36.5</b>	certificateWidget	The URL supporting the display of widgets/badges.	0-1	URL Example: <a href="https://www.cert.example/badge.jpg">https://www.cert.example/badge.jpg</a>

<sup>33</sup> <https://doi.org/10.48440/re3.013>

### Example XML

```

<repository>
  ...
  <certificate>
    <certificateName>CoreTrustSeal</certificateName>
    <certificateUrl>//cert.example/cert.pfd</certificateUrl>
    <certificateWidget>//cert.example/badge.jpg</certificateWidget>
    <certificateEndDate>2024-01-01</certificateEndDate>
  </certificate>
  ...
</repository>

```

### 3.2.2. API description

FAIRsFAIR Work Package 2 follows a modern linked data approach and fosters the implementation of the FAIR Data Points<sup>34</sup> within repositories, still re3data is currently not aware of RDF and cannot support DCAT for now. But as the API-Property isn't limited by a controlled vocabulary and upcoming releases of re3data will support specifying Internationalized Resource Identifiers (IRIs) to DCAT respective to its adoption as FAIR Data Point as an initial approach to solve the problem. This may enable a first support for linking the repository description with the data sets, further RDF support will be followed up after the project (see 4.2). See Table 2 below for the API properties included in Version 3.1 of the re3data metadata schema?

Table 2. API field in the re3data metadata schema

#	Property	Description	O/C	Vocabulary / Values
<b>28</b>	api	The API supported by the RDR (wrapper element).	0-n	
<b>28.1</b>	apiType	The type of the API.	1	Example: REST, SPARQL, OAI-PMH
<b>28.2</b>	apiUrl	The URL of the API for machine processing of data or metadata.	1	Example: <a href="https://datadryad.org/api/v2/">https://datadryad.org/api/v2/</a>
<b>28.3</b>	apiDocumentation	A link referring to the API documentation, a website that states its availability	1	Example: <a href="https://datadryad.org/api/v2/docs/">https://datadryad.org/api/v2/docs/</a>

<sup>34</sup> <https://www.fairdatapoint.org/>

		and other information for using the API.		
--	--	--	--	--

#### Example XML

```

<repository>
  ...
  <api>
    <apiType>FAIR Data Point</apiType>
    <apiUrl>https://api.example/catalog</apiUrl>
    <apiDocumentation>https://api.example/doc</apiDocumentation>
  </api>
  ...
</repository>

```

### 3.2.3. Community profile property

As DataCite Commons enables discovery and linkage of different entities in the PID-graph, it is desirable to preserve the current functionality of the *Repository Finder* for researchers, which provides such users with a list of repositories, based on the predefined search criteria made available to them. As a pilot, the set of FAIR-enabling repositories in the earth, space and environmental sciences had been implemented in the Repository Finder based on COPDESS community developed criteria.<sup>35</sup> Currently this subset of these FAIR-enabling repositories is not tracked in the metadata but are nevertheless still relevant to researchers looking for those repositories. With the profile property added to Version 3.1 of the re3data metadata schema, these selections can now be materialized and will be added to the search and faceting of re3data (See Table 3). In addition to making FAIR-enabling repositories identifiable, this will allow re3data to define profiles in collaboration with different research communities to provide guidance on recommended repositories.

Table 3. Profile field in the re3data metadata schema

#	Property	Description	O/C	Vocabulary / Values
37	profile	The selection of repositories based on a set of community-developed criteria (wrapper element).	0-n	

<sup>35</sup> Connecting Researchers to Data Repositories in the Earth, Space, and Environmental Sciences. [https://doi.org/10.1007/978-3-030-11226-4\\_7](https://doi.org/10.1007/978-3-030-11226-4_7)



<b>37.1</b>	profileTitle	The name of the profile the RDR is associated with.	1	Example: Enabling FAIR Data (COPDESS/AGU) WDS Regular Members
<b>37.2</b>	profileUrl	The URL to a source naming the entity creating the profile and outlining the profile / describing the selection criteria.	1	Example: <a href="http://www.copdess.org/enabling-fair-data-project/enabling-fair-data-faqs/">http://www.copdess.org/enabling-fair-data-project/enabling-fair-data-faqs/</a> <a href="https://www.worlddatasytem.org/community/membership/regular-members">https://www.worlddatasytem.org/community/membership/regular-members</a>

#### Example XML

```
<repository>
...
<profile>
  <profileTitle>FAIR-enabling</profileTitle>
  <profileUrl>fair-community.example/profile-description</profileUrl>
</profile>
...
</repository>
```

### 3.3. API

Transitioning from the Repository Finder to DataCite Commons, re3data is providing the necessary interface to integrate repository descriptions within the DataCite services. The first iteration of the repository search feature in DataCite Commons is limited to repositories within the DataCite scope, but further work will be quickly undertaken to allow for the discovery of all repositories listed in re3data.

The implementation of a lightweight API to set the certification of a repository by the certification authority is scheduled for March (after the end of the FAIRsFAIR project). This will improve the quality of the metadata records displayed in DataCite Commons and will ease the curation process for repository owners and the re3data editorial board as the metadata on certification has not to be added and checked manually then.

### 3.4. Editorial Board support

As part of the work undertaken within FAIRSFair Work Package 4, 10 repositories were selected (following their application) for support in obtaining CoreTrustSeal certification<sup>36</sup>. An additional 12 repositories were also selected for support by FAIRSFair Work Package 2 with the aim of them becoming FAIR Data Points. 3 of the repositories supported by FAIRSFair had been missing in the re3data database and had been ingested and checked by the re3data Editorial Board as well as the existing ones have been updated.

Furthermore, the links provided in DataCite Commons to re3data entries will be analyzed by the re3data Editorial Board with the aim of providing updates and extensions, thus improving data discovery and linkage in the PID-Graph. The re3data Editorial Board has started to index RORs for repository institutions which will be synced and linked with related entities. Further links enabled by the PID-Graph are being considered for implementation in future iterations of DataCite Commons.

## 4. Future work and sustainability

### 4.1. Next steps for DataCite Commons and Re3data Integrations

Currently DataCite Commons supports the searching of repositories run by DataCite members and provides links to corresponding re3data landing pages with detailed repository metadata. In the next iteration, we plan to include all repositories from re3data within DataCite Commons in a similar fashion to which ORCID and ROR records are included. This will enable users to search for repositories from DataCite members and the re3data registry. With the project start an extension of the Repository Finder had been envisioned with a simplified view on a FAIR enabling repository subset of re3data. The concept of a simple search has been discarded in favor of the transition to DataCite Commons. The DataCite PID-Graph allows for a deeper integration of re3data repository descriptions within DataCite as well as enriching those entries in re3data. These options will be explored in upcoming development cycles with feedback from the community.

### 4.2. FAIR Data Point

Following the ‘Make your repository FAIR enabling: FAIRSFair Hackathon’<sup>37</sup> (held in October 2021), which acted as an introduction to the current FAIR Data Point concept, support by the re3data service had been discussed. The current concepts, implementations and prototypes developed thus

---

<sup>36</sup> <https://doi.org/10.5281/zenodo.5113523>

<sup>37</sup> <https://www.fairsfair.eu/events/make-your-repository-fair-enabling>

far in WP2, shall be continued beyond the FAIRsFAIR Project, with the opportunities identified regarding the potential benefits that a Semantic Web/Linked Data approach provides taking center stage in any future considerations. As a first step, the current mapping<sup>38</sup> of the re3data metadata schema to RDF will be considered for an update with the aim of making the Resources accessible. The team envisage this being achieved through the use of the re3data namespace and the respective (resolvable) IRIs. Potential solutions for implementing the FAIR Data Point concepts into re3data including update mechanisms for repositories and properties will be evaluated. Those efforts need to be embedded within the broader community. This is not required for the DataCite Commons integration, but might support services within the semantic web / linked data world.

### 4.3. Sustainability commitment - DataCite and re3data partnership

Alongside the joint efforts within FAIRsFAIR, the integration of Repository Finder within DataCite Commons has led to DataCite and re3data beginning a process of renewing their collaboration agreements, alongside a revision of the of the roles and responsibilities required for a mutual service operation that were outlined in the previous collaboration agreement(s). This process has been recognized as an essential component to complement the technical development. and is a crucial element in sustaining the current efforts tasked with developing and improving existing solutions beyond the FAIRsFAIR project. Thus, any future partnership between re3data and DataCite will seek to maintain and build upon FAIRsFAIR results.

---

<sup>38</sup> <https://github.com/FAIRDataTeam/FAIRDataPoint-Spec/blob/master/spec.md>

## 5. Bibliography

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, re3data, 37 p. <https://doi.org/10.48440/re3.010>

re3data.org Schema 3.1 XML Schema Definition, <http://doi.org/10.48440/re3.011>