# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who we are</td>
<td>3</td>
</tr>
<tr>
<td>Message from the... Executive Director, Matt Buys</td>
<td>6</td>
</tr>
<tr>
<td>Message from the... Executive Board President, John Chodacki</td>
<td>7</td>
</tr>
<tr>
<td>Message from the... Executive Board Treasurer, Marco Marsella</td>
<td>8</td>
</tr>
<tr>
<td>Community</td>
<td>9</td>
</tr>
<tr>
<td>Services &amp; Technology</td>
<td>18</td>
</tr>
<tr>
<td>Funded Projects</td>
<td>20</td>
</tr>
<tr>
<td>Strategic Initiatives</td>
<td>24</td>
</tr>
<tr>
<td>Financial Overview</td>
<td>26</td>
</tr>
<tr>
<td>Executive Board</td>
<td>29</td>
</tr>
<tr>
<td>Staff</td>
<td>33</td>
</tr>
<tr>
<td>Research Publications</td>
<td>37</td>
</tr>
</tbody>
</table>
Who we are
We are a global community that shares a common interest: to ensure that research outputs and resources are openly available and connected so that their reuse can advance knowledge across and between disciplines, now and in the future.

DataCite makes research more effective by connecting research outputs and resources – from data and preprints to images and samples. We support the creation and management of DOIs and metadata records, enhance research workflows with service integration, and enable the discovery and reuse of research outputs and resources.

DataCite is mission-driven and value-led.

Connecting research, identifying knowledge.

Our mission is to be a world-leading provider of persistent identifier services to help make research outputs and resources findable, citable, connected, and reused globally. We seek to create value for our members through community-driven, innovative, open, integrated, usable, and sustainable services for research.

Our guiding values:

**Reliability:** We exist to reliably support our community and ensure our infrastructure is sustainable.

**Transparency:** We make transparent decisions and engage openly with the community in all that we do.

**Trust:** We seek to be a trusted partner for our members and other community stakeholders by delivering services that uphold community principles.

**Inclusivity:** We support a global community and value diverse perspectives.
Message from the...
Message from the...

**Executive Director, Matt Buys**

It is my pleasure to present the DataCite Annual Report for 2022. This year has been a testament to the strength and resilience of our global community, and I am proud of the work that we have accomplished together.

We are excited about the launch of the Data Citation Corpus project and the Global Access program in 2023, both of which will play a crucial role in advancing open scholarly communication infrastructure and adoption. Our commitment to financial sustainability has allowed us to continue supporting our community and making a difference in the world of research.

As a global community, we share a common interest in ensuring that research outputs and resources are openly available and connected so that their reuse can advance knowledge across and between disciplines, now and in the future. Through the creation and management of persistent identifiers, service integration, and the discovery and reuse of research outputs and resources, we aim to make research more effective and accessible.

I would like to extend my gratitude to each and every one of you for your contributions, collaboration, and support. Together, we are making a difference in the world of scholarly communication and I look forward to our continued work together in the future.

Matthew Buys  
DataCite Executive Director
Executive Board President,
John Chodacki

This year, DataCite has continued to lead the way in providing persistent identifiers (PIDs) for research outputs and resources, making it easier for researchers to find and reuse research.

We are proud to have supported the research community through our training and outreach initiatives, including providing guidance on best practices in research data management and sharing. Our partnerships with organizations like FORCE11, RDA, and others have helped us to advance the cause of open research and we will continue to work with these partners to promote the use of PIDs.

In 2022, we launched new partnerships including our support for IGSNs. This collaboration acts as a model for how DataCite can leverage our robust infrastructure to accommodate new use cases and priorities from a wide range of research domains and communities. We see this work as the first of many and invite other identifier communities to contact the staff if you too would like to collaborate with DataCite in this way.

Our members are our strength. I am pleased to see that the results of our member survey in 2022 demonstrated high levels of satisfaction with our services and the desire for further collaboration and community building. We will continue to work closely with our members to ensure their needs are met. I want to thank our members, supporters, and the wider research community for their ongoing partnership and dedication to promoting the adoption of open infrastructure. DataCite’s success would not be possible without your support.

John Chodacki
DataCite Board President
I am pleased to report that DataCite closed the 2022 fiscal year with a positive overall result, with total income being €336,556 above budget and expenses €233,003 below budget. This was due to a combination of factors, including the ongoing COVID-19 travel restrictions and membership outreach expense reductions, as well as additional expenses in the licenses budget related to automation and professionalization efforts.

I am also pleased to announce that in 2022, we welcomed 28 new members and 160 new Consortium Organizations to the DataCite community, and we have received additional funding from various sources including the EU Commission, the Alfred P. Sloan Foundation, and the German Research Foundation. Our cash balance on December 31, 2022 was €1,865,484.21, and we do not anticipate any cash management issues through December 2024.

Looking ahead to 2023, we are forecasting membership revenue of €1,863,480, with consistent new membership and consortia revenue growth. We also expect to receive additional project funding from the National Institutes of Health, the Wellcome Trust, and the Chan Zuckerberg Initiative. Our budget for 2023 includes increased administrative and licensing costs, as well as travel and infrastructure costs, to align with our growing needs and the launch of new services.

Finally, I would like to thank our members and partners for their continued support and commitment to DataCite’s mission.

Marco Marsella
DataCite Board Treasurer
Community
Membership

**DataCite members** are at the heart of our community and we were very happy to welcome 28 new members and 160 new Consortium Organizations to the DataCite community.

We were particularly pleased with the emergence of four new consortia in 2022. Two of these focus on providing service in specific countries (Panama and China), one consortium focuses on large data from observatories and the final one was started by the National Institutes of Health in the USA to enable the implementation of their data policies.

Another important development in 2022 was that several IGSN members joined the DataCite Member community. This was part of the partnership between DataCite and the International Generic Sample Number (IGSN) e. V. that was announced in 2021 to foster worldwide adoption, implementation, and utilization of persistent identifiers for material samples.

We are very pleased to report that the transitioning step in this project has now been completed. In particular, all IGSN Allocating Agents have joined the DataCite membership. At the beginning of the project, eight of the 14 Allocating Agents were Direct Members or Consortium Organizations. By the end of 2022, four Allocating Agents had joined DataCite as Direct Members, and two others joined as Consortium Organizations.

---

1 At the time of writing, one Allocating Agent is changing its status to become a Direct Member in its own right, rather than having membership through its host institute.
New Members & Consortium Organizations

New Members:

1. CSIRO
2. Consorci per a la Construcció, Equipament i Explotació del Laboratori de Llum Sincrotró (CELLS - ALBA)
3. Data Observatory (Consortium Lead)
4. Addgene
5. Castor
6. SENACYT Panama (Consortium Lead)
7. Doctors Without Borders/ Médecins Sans Frontières
8. Lithodat
9. University of Chicago Library (Member Only)
10. National Yang Ming Chiao Tung University
11. European Space Agency
12. The Allen Institute (for Immunology)
13. Korea Institute of Geoscience and Mineral Resources (KIGAM)
14. Computer Network Information Center (CNIC), Chinese Academy of Sciences
15. Université de Mons
16. Duke University Libraries
17. Arcadia Science
18. Luxembourg Institute of Science & Technology
19. Geoscience Australia
20. The National Institutes of Health (NIH) (Consortium Lead)
21. West African Science Service Center on Climate Change and Adapted Landuse (WASCAL)
22. Hugging Face
23. Instituto Nacional de Tecnología Agropecuaria
24. Fundación DiaNet
25. Pew Research Center
26. Beijing Wanfang Data Co., Ltd (Consortium Lead)
27. Woods Hole Oceanographic Institution – Ocean Observatories Initiative
28. Konsortium der sächsischen Hochschulbibliotheken c/o SLUB Dresden (Member Only)
29. NUHS
30. Université Savoie Mont Blanc
31. Universität für angewandte Kunst
32. Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen e.V.
33. Hochschule für nachhaltige Entwicklung Eberswalde
34. Science Communications World Wide
35. Fachhochschule Burgenland
36. Recherche Data Gouv France
37. New Zealand eScience Infrastructure
38. Wrocław University of Environmental and Life Sciences
39. Adam Mickiewicz University in Poznań
40. AZTI
41. NHM Vienna
42. Christian-Albrechts-Universität zu Kiel
43. Deutsches Forschungszentrum für Künstliche Intelligenz GmbH
44. Bundesamt für Seeschifffahrt und Hydrographie
54. Amt für Statistik Berlin-Brandenburg
55. British Council
56. Dauphin Island Sea Lab
57. Future Science Group
58. National Computational Infrastructure
59. Australian Catholic University
60. Department of Agriculture, Water and the Environment - ABARES
61. University of Southern Queensland
62. T-CAIREM
63. Universidad de Cartagena
64. Benemérita Universidad Autónoma de Puebla (BUAP)
65. University Autónoma de Baja California (UABC)
66. Carl von Ossietzky University of Oldenburg
67. Arolsen Archives (ITS)
68. Universidad Tecnológica de Panamá
69. Universidad Especializada de las Américas
70. Universidad Metropolitana de Educación, Ciencia y Tecnología
71. Universidad Católica Santa María La Antigua
72. Association française de Cartographie de la Faune et de la Flore
73. Mauthausen Memorial
74. Scotland’s Rural College
75. NHS Digital
76. Technische Hochschule Ingolstadt
77. Jönköping University
78. Tshwane University of Technology
79. Österreichische Galerie Belvedere
80. National Park Service
81. Universität Gaston Berger
82. Cahiers de Biothérapie
83. Réseau d’épidémiosurveillance en pathologie équine
84. Institut de Radioprotection et de Sûreté Nucléaire
85. Hochschule für angewandte Wissenschaften Hof
86. Benemérita Universidad Autónoma de Puebla
87. Autonomous University of Baja California
88. Institute of Methodologies for Environmental Analysis
89. Scuola Superiore Sant’Anna
90. Royal Netherlands Academy of Arts and Sciences - KNAW
91. Energy Systems Catapult
92. Coventry University
93. Digital Sample Freezing Platform (DSFP)
94. Macquarie University
95. Association pour le Traitement Automatique des Langues
96. Georgetown University
97. te.ma.gGmbH
98. INM - Leibniz-Institut für Neue Materialien gGmbH
99. Deutsches Institut für Urbanistik gGmbH
100. Deutsche Schillergesellschaft e. V. (Deutsches Literaturarchiv Marbach)
101. Campus Kristiania
102. CONSORTIA PAPYRUS
103. SHOT
104. LSTMed
105. Centre international de Ressources Microbiennes – Levures
106. Société Royale Belge d’Anthropologie et de Préhistoire (SRBAP)
107. AgResearch
108. Advocacy Unified Network
109. St. Anna Kinderkrebsforschung GmbH
110. LawThek GmbH
111. Ludwig Boltzmann Gesellschaft
112. Qatar National Library
113. ELI-ALPS
114. DDL Conference (Drug Delivery to the Lungs)
115. Abertay University
116. Rostocker Institut für Sozialforschung und gesellschaftliche Praxis e.V.
117. Spanish Foundation for Science and Technology
118. Pädagogische Hochschule Ludwigshurg
119. Badische Landesbibliothek
120. University of Technology Sydney
121. Universidade Federal Rural da Amazônia
122. Association pour les Rencontres Internationales en Science Fondamentale
123. National Institutes of Health, Office of the Director (NIH OD)
124. Eunice Kennedy Shriver National Institute of Child Health and Human Development
125. National Heart Lung and Blood Institute
126. BCU
127. Middlebury
128. The University of Notre Dame Australia
129. Paris West University Nanterre La Défense
130. Universität Catholique du Graven
131. Universidad Externado de Colombia
132. The Metabolomics Innovation Centre
133. McMaster University
134. MapBiomas
135. National Institute of Biomedical Imaging and Bioengineering
136. National Institute of Diabetes and Digestive and Kidney Diseases
137. Università di Roma Tor Vergata
138. American University of Bahrain
139. Florida State University
140. Tampere University
141. German Centre for Higher Education Research and Science Studies
142. National Center of Social Research – EKKE
143. Czech Social Science Data Archive (CSDA)
144. Institute for Employment Research (IAB)
145. Leibniz Institute for Research and Information in Education (DIfE)
146. FH CAMPUS D2
147. National Cancer Institute
148. National Institute of Allergy and Infectious Diseases
149. IDD
150. GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel
151. ILS – Institut für Landes- und Stadtentwicklungsforschung gGmbH
152. Tallinn University
153. WIPO
154. ICI Berlin
155. Jagiellonian University Medical College
156. AgEcon
157. University of Gävle
158. University of the Sunshine Coast
159. Torrens University Australia
160. USP - Universidade de São Paulo
DataCite Member Survey

At the end of 2022, we conducted our annual member survey, asking our members about their experience being part of the DataCite community and how DataCite services might help with any challenges they are facing. We thank the many members who participated in the survey for sharing their valuable perspectives with us. After analyzing their insights, we are pleased to publish the following summary of the main outcomes.

GENERAL

Similar to previous years, approximately 25% of DataCite members responded to our request and participated in the survey. Respondents

- Came from all continents (except Antarctica),
- Mostly came from Europe, and
- In 75% of cases, represented research and academic institutions.

This year, 87.5% of our members gave a minimum rating of 8 out of 10 when asked if they’d recommend DataCite to other organizations. This meant that the overall average rating was higher than in 2020 (83.3%) and 2021 (84.7%).

When asked what we do well, the top three areas mentioned were:

- Community engagement
- DOI registration
- Providing support

We believe the development and execution of our communication and event strategies contributed to community engagement being mentioned frequently. We also received positive feedback from many participants on our efforts to make the metadata schema development process more open and transparent.

As for what we can do better, the top three areas are

- Fabrica
- Documentation
- Statistics/metrics

The first of these is a reminder of the important role Fabrica plays for many of our members. This has been heard, and we hope that the latest release of Fabrica will address the needs of our users. In terms of enhancing our documentation, one of the goals of our annual vision for 2023 is the creation of a content strategy for documentation. Finally, with regards to metrics, we expect the Open Global Data Citation Corpus to transform the data citation landscape and address your metrics needs.

PRODUCT DEVELOPMENT

We also asked members about the top three services or features they look forward to in 2023 and these were:

- Bibliometrics dashboard
- FAIR visualizations
- Batch metadata update
For the first time, we also asked what challenges organizations are currently facing that DataCite services could help with. The three most common responses were:

- Citation metrics
- Batch DOI registration via API
- Relations to other resources and research outputs

**SUPPORT AND OUTREACH**

Similar to the general ratings (see above), members’ experiences of the DataCite helpdesk and our support pages improved compared to previous years, with median ratings of 9 and 8, respectively.

The top three topics for which respondents would like to see additional support documentation are

- Metadata (schema)
- Beginners guide, providing basics
- APIs

We are already beginning to address the first of these by including a greater number of examples in DataCite Metadata Schema 4.5. In addition, we have updated the metadata schema pages to provide more information about the metadata schema and how to contribute towards its future development.

Finally, the top three topics respondents would like to learn more about in DataCite events and meetings are

- Metadata
- Best practices
- APIs

These topics were the focus of last year’s DataCite Member Meeting and we are planning to intensify our efforts around member training, initially with the DataCite Open Training Days on 26-27 April 2023, where we will offer sessions on Metadata, Fabrica, and APIs.
Outreach

COMMUNICATION STRATEGY

In 2022, together with the CESG and EMEA EG we conducted stakeholder mapping exercises. With the stakeholder map in place, we were able to finalize a new communication strategy that defines our approach to communication over the next three years, in line with the strategic plan. While our communication strategy addresses communication on a strategic level, the actual implementation of communication activities is dealt with in separate outreach plans. This way, the communication strategy serves as a point of reference for the different communication activities.

EVENT STRATEGY

The communication strategy was complemented by the DataCite Event Strategy. Putting our Event Strategy together, we focused on two things: we wanted to be more proactive (than reactive) with our events planning, and we wanted to make the most of both organizing and participating in events, to better support our community. As a small non-profit organization working with a global community where there are a lot of topics to address and hundreds of events being organized, it is very important to prioritize and have a strategy in place for decision-making.

Both strategies are put into practice with the annual DataCite outreach plan which helps streamline the coordination and organization of our communications (blog posts, newsletter, mailings) and events (meetings, webinars, Open Hours).

VIRTUAL EVENTS

In 2022 we organized and hosted the following virtual events.

- Increase your impact through better metadata! (March 2022)
- Make Data Count webinar series
  - FORAGE: The hunt for existing data citations (March 2022)
  - EXPLORE: The need for an open classification system (April 2022)
  - BEGIN: Metadata for meaningful data metrics (May 2022)
- IGSN
  - Introducing the IGSN and DataCite Partnership (March 2022)
  - IGSN IDs – All Sample Types, All Disciplines (December 2022)
- Better Together webinar series (in cooperation with Crossref and ORCID)
  - Open new possibilities with Open Infrastructure (June 2022)
  - Facilitating FAIR Research Output Sharing (September 2022)
  - Complete Metadata as Robust Infrastructure (November 2022)
- 2022 DataCite Annual Member Meeting
  including training sessions about metadata (basics, schema, and formats), Fabrica, DataCite Commons, IGSN, communications, and APIs.
Adoption

DATAcite SUPPORT AND TRAINING

In 2022, we continued to provide technical support to the DataCite community through our help desk, support documentation, and training offerings. As part of our DataCite Annual Member Meeting, we offered two interactive training tracks: Core Training and Advanced Training. The core training covered DOI and metadata basics, an introduction to DataCite Fabrica and DataCite Commons, and an overview of the DataCite community and communication channels. The advanced training focused on metadata and technical topics, including creating connections between PIDs, DataCite APIs, and IGSN IDs. We also held training sessions tailored for consortium leads focused on account management and API basics.

DATAcite METADATA

To support the evolving use cases for DataCite DOIs, the DataCite Metadata Working Group invited feedback on the draft metadata schema 4.5 in 2022. The DataCite community provided input on the proposed changes, which include support for publisher identifiers and new resource types for instruments and study registrations.

FAIR WORKFLOWS

In the Implementing FAIR Workflows Project we track, implement, and share with the community FAIR practices in the neuroscience research processes. The Workflow Specification deliverable published in November 2022 outlined the key activities in the PID-centric workflows for the researchers and a number of DataCite DOI integrations for research-supporting tools and platforms.

COMMUNITY MEETINGS

DataCite Open Hours, Consortium Leads Open Hours, and our quarterly calls for Service Providers continued to be spaces for our community to convene, ask questions, and learn from each other. This year, our Open Hours covered topics such as metadata enrichment and IGSN IDs, featuring both staff and guest speakers.
Updates from the Steering and Expert Groups

In 2022, the Community Engagement Steering Group continued to meet bi-monthly and worked on the following topics:

- Communication and event strategies
- Defining the DataCite community
- A broader PID value proposition
- Setting metrics for engagement
- Outreach around the IGSN and FAIR workflows projects

The three regional expert groups focused on:

The Asia Pacific (APAC) Expert Group convened on the topics of regional engagement initiatives, through the discussion the APAC facing Better Together event series was incepted and then developed into a successful collaborative effort among DataCite, Crossref, and ORCID that delivered three webinars in 2022. The group also provided initial inputs for the DataCite Global Access Program which was launched by the end of the year and shared perspectives about live and recorded trainings, which informed part of our training strategy.

The Americas Expert Group discussed a variety of topics including how DataCite members deliver community support, strategies for creating connections between PIDs in DOI metadata, and ways to communicate the value of PIDs to different stakeholders.

In 2022, the EMEA Expert Group gave input and feedback on the DataCite stakeholder mapping and communication strategy. In addition, the topics of (national) PID strategies and outreach to discipline-specific communities were discussed.

In 2022, the Services and Technology Steering Group met every two months for 60 minutes. The group mainly worked together on an Opportunity Canvas to start building the strategy for DataCite Commons. In addition, five group members presented how they use DataCite services within their organizations.
Fabrica

Fabrica is our main service for DataCite Members to manage their account information, repositories, prefixes, and DOI registration. In 2022, we worked on:

- Improving the Fabrica interface to improve user experience following user feedback and behavior. The improvements include creating DOIs from any page, additional vocabulary terms for repositories platform and organization types, sorting the DOI list alphabetically by title, basic DOI metadata records download in CSV format, improved guidance and help messages, format and language field autosuggest with support for custom inputs, improved guidance and help messages, LaTeX support for metadata display and a new navigation bar.
- Providing full support for metadata schema 4.4 to enable repositories to create and update DOIs with RelatedItem metadata via the Fabrica form.
- Adding a DOI estimate field for Consortium Organizations in the Fabrica interface to enable Consortium leads to add the annual estimated registered DOI numbers to the settings of the Consortium Organization account when a new Consortium Organization account is created.
- Improving the prefix assignment workflow in Fabrica by streamlining prefix assignment to repositories by automatically assigning a prefix on repository creation by the user.
- Updating various software packages to stable releases, improving tests and workflows to improve the maintainability of the system and the productivity of the team.

Commons

DataCite Commons provides an integrated system for querying different persistent identifiers (DOIs, ROR IDs, ORCID iDs) and viewing their connections. We improved DataCite Commons to enable users to search for repositories. The functionalities include:

- Search DataCite Member repositories and repositories that are registered in the Registry of Research Data Repositories (re3data).
- Filter the search results according to compliance criteria (e.g. FAIRsFAIR Project), certificates (e.g. CoreTrustSeal), and software of the repository.
- A dashboard that contains relevant information about the repository as well as metrics about deposits and citations of research outputs in the respective repository.

APIs

DataCite provides multiple APIs to retrieve, query, and browse DataCite DOI metadata records. In addition, DataCite members have access to APIs to manage repositories, prefixes, and register DOIs. In 2022, we worked on:

- Providing full support for metadata schema 4.4 to enable repositories to create and update DOIs with RelatedItem metadata via the REST API.
- Updating various software packages to stable releases to improve the maintainability of the system.
Funded Projects
**FAIRsFAIR**

**PROJECT TITLE:** FAIRsFAIR  
**FUNDER:** European Commission, Horizon 2020  
**DATACITE FUNDS:** 242,500 €  
**DURATION AND START DATE:** 36 mo., March 1, 2019  
**PARTNERS:** The project is a pan-European collaboration of 22 partners with six core partners: DANS in the Netherlands (lead), CSC in Finland, the DCC and the STFC in the UK, Trust-IT in Italy, and the European University Association (EUA) in Belgium.  
**OBJECTIVES:** This project supplies practical solutions for the use of FAIR data principles throughout the research data life cycle. DataCite will collaborate with Re3data to make FAIR repositories discoverable.  

**re3data COREF**

**PROJECT TITLE:** re3data COREF  
**FUNDER:** Deutsche Forschungsgemeinschaft (DFG)  
**DATACITE FUNDS:** 130,406 €  
**DURATION AND START DATE:** 36 mo., January 1, 2020  
**PARTNERS:** DataCite, Karlsruhe Institute of Technology (KIT), the Berlin School of Library and Information Science at the Humboldt-Universität zu Berlin and the Helmholtz Open Science Office at the German Research Centre for Geosciences (GFZ)  
**OBJECTIVES:** re3data COREF aims to connect re3data as the reference for research data repositories with other services and infrastructures. By providing customizable and extendable core repository descriptions that are persistently identifiable and can be referred and cited in appropriately, re3data shall be advanced to facilitate the reuse of reliable and trustworthy information on research data repositories.  

**ORCID DE 2**

**PROJECT TITLE:** ORCID-DE 2  
**FUNDER:** Deutsche Forschungsgemeinschaft (DFG)  
**DATACITE FUNDS:** 74,940 €  
**DURATION AND START DATE:** 30 mo., January 1, 2020  
**PARTNERS:** DataCite, Helmholtz Open Science Office at the German Research Centre for Geosciences GFZ, the German National Library, Bielefeld University Library and the Technische Informationsbibliothek (TIB)  
**OBJECTIVES:** During its 30-month project period ORCID DE 2 seeks to expand and consolidate the existing network of scientific institutions that have already integrated ORCID into their infrastructures. One major goal is to increase the support for institutions and target groups interested in and using ORCID. Furthermore, in the course of ORCID DE 2, a survey about the need for an identification system for organizations will be conducted and implementation explored.
**DICE - Data Infrastructure for EOSC**

**PROJECT TITLE:** DICE - Data Infrastructure for EOSC  
**FUNDER:** European Commission, Horizon 2020  
**DATACITE FUNDS:** 125,668 €  
**DURATION AND START DATE:** 30 mo., January 1, 2021  
**PARTNERS:** The project is a pan-European collaboration of 24 partners with seven core partners: CINECA in Italy (lead), CSC and EUDAT in Finland, National Infrastructures for Research and Technology GRNET in Greece, Forschungszentrum Jülich FZJ and MPG in Germany, and Barcelona Supercomputing Center BSC in Spain.  
**OBJECTIVES:** The Data Infrastructure Capacities for EOSC (DICE) consortium brings together a network of computing and data centres, research infrastructures, and data repositories which propose to enable a European storage and data management infrastructure for EOSC, providing generic services and building blocks to store, find, access, and process data in a consistent and persistent way.

**FAIR Workflows**

**PROJECT TITLE:** FAIR Workflows  
**FUNDER:** Templeton World Foundation  
**DATACITE FUNDS:** 304,963 €  
**DURATION AND START DATE:** 36 mo., October 1, 2021  
**PARTNERS:** California Digital Library, Crossref, DataCite, Stanford University, Australian Research Data Commons, Dryad, Max-Planck-Institut für empirische Ästhetik, ChronosHub, Crossref, Center for Open Science  
**OBJECTIVES:** FAIR Workflows is aimed at building and implementing an exemplar FAIR and Open research workflow based on the reality of an entire research lifecycle. The project will produce a practical and easy-to-use guide for other scientists to improve the FAIRness of their own research.  
https://doi.org/10.5281/zenodo.7382642  
https://doi.org/10.5281/zenodo.7034858  
https://doi.org/10.5281/zenodo.6302912

**IGSN**

**PROJECT TITLE:** IGSN 2040 Vision to Execution: Institutionalizing IGSN ID services  
**FUNDER:** Alfred P. Sloan Foundation  
**DATACITE FUNDS:** 219,352 €  
**DURATION AND START DATE:** 15 mo., February 1, 2022  
**PARTNERS:** IGSN e. V.  
**OBJECTIVES:** IGSN and DataCite have a common purpose, and a close relationship in the future will provide mutual benefit to our shared vision of connecting research and identifying knowledge. The project brings years of experience across our organizations and communities to scale sample community engagement, develop sample identifier practice standards, and increase adoption globally.
**FAIR-IMPACT**

**Funder:** European Commission, Horizon Europe Framework Programme (HORIZON)

**DataCite Funds:** 333,750 €

**Duration and Start Date:** 36 mo., June 1, 2022

**Partners:** Led by DANS, 28 partners

**Objectives:** The project goal is to realise an EOSC of FAIR data and services, by supporting the implementation of FAIR-enabling practices, tools and services across scientific communities at a European, national, and institutional level.

**FAIRCORE4EOSC**

**Project Title:** FAIRCORE4EOSC

**Funder:** European Commission, Horizon Europe Framework Programme (HORIZON)

**DataCite Funds:** 720,000 €

**Duration and Start Date:** 36 mo., June 1, 2022

**Partners:** Led by CSC, 22 partners

**Objectives:** The project will deliver nine EOSC-Core components for a FAIR research life cycle, bridging the gaps identified in the EOSC Strategic Research Innovation Agenda. The components will enable an EOSC PID and research software infrastructure, sharing and access to metadata schemas, crosswalks, and improve discovery services over all EOSC repositories.

**FAIR Workflows**

**Project Title:** FAIR Workflows: Connecting Engineering Research Outputs using PID Infrastructure

**Funder:** Deutsche Forschungsgemeinschaft (DFG)

**DataCite Funds:** 49,725 €

**Duration and Start Date:** 9 mo., April, 2022

**Objectives:** The aims of this project are twofold. Firstly, we will address the need to enable and support the use of PIDs and metadata within NFDI4Ing as an essential component for the implementation of the FAIR principles for research data. Secondly, we will build interactive dashboards on the consortium output, show links and data reuse via event data and the PID graph.
Strategic Initiatives
In line with our mission and vision, DataCite actively participates and leads various initiatives through collaboration with stakeholders in the community.

**Data metrics**

We help further the adoption and implementation of responsible data metrics with, for example, the Make Data Count initiative.

Make Data Count is a global, community-led initiative focused on the development of open research data assessment metrics. The principles of our social and technical infrastructure are rooted in transparency and accessibility.

**Identifier registries**

We support community-led registries of identifiers such as the Research Organization Registry (ROR).

ROR is the Research Organization Registry – a global, community-led registry of open persistent identifiers for research organizations. The registry currently includes globally-unique persistent identifiers and associated metadata for more than 103,000 research organizations.

**Repository discovery**

We contribute to the development of repository discovery initiatives such as re3data with collaboration and financial support.

The Registry of Research Data Repositories (re3data) is an open science tool that offers researchers, funding organizations, libraries and publishers an overview of existing international repositories for research data.
Financial Overview
**Revenue**

Revenue continues to track positively with 28 new Direct Members and 160 new Consortium Members during 2022. Membership revenue was €43.5 K (2.6% below budget); and project funding was €336.5 K (13.4% above budget) mainly due to pre-payments for the FAIR-IMPACT and FAIRCORE4EOSC projects funded by the EU Commission. Total revenue was €2.844.3 K.

Based on pre-audited financials, our 2022 projected surplus is €569.5 K higher than expected.

This can be attributed to reduced spending and additional project funding. Our total income was exceeded by €336.5 K (13%) and our total expenditures are €233.0 K (9%) lower than expected. Our 2022 cash balance is compliant with § 82 Abs. 1 Nr. 3 AO (10%) considered as free reserves and the remaining reserves are used for project pre-financing.
Expenses

Total 2022 expenses were €233.0 K (9 %) under Budget. It can be noted that the lower expenses are due to ongoing COVID-19 travel reductions and membership outreach reductions. We recorded additional expenses in the licenses budget due to automation and professionalization efforts related to our strategic Pillar 1: Provide easy, efficient, and responsive community services to support the needs of our community.

We record project income and expenditures including project personnel expenses and other project-related expenses e.g. travel separately from membership income and expenses in the budget. This enables us to monitor the balance of our two funding sources, membership fees, and project funding.

1.4 % Outreach
5.8 % Infrastructure
5.4 % Licensing
4.8 % Travel
2.0 % Administrative
34.5 % Project Expenses
49.1 % Personnel
2022 Executive Board Roster

John Chodacki

John Chodacki (President) is responsible for overseeing the strategic planning, development, and management of the California Digital Library’s digital curation group, University of California Curation Center.

Marco Marsella

Marco Marsella (Treasurer) Senior Advisor Global Information System at FAO, where he contributes to the design of the system and the promotion of DOIs in the Plant Genetic Resources community.

Mohamed Ba-Essa

Mohamed Ba-Essa is Manager of Preservation and Digital Services at King Abdullah University of Science and Technology (KAUST), where he is responsible for Research Repository and Research Data Services, University Records Management and Archive, and Library Systems and Digital services.

Rebecca Ross

Rebecca Ross (Deputy President) is the Director of Strategy and Engagement at the Canadian Research Knowledge Network (CRKN), Canada’s national collective negotiating body for academic libraries in Canada. Rebecca is responsible for overseeing strategic planning, stakeholder engagement, and marketing and communications of CRKN’s diverse portfolio of licensing and digitization, preservation and access.

Irina Sens

Irina Sens (Business Officer) is the Deputy Director at the German National Library of Science and Technology (TIB). She led the German project “Registration of Scientific primary data” which was responsible for the introduction of DOIs for research data.

Adrian Burton

Dr. Adrian Burton is Director of Data Policy and Services at the Australian Research Data Commons which is a government-funded research infrastructure for data, cloud, and tools.
Ingrid Dillo

Dr Ingrid Dillo is Deputy Director at DANS (Data Archiving and Networked Services) in the Netherlands. She holds a PhD in history and has worked in the field of policy development for the last 30 years, including as senior policy advisor at the Dutch Ministry of Education, Culture and Science and the National Library of the Netherlands (KB). Among her areas of expertise are research data management and the certification of digital repositories. Ingrid is Co-chair of the Research Data Alliance (RDA) Council and Treasurer of the Board of CoreTrustSeal (CTS). Ingrid is project coordinator of the EU-funded project FAIR-IMPACT.

Viv Hutchison

Viv Hutchison is the Science Data Management Chief for the US Geological Survey. Her team develops and maintains a suite of science data management applications and leads a prominent open forum called the Community for Data Integration.

Tiwonge Banda

Tiwonge Banda has over 15 years experience in ICT4D project management, community engagement and organisational administration. Serving in various roles at UbuntuNet Alliance, he has spent his whole career working with and promoting the establishment of National Research and Education Networks in Eastern and Southern Africa (ESA); and promoting the use of digital technologies in research, teaching and learning. He is part of the core team that established the Internet network connecting 1,000+ universities, serving 2,000,000+ users in 12 countries in ESA under the AfricaConnect programme, funded by the European Union.

Mark Hahnel

Mark Hahnel is the founder of figshare, a repository where users can make all of their research outputs citable, shareable and discoverable. He has a PhD in stem cell biology at Imperial College London.

Salvatore Mele

Salvatore Mele is head of Open Access at CERN, where he co-architected the SCOAP3 initiative. His team develops solutions for Open Data and runs INSPIRE, a global Open Access digital library for High-Energy Physics.
Jennifer Gibson

Jennifer is Executive Director of Dryad, the open data publishing platform and community committed to the reuse of all research data. Since 2005, she has worked with scientists, funders, publishers, libraries, developers and others to explore fresh paths toward accelerating discovery through open research communication and open-technology innovation. Prior to Dryad, Jennifer was Head of Open Research Communication and founding Head of Marketing and Communications for eLife, a non-profit and research funder-backed initiative to transform science publishing.

Lynn Woolfrey

Lynn Woolfrey is a data scientist with 30 years’ experience curating and sharing African data. She manages the research data service at DataFirst at the University of Cape Town, South Africa and Africa’s only CoreTrustSeal certified research data repository. Lynn trains data managers across Africa and participates in UN projects to assist governments to assess their data-readiness for SDG planning and monitoring. Her research field of interest is open government data and information ecosystems for data-directed policy-making in African countries.

Torsten Reimer

Torsten Reimer is University Librarian and Dean of the University Library at the University of Chicago. He holds a PhD from the University and Munich and has a background in digital humanities and scholarly communication, having dedicated much of his work to research infrastructure. Before joining UChicago, Torsten was a member of the leadership team at the British Library, where he oversaw content and research services, including the UK DataCite consortium and the BL’s repositories. In previous roles, he worked at Imperial College London, Jisc, King’s College London and the Bavarian State Library in Munich.

Jennifer Gibson

Jennifer is Executive Director of Dryad, the open data publishing platform and community committed to the reuse of all research data. Since 2005, she has worked with scientists, funders, publishers, libraries, developers and others to explore fresh paths toward accelerating discovery through open research communication and open-technology innovation. Prior to Dryad, Jennifer was Head of Open Research Communication and founding Head of Marketing and Communications for eLife, a non-profit and research funder-backed initiative to transform science publishing.

Lynn Woolfrey

Lynn Woolfrey is a data scientist with 30 years’ experience curating and sharing African data. She manages the research data service at DataFirst at the University of Cape Town, South Africa and Africa’s only CoreTrustSeal certified research data repository. Lynn trains data managers across Africa and participates in UN projects to assist governments to assess their data-readiness for SDG planning and monitoring. Her research field of interest is open government data and information ecosystems for data-directed policy-making in African countries.

Torsten Reimer

Torsten Reimer is University Librarian and Dean of the University Library at the University of Chicago. He holds a PhD from the University and Munich and has a background in digital humanities and scholarly communication, having dedicated much of his work to research infrastructure. Before joining UChicago, Torsten was a member of the leadership team at the British Library, where he oversaw content and research services, including the UK DataCite consortium and the BL’s repositories. In previous roles, he worked at Imperial College London, Jisc, King’s College London and the Bavarian State Library in Munich.
2022 Staff Roster

Matthew Buys  Executive Director
Matthew Buys is the Executive Director of DataCite where he leads a passionate and committed team who provides the means to create, find, cite, connect, and use research globally. Matt’s work is focused on building DataCite as a sustainable global community, through efforts that support and advance community-driven practices.

Britta Dreyer  Director of Operations & Finance
Britta joined DataCite in 2015. She is responsible for budget and fiscal functions, policies and procedures, while managing the day-to-day operations of the DataCite business office in Hannover. She holds a BSM from Pepperdine University and an MBA in Entrepreneurship. Before coming to DataCite, Britta managed the human resources and marketing functions of the family business. Britta is passionate about Open Science and her onboarding sessions with new DataCite Members.

Helena Cousijn  Community Engagement Director
Helena leads the DataCite Engagement team and is responsible for all DataCite’s membership and community activities. She’s committed to DataCite’s mission and passionate about making DataCite services available globally. Before joining DataCite, Helena worked as Senior Product Manager for Research Data Management Solutions at Elsevier. She holds a DPhil in Neuroscience from the University of Oxford.

Sarala Wimalaratne  Engineering Director
Sarala is responsible for the technical and product vision at DataCite. She leads the technical and product teams to deliver DataCite services. She collaborates with external and internal stakeholders to improve existing member services and identify potential member service opportunities. She has been working with the Open Science and PID communities for many years. Before joining DataCite, Sarala spent ten years at the EMBL-EBI leading multiple data integration projects including the Identifiers.org resource. She holds a PhD in Bioengineering and a BE in Software Engineering from the University of Auckland.

Mike Bennett  Application Developer
Mike hails from sunny Scotland and his background is as a self-taught developer with an interest in building open tools and making things accessible. Prior to joining DataCite, he was the technical lead on the Bodleian Library’s Broadside Ballads Online project and a developer with the University of Edinburgh Digital Library. In his spare time, he can be found bothering a variety of string instruments or trying to solve a cryptic crossword.
Xiaoli Chen  FAIR Workflows Project Lead

Xiaoli leads the Implementing FAIR Workflows project, working on building exemplar FAIR practices throughout the research lifecycle with the project team and partners. She is also responsible for the outreach effort in the APAC region while co-chairing the DataCite APAC Expert Group. She holds an MLIS from Syracuse University and is working towards her Ph.D. in the same field from the University of Sheffield. Before joining DataCite, Xiaoli worked at CERN in parallel with her Ph.D. research on Open Science research practices among high-energy physicists.

Richard Hallett  Software Engineering Manager

Richard Hallett is a technical lead at DataCite, with nearly two decades of experience in software development. He joined DataCite in 2017 and has become a vital member of the team, leading the development of new products and services. Richard has participated in various grant-funded projects aimed at furthering DataCite’s mission. He is passionate about open data and open science, making valuable contributions to the organization’s efforts.

Veronica Gomes Ferreira  Intern

Veronica is our intern from Universiteit van Amsterdam finishing her master’s in Urban and Regional Planning with a specialization in Spatial Sustainability Studies. She works closely with our Community Engagement Team to provide support in our expansion to Latin America. Veronica is particularly interested in how data management helps improve and contribute to the global research community.

Kristian Garza  Product Design Manager

Kristian Garza is the Product Design Manager at Datacite, focused on designing state-of-the-art PID services using cutting-edge design thinking practices. He collaborates with RDA, Metadata Working Group, and MDC initiatives to support the Open Science Community. Kristian’s responsibilities include information architecture, user research, and design to ensure an intuitive user journey for all DataCite products. With a Ph.D. in Computer Science from the University of Manchester, he previously held research positions at MSSL and ELIXIR, and has made significant contributions to Datacite’s technical architecture.

Rorie Edmunds  Samples Community Manager

Rorie leads the partnership between DataCite and IGSN e.V. He is responsible for driving adoption of PIDs for physical samples, advocacy activities to scale samples community engagement, and the development of sample PID practice standards. Based in Tokyo, Rorie has more than 10 years’ experience of coordinating international initiatives and scaling communities globally. Before joining DataCite, he was Executive Director of ISC’s World Data System, and Head of Secretariat of CoreTrustSeal. Rorie’s research background is highly multidisciplinary. His first degree is Mathematics, and he holds a Ph.D. in Structural Geology.
Liz leads development of software and infrastructure for the Research Organization Registry (ROR). Liz has over a decade of experience in technology and higher education. Before joining DataCite, Liz was a technical lead and software developer at ORCID. She also has an MLIS from the University of Wisconsin-Madison and previously worked in various academic library and IT roles.

Gabi is our Community Manager and leads DataCite participation in two European Commission funded projects: FAIR-IMPACT and FAIRCO-RE4EOSC. As part of our Engagement team, she contributes to our outreach efforts and collaborates with our broader community. Gabi also volunteers in many scholarly communication organizations and initiatives to promote open research. Before joining DataCite, she worked at ORCID focusing on community engagement, driving membership and adoption across the Europe, Middle East and Africa region and within ORCID consortia.

Bryceson joined Datacite in September 2022 and recently moved from the United States to Portugal. Before joining DataCite, he was a software development engineer at Cambia Health Solutions, where he worked on a large-scale care gap management engine that assists health professionals in identifying gaps in member care, allowing them to provide better coverage. Prior to that, he was also involved in research at Texas A&M University. Outside of work, he spends much of his time playing music and rock climbing.
**Suzanne Vogt  Application Developer**

Suzanne joined DataCite in November 2020 as an application developer. She lives in the seacoast area of New Hampshire with her family. Suzanne contributes to the DataCite technical infrastructure. She has worked as a software developer for a number of companies in the U.S. with a wide range of technologies and applications, most recently in higher education at the University of New Hampshire library moving forward the modernization of their library digital repository. Outside of work, I play the violin in the New Hampshire Philharmonic, a local community orchestra.

**Kelly Stathis  Technical Community Manager**

Kelly leads adoption efforts and engages with the DataCite community on technical implementation. Before joining DataCite, Kelly worked at the Portage Network of the Digital Research Alliance of Canada, where they contributed to national data discovery initiatives and supported the DataCite Canada Consortium. Kelly holds an MLIS from the University of British Columbia, a BS in Computer Science, and a BMus in Music Composition.

**Paul Vierkant  Outreach Manager**

Since 2020, Paul is DataCite’s outreach manager. In addition to his communications work, he is involved in the re3data COREF, PID Network Germany, and NFDI4Ing projects. Before joining DataCite, Paul worked for the Helmholtz Open Science Office and several universities where he was involved in building publication and data repositories. Paul holds an M.A. in North American Studies, Italian, and ethnology. A dedicated advocate of Open Science, Paul strives to spread the idea of openness in scientific communication. Paul loves beautiful infographics and even more his crazy family.

**Joseph Rhoads  Application Developer**

Joseph works on the development team and contributes to the technical infrastructure. He lives in Rhode Island with his family. Prior to joining Datacite, Joseph led the Repository team at the Brown University Library. He studied mathematics and holds degrees in Computational Mathematics and Biomedical Mathematics. When not working, Joseph spends his time traveling with his family, playing board games, and reading as much science fiction and fantasy as he can get his hands on.
Research Publications


