

DataCite Annual Report 2021



CONNECTING RESEARCH,
IDENTIFYING KNOWLEDGE

Table of Contents

Who we are	3
Message from the... Executive Director, Matt Buys	6
Message from the... Executive Board President John Chodacki & Deputy President Rebecca Ross	8
Message from the... Executive Board Treasurer, Marco Marsella	9
DataCite Community	10
DataCite Services & Technology	16
Current Projects & Funded Initiatives	19
Financial Overview	22
2021 Executive Board Roster	26
2021 Staff Roster	28
Research Publications	31

Who we are

DataCite is a global community of organisations and researchers identifying and citing research outputs and resources. We provide services to create persistent records of research, enable discovery and reuse, and support workflows throughout the research lifecycle.

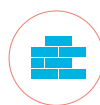
OUR VISION

**Connecting research,
identifying knowledge.**

OUR MISSION

DataCite's 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 develop and support reliable persistent identifier services.



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.



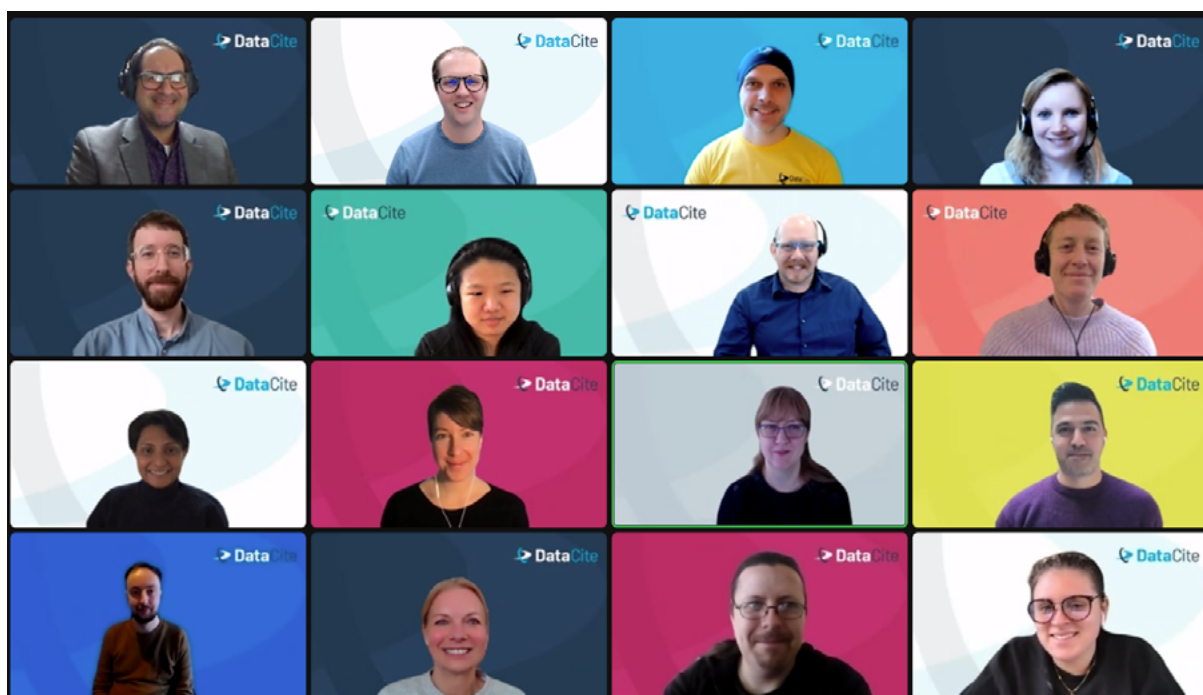
Inclusivity: We support a global community and value diverse perspectives.

**Message from
the...**

Message from the...

Executive Director, Matt Buys

Over the last two years, the global pandemic compelled us all to spend a little more time reflecting. During 2021, as a community, DataCite spent considerable time reflecting on our roots and discussing our future. We conducted an extensive strategic consultation process which ultimately led to the development of our multi-year strategic plan¹. This was an incredibly rewarding process and provided clear direction for our path forward.



DataCite team in March 2022

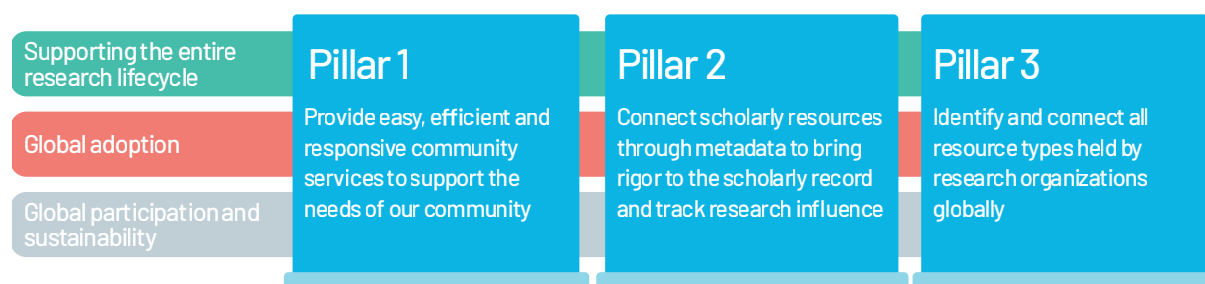
Through the year, we welcomed 32 new Members and 137 Consortium Organisations. Our community grows from strength to strength and is well positioned to execute on our strategic plan in the coming years. Our team expanded to support the growing need of our community and we launched some exciting new project initiatives that further our efforts to deliver value to our members.

As we move into 2022, we will be focussing on the foundational steps in executing the strategic plan. Underpinning our collective efforts are the DataCite core values. During the 2021 strategic visioning consultation, we reaffirmed our values with our members and team.

¹ Buys, M. (2021). Connecting research, identifying knowledge | DataCite vision to reality. DataCite. <https://doi.org/10.14454/79CY-QE15>



Our values support our mission driven strategy and bring focus to our efforts. In aligning with our values and mission, during 2022, we will be undertaking to execute three cross pillar themes. Through a series of engagements in 2022 we will be sharing more about our planned activities and seeking further input from the community:



We are proud of the work put into our strategic plan and look forward to working with the community on furthering our mission. Our work can only be successful if we rally together as a community and work with a common purpose. Our annual report is a testament to our collective community power. Together, we can help shape a better landscape for research by creating value through community-driven, innovative, open, integrated, usable, and sustainable services.

Message from the...

Executive Board President John Chodacki & Deputy President Rebecca Ross



This was a very productive year for DataCite. Under the amazing leadership of our Executive Director and staff, we continue to increase the value and service for our community. Of course, DataCite's work requires that we must iterate our approaches regularly to ensure we are meeting the needs of our members. To this aim, in 2021, we embarked on a thorough process of strategic (re)visioning.

The outcomes of our exercise have been well received by our communities and will help to reinforce the uniqueness of DataCite and help us effectively articulate how we support research institutions around the world.

It is important for any community to regularly analyze its approach to success. We are very proud of the work that was done this year and believe the outputs will help frame our strategies as we move forward. Thank you to our fellow Board members for their service and to everyone who makes DataCite such a great place to achieve our common goals.

Message from the...

Executive Board Treasurer, Marco Marsella



During fiscal year 2021, while still challenged by the uncertainties of the COVID-19 pandemic, we were able to meet our financial projections and continue on our positive path to financial sustainability. Alongside the outstanding growth of 32 new Members and 137 new Consortium Organizations, we successfully implemented the updated member model in 2021. I would like to recognize the excellent contributions of the DataCite members and staff during this transition period.

In 2021, we continued our prudent approach, with detailed monthly financial indicators being reviewed by the Board and cost reduction where possible. Additionally, we received project funding for two new community projects and the final FREYA payment. This led to the expected shift back into black of our bottom line. Our financial indicators continue to track positively for the upcoming years.

DataCite Community

DataCite is defined by our community.
We are very pleased to represent over
2442 repositories managed by over
1076 organizations in 50 countries.

New Members & Consortium Organizations

In 2020 the new DataCite member model was approved, which introduced the concept of consortia. Our community growth in 2021 demonstrates the value of consortia. Looking at the new member applications, 15% came from consortia, which included consortia in Latin-America, Africa, and the Middle-East. We are very pleased to see our community grow in these regions.

Another important development was that 137 Consortium Organizations joined the DataCite community. This illustrates that the new member model scales well and makes it easier for organizations to start working with DataCite. We would like to welcome all these new Consortium Organizations and we thank our Consortium Leads for the work they do to onboard and support new organizations.

NEW MEMBERS:

- | | |
|--|--|
| 1. National Institute for Materials Science (Japan) | 18. Stiftung Preußischer Kulturbesitz: Staatsbibliothek zu Berlin |
| 2. EuroFIR AISBL | 19. CESNET, z. s. p. o. |
| 3. Adelphi University | 20. National Tibetan Plateau/Third Pole Environment Data Center |
| 4. Riga Stradiņš University | 21. eScire (Consultoría, Tecnologías y Gestión del Conocimiento S.A. de C.V., Consortium Lead) |
| 5. National Center for Supercomputing Applications | 22. Eurac Research |
| 6. EUDAT (Consortium Lead) | 23. Meconed |
| 7. Yale Center for Biodiversity and Global Change | 24. Colorado School of Mines Arthur Lakes Library |
| 8. Jet Propulsion Laboratory | 25. Instituto de Pesquisas Energéticas e Nucleares (IPEN) |
| 9. UbuntuNet Alliance (Consortium Lead) | 26. KU Leuven |
| 10. Librairie Classiques Garnier | 27. University of Melbourne |
| 11. TÜBİTAK ULAKBİM | 28. Open Geographic Modeling and Simulation (OpenGMS) |
| 12. Exaptive | 29. Consortia SAS (Consortium Lead) |
| 13. Trase | 30. Insper |
| 14. Helsinki University Library | 31. UiT The Arctic University of Norway |
| 15. The Chinese University of Hong Kong | 32. International Institute for the Study of Kurdish Societies e. V. (IISKS, Consortium Lead) |
| 16. Bundesamt für Eich- und Vermessungswesen / Federal Office of Metrology and Surveying | |
| 17. Facebook (Open Research & Transparency) | |

NEW CONSORTIUM ORGANIZATIONS:

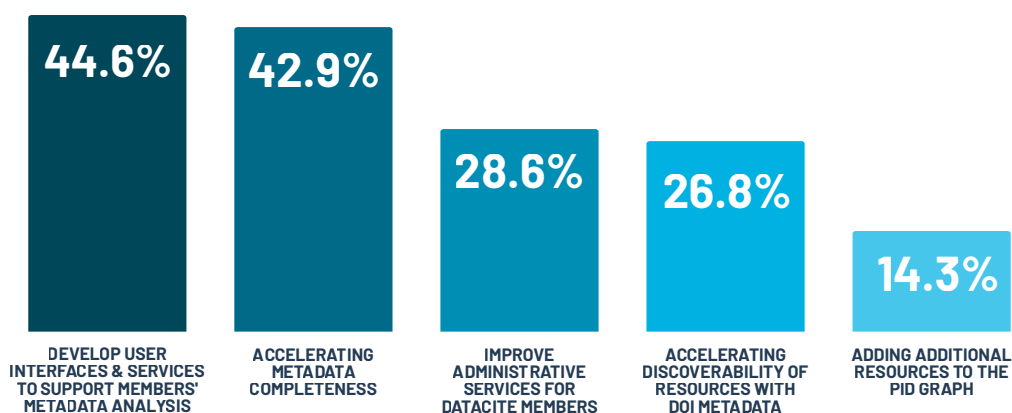
1. Competence Center of Urban and Regional Planning
2. DEPP: Direction de l'évaluation, de la prospective et de la performance
3. Direction des ressources et de l'information scientifique
4. SFA : Société Française d'Acoustique
5. FH Südwestfalen
6. Finnish Environment Institute
7. Hertie School gGmbH
8. National University of Ireland, Maynooth
9. Hochschule Aalen
10. Landesamt für Bergbau, Energie und Geologie (LBEG)
11. Wuppertal Institute
12. Kenyon College Library
13. Rockefeller University
14. University of Tampa
15. Living Lakes Canada
16. Lowitja Institute
17. Université Rennes 2
18. Living Goods
19. Zurich University of the Arts
20. Radiotekhnika
21. IPR Media
22. Swansea University
23. Robert Gordon University
24. University of Coimbra
25. Leibniz-Institut für Wissensmedien (IWM)
26. Hochschule Kaiserslautern
27. Landesinstitut für Arbeitsgestaltung des Landes Nordrhein-Westfalen (LIA.nrw)
28. Yaroslav-the-Wise Novgorod State University
29. Instytut Badawczy Leśnictwa
30. Parafia Ewangelicko-Augsburska w Gdańsku
31. IsoArch database
32. University of Paris-Saclay
33. Leuphana Universität Lüneburg
34. Zeppelin Universität gemeinnützige GmbH
35. Hochschule Offenburg
36. Fachhochschule Westküste
37. Pawsey Supercomputing Centre
38. E-LINE PRESS
39. Société informatique de France
40. Mercator Ocean International
41. Association QualiREL Santé
42. Istituto Italiano di Tecnologia
43. Federal Office for Migration and Refugees (BAMF)
44. ICM
45. University of Bonn
46. Ostwestfalen-Lippe University of Applied Sciences and Arts
47. BL Repository Services
48. Sociaal en Cultureel Planbureau
49. Flanders Hydraulics Research
50. Militärakademie an der ETH Zürich
51. Bundesamt für Kartographie und Geodäsie
52. Pädagogische Hochschule Freiburg
53. Lyrasis
54. Brandeis University
55. North Dakota State University
56. University of Pennsylvania libraries
57. The University of Queensland
58. OLOS
59. CyVerse
60. Brown University
61. The Cancer Imaging Archive
62. Trenholm State Community College
63. Corporacion Ecuatoriana Para El Desarrollo De La Investigacion Y La Academia (CEDIA)
64. GWDG
65. Collège de France
66. University of Lausanne
67. University of Concepción
68. Hochschule Düsseldorf University of Applied Sciences
69. Fachinformationszentrum der Bundeswehr (FIZ Bw)
70. acatech – Deutsche Akademie der Technikwissenschaften
71. Technische Hochschule Georg Agricola
72. WU Vienna University of Economics and Business
73. Universität Mozarteum Salzburg
74. Centerdata
75. German Center for Neurodegenerative Diseases
76. Institut für Kulturforschung Graubünden
77. Pacific Salmon Foundation
78. Department of Defense, Oak Ridge National Laboratory, Geospatial Science and Human Security division (ORNL-GSHS)
79. BAuA
80. Fachhochschule Bielefeld
81. Universidad Autonoma de Zacatecas
82. Deutsches Zentrum für Schienenforschung - DZSF
83. Institute for Advanced Study
84. University of Alabama
85. University of Wyoming
86. Maastricht University
87. Frankfurt University of Applied Sciences
88. Helmholtz-Zentrum für Umweltforschung GmbH - UFZ
89. JIV-ERIC
90. The Banco de España data Laboratory (BELab)
91. Herder-Institut für historische Ostmitteleuropaforschung
92. Karolinska Institutet
93. MF Norwegian School of Theology, Religion and Society
94. Instytut Fizyki Jądrowej PAN
95. CRITAC
96. RTI International
97. Musée du Quai Branly – Jacques Chirac
98. Union des Professeurs de Physique et de Chimie
99. Société Herpétologique de France
100. Akofena
101. Société Française de Statistique
102. National Institute for Demographic Studies
103. Polska Platforma Medyczna
104. Mangosuthu University of Technology (MUT)
105. Phoenix Bioinformatics
106. Carnets de la Consommation
107. Mondragon Unibertsitatea
108. Agence spatiale canadienne / Canadian Space Agency
109. Arctic Research Foundation
110. Centre for Addiction and Mental Health
111. Finnish Meteorological Institute
112. Kyiv Academic University
113. Opole University
114. Hochschule Osnabrück
115. ADMEE-Europe
116. National Academic Digital Repository of Ethiopia
117. Geologische Bundesanstalt (Geological Survey of Austria)
118. FH Campus Wien
119. Institute of Science and Technology Austria (IST Austria)
120. National Library of Technology
121. Wrocław University of Economics and Business
122. Masaryk University
123. Institut universitaire du Sud
124. Institut Pasteur
125. Société Publique Locale «L'Eau des Collines»
126. Leibniz-Institut Hessische Stiftung Friedens- und Konfliktforschung (HSFK)
127. IPH – Institut für Integrierte Produktion Hannover gGmbH
128. MIT Libraries
129. University of Turku
130. Klassik Stiftung Weimar
131. Universitätsbibliothek Johannes Gutenberg Universität Mainz
132. Geological Survey of Austria
133. Universität für Musik und darstellende Kunst Graz
134. Vorarlberger Landesbibliothek
135. Veterinärmedizinische Universität Wien
136. TH Köln – University of Applied Sciences
137. Cooperativa de Ensino Superior Politécnico e Universitário

DataCite Member Survey

In the 2021 DataCite member survey, we asked our members what they thought about DataCite in general, our services, support, and the DataCite community in 2021. Overall the responses we received reflect the goals of our strategic plan and annual vision 2022. We were happy to see that, just like in 2020, over 80% of the participants responded with a rating of 8-10 out of 10 when asked if they would recommend DataCite to other organizations (median rating of 9). Likewise the ratings for the experience with the DataCite helpdesk and support pages stayed the same (median rating of 8).

The top three things respondents think we do well are registering DOIs, community engagement, and support. In contrast to 2020, where both the DataCite Metadata Schema and our documentation were not mentioned at all, this year they were mentioned multiple times (10 in total). The top three things respondents think we could do better are community engagement, statistics /metrics, and documentation. Although this might seem contradictory to the responses of the previous questions, it resonates with what our focus is within our vision 2022 (e.g. developing a communication strategy and implementing the design manual) and the strategic plan (e.g. development of the bibliometrics dashboard).

When asked in which services DataCite should improve next year, respondents chose the following top 5 areas¹:



We are glad that we are already addressing these demands in our product roadmap (e. g. metadata completeness dashboard, Fabrica improvements etc.) as well as in pillar 1 and 2 of our strategic plan.

In the final question of the DataCite member survey 2021 we asked what the respondents see as the shared goal of the DataCite community. The top three goals mentioned are:

- Embedding PIDs (in particular DOIs) into the scholarly ecosystem
- Collaboration between stakeholders to enable open science
- Discoverability and access of research outputs

The top three DataCite services or features respondents are looking forward to in 2022 are:

- Bibliometrics dashboard (23,7%)
- DOI metadata enrichment (20,3%)
- Batch metadata update (20,3%)

¹ Multiple responses possible

DataCite Communications

DESIGN MANUAL

The first strategic priority of DataCite's vision 2021 was "Building on the DataCite identity". With the publication of the DataCite design manual, we took the first step in our journey to create a consistent visual identity for DataCite. Together with designers, we worked for several months on a variety of aspects of our design. In an iterative feedback process including the input from the Community Engagement Steering Group, we sought to update and harmonize existing design elements such as logos and colors, and at the same time introduce a scalable sub-brand system as well as our new DataCite community badges. Step by step, the new design is and will be applied to DataCite's resources, communication channels, and services. DataCite Members, Consortium Organizations, and registered Service Providers may use DataCite's name, logo, and the relevant community badge to announce their status and their activities. DataCite's logos and badges can be embedded into your websites and resources following the documentation on our support page.

BLOG PLATFORM

In December 2021, we migrated our blogging platform to WordPress. This transfer improves our blog post publication workflow. The design has not yet changed as the implementation of the new DataCite design manual will be part of the design transition roadmap in 2022.

RETIREMENT OF MAILING LISTS

In spring 2021, we retired the existing mailing lists and successfully transitioned to new mailing lists that are based on member information available in DataCite Fabrica. By doing so, contacts of our members receive relevant information depending on the roles they have assigned to their contacts. This enables voting representatives to receive notifications for board elections, service contacts to receive invitations to open hours or consortium leads to receive announcements for events, to give a few examples. We have made some changes in Fabrica to enable you to add more contacts and easily assign roles to them.

DataCite Community Activities

DATA CITE SUPPORT

The help desk is an important point of contact for DataCite members, incorporating technical support, bug reporting and product feedback. In 2021 we offered advanced support training for existing Consortia and we now offer a specialised support onboarding to all new Consortium Leads enabling them to provide first line support to all of the organizations in their consortium structure. We produced new materials for the support site including Getting Started guides (in English and Spanish), documentation requested by members and new guides to support new releases. The Knowledge Base and Community sections continue to be updated monthly and supplemented with short tutorials on the YouTube channel.

COMMUNITY MEETINGS

DataCite Open Hours is a meeting place where DataCite Members and Consortium Organizations are invited to share, interact and learn. In 2021 we

relaunched Open Hours, with part of the meeting now dedicated to giving members a chance to present their metadata schema suggestions. We also offered quarterly meetings to the DataCite Consortium Leads and Service Providers, providing an opportunity for them to hear the latest developments and learn from each other.

THE PID FORUM

DataCite is using advantage of the PID Forum to facilitate communication and interaction with and between members. There are currently two categories: the DataCite Chat Room, where DataCite shares news and updates including job postings, the newsletter and invites to upcoming events and the DataCite Consortium Leads group which is primarily being utilised as a platform for resource-sharing.

Updates from the Steering and Expert Groups

The DataCite Community Engagement team works with a number of groups that advise and support us. In 2021, we established Terms of Reference for these groups to clarify their roles and responsibilities. Currently, the Community Engagement Steering Group is responsible for identifying outreach needs for the DataCite community, including current and potential DataCite members as well as the broader research community. This includes: monitoring and assessing community needs and providing feedback on DataCite's priorities and services; recommending strategies and tools to communicate about DataCite's services, features, and plans; helping to build relationships, cultivate discussions and promote collaboration among the DataCite community; and involvement in DataCite's annual meetings. The CESG works with three regional Expert Groups: the APAC Expert Group, the EMEA Expert Group and the Americas Expert Group (the latter two of which were established in 2021).

In 2021, these groups contributed to:

- Developing key stakeholder messaging on the value and use of DataCite services
- Providing feedback on DataCite materials e.g. infographics
- Providing feedback on DataCite activities e.g. communications plan, support plan, strategic visioning, DataCite Zenodo community
- Providing feedback on new DataCite functionality and features e.g. Bibliometrics dashboard, DMP IDs
- Contributions to the translation of DataCite outreach materials to a variety of languages
- Discussion of broader PID landscape developments and their relevance to DataCite e.g. PID graph development, Research Data Alliance PID IG and National PID Strategies WG

DataCite Services & Technology

Development Journey

In 2021, we introduced the development journey, to help us deliver our vision to the DataCite Membership. We work in a dual-track workflow, in which the product team works on validating new features, and the development team works on building these features into the software. This workflow is iterative and aims to be user-driven. We have specific stages in which we involve the membership.

The first stage is that we collect community and membership insights – we use the DataCite portal Roadmap for this purpose. The membership and the community have provided 249 insights or suggestions alone in the last twelve months. That is more than insights collected in the previous two years combined (2019 [50], 2020 [93]). These insights come from all types of organisations and geographical locations. We have captured insights from 79 different organisations in the last twelve months.

The second stage is the validation section with users. Here we invited specific stakeholders that would help us validate that the features we are designing address member and user needs. This year, 37 stakeholders have contributed time to validating new products or features. That figure is not limited to Members; it also includes stakeholders from other RAs and organizations such as AGU. So there has been an excellent effort for the community of stakeholders to validate new DataCite features and services.

The insights collected and our validations have already shaped our vision for DataCite products in 2022 and will continue to do so in the future.

Fabrica

Fabrica is our main service for DataCite Members to manage their account information, repositories, prefixes and DOI registration. In 2021, we improved the workflow for adding contacts to member accounts. As part of this work, we made three fields (service, voting and billing contacts) on the member form mandatory to allow us to capture vital information to contact our members when needed.

Following the release of metadata schema 4.4, DataCite updated the services to enable users to create and update DOIs with the new metadata schema. The changes to the DataCite services to support metadata schema 4.4 was carried out in two phases. Phase one was completed last year, which included improvements to Fabrica interface to support the extended list of values for the 'resourceTypeGeneral' property. This year we will be working on the second phase to include adding 'relatedItem' information via Fabrica.

Commons

DataCite Commons provides an integrated system for querying different Persistent Identifiers (DOIs, ROR IDs, ORCID iDs) and viewing their connections. As part of the PARSEC project funded by the Belmont Forum (Collaborative Research Action on Science-Driven e-Infrastructures Innovation) managed through the National Science Foundation), we improved DataCite Commons on several fronts:

- **ORCID search and link** – allow researchers to manually claim a DOI for their work and link it to their ORCID records.
- **Accessibility Achievements** – provide information about the openness of a researcher's works based on the license information in the DOI metadata. The license data of DOIs are used to assign a badge: Open Hero, Open License, Open Science Triathlete or Open Access.
- **Co-author list** – provides a way to visualize the co-authors of a particular researcher. The users can use this feature to analyze the network of a particular researcher's collaborations and explore their profiles.

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. Following the release of metadata schema 4.4, DataCite APIs were improved to support creating, updating and retrieving DOI metadata following the latest version.

The EZ API was developed in 2017 as a temporary solution to assist those Members that transitioned from California Digital Library (CDL) EZID service. After four years of this service, we successfully retired the production EZ API on the 31st of December 2021. All members who were previously using the EZ API are now using our REST API.

Metadata

The DataCite Metadata Schema is the key to making data accessible, citable and discoverable. Last year, DataCite released Metadata schema 4.4 which includes:

- **Addition of the new subproperty "classificationCode" in the Subject property.**
- **Addition of new values to the resourceTypeGeneral property**
- **Addition of a new relatedItem property, with subproperties to contain specific details for containing publication information**

Current Projects & Funded Initiatives



FAIRsFAIR

PROJECT TITLE: FAIRsFAIR**FUNDER:** European Commission, Horizon2020 program**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.**REPORTS:** Sarala Wimalaratne, & Robert Ulrich. (2021). M4.8 Introduce additional components and practices to metadata schema and align them with FAIR data practices (1.0). Zenodo. <https://doi.org/10.5281/zenodo.5473107>

ORCID DE ORCID-DE 2

PROJECT TITLE: ORCID-DE 2**FUNDER:** German Research Foundation (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 in their infrastructures. One major goal is the increase of 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 of an identification system for organizations will be conducted and implementation explored.

PARSEC

PROJECT TITLE: PARSEC: Building New Tools for Data Sharing and Reuse through a Transnational Investigation of the Socioeconomic Impacts of Protected Areas**FUNDER/ SPONSORS:** French National Research Agency, France; National Science Foundation, USA; São Paulo Research Foundation, Brazil; Japan Science and Technology Agency, Japan**DATACITE FUNDS:** 177.070 €**DURATION AND START DATE:** 24 mo., July 1, 2019**PARTNERS:** American Geophysical Union, Ecological Society of America, European Geosciences Union, Japan Geoscience Union, Earth Science Information Partners, ORCID, DataCite, Research Data Alliance.**OBJECTIVES:** PARSEC will address the following two issues: 1) large amounts of research data related to the Earth and its ecosystems are either not well preserved or preserved at all and 2) there is also limited information on how diverse data are re-used for research and quantifying the value of curated data for such purposes, and how the quality of data preservation affects these outcomes.RESEARCH ORGANIZATION REGISTRY
(ROR)**PROJECT TITLE:** Research Organization Registry (ROR)**START DATE:** July 1, 2018**PARTNERS:** California Digital Library, Crossref, DataCite, Digital Science**OBJECTIVES:** ROR was launched in January 2019 and provides organization identifiers that are globally unique, stable, discoverable, and resolvable. In addition, ROR will develop an appropriate metadata schema for organizations and explore interoperability with other identifiers through relationship metadata.

PROJECT TITLE: re3data COREF

FUNDER: German Research Foundation (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 an appropriate manner, re3data shall be advanced to facilitate the reuse of reliable and trustworthy information on research data repositories.

REPORTS: 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): re3data Conceptual Model for User Stories, re3data, 10 p. <https://doi.org/10.48440/re3.012>

PROJECT TITLE: Make Data Count 3

FUNDER: Alfred P. Sloan Foundation

DATACITE FUNDS: 193.647 €

DURATION AND START DATE: 24 mo., April 1, 2020

PARTNERS: California Digital Library, Crossref, DataCite, DataONE, ScholCommLab, University of Ottawa, and ZBW – Leibniz Information Centre

OBJECTIVES: 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.

PROJECT TITLE: DICE – Data Infrastructure for EOSC

FUNDER: European Commission, Horizon2020 program

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, 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.

Financial Overview

Based on pre-audited financials, our 2021 projected surplus is €249.5K higher than expected.

This can be attributed to reduced spending, additional project funding and higher membership fee income. Our total income was exceeded by €104.9K (5.6%) and our total expenditures are €144.6K (8%) lower than expected. Our 2021 cash balance is compliant with § 62 Abs. 1 Nr. 3 AO (10%) considered as free reserves and the remaining reserves are used for project pre-financing.

Revenue

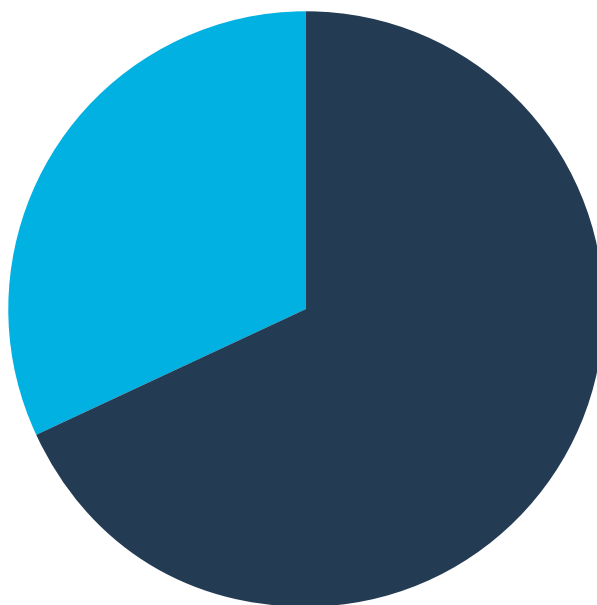
Revenue continues to track positively with 32 new Members and 129 new Consortium Members during 2021. Membership revenue was €7.5K (0.5% above budget); and project funding €97.4K (18.5% above budget) due to additional project funding from the Templeton World Charity Foundation for the FAIR Workflows Project for three years starting September 1, 2021 and the higher as budgeted final FREYA payment. Total revenue was €1,961.1K.

Project Fundings

31,7 %

Membership Fees

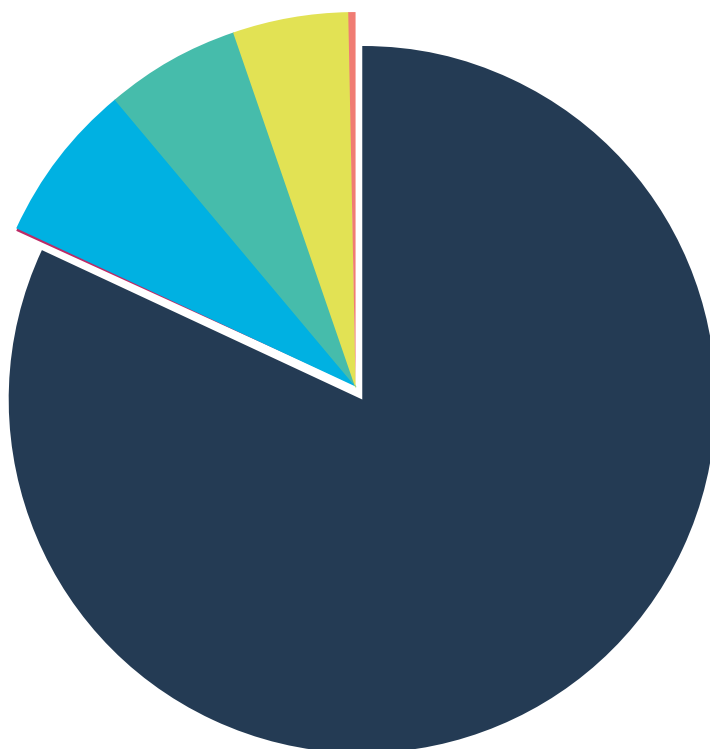
68,3 %



Expenses

Total 2021 expenses were €144.6K under Budget (8%). It can be noted that the lower expenses are due to travel reductions and membership outreach activities. These have been partially relocated towards the implementation of the outreach related strategic goal “building on the DataCite identity”. The dedicated cost management of our infrastructure continues to have a positive impact on our hosting costs. We recorded additional expenses for legal counseling due to continued tax compliance efforts.

0,5 %	Outreach
4,8 %	Administrative
5,9 %	Licensing
6,9 %	Infrastructure
0,1 %	Travel
81,8 %	Personnel¹



¹ Personnel costs represent a higher percentage of our costs due to COVID related restrictions to travel and membership outreach.

Executive Board & Staff Roster

2021 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.

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.

MARCO MARSELLA



Marco Marsella (Treasurer) is 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.

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.

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.

ADRIAN BURTON



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



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. She is also Treasurer of the Board of CoreTrustSeal (CTS) and Vice Chair of the Scientific Committee of the ISC/World Data System (WDS). Ingrid is project coordinator of the European H2020 project FAIRsFAIR.

ELEONORA PRESANI



Eleonora Presani is the Executive Director at **arXiv** and leads daily operations, including technical and business planning, governance, and fundraising efforts. After a PhD and postdoc in astroparticle physics, her career moved to scientific publishing and, later, to digital product development, servicing the research community.

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.

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.

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.

TORSTEN REIMER



Torsten Reimer is Head of Research Services at the **British Library**. His teams are developing the collection and services supporting researchers and research organizations, onsite and online, including digital preservation, repository, data and identifier services.

2021 Staff Roster

MATTHEW BUYS



Matthew Buys, Executive Director, leads a passionate and committed team at DataCite who provide the means to create, find, cite, connect, and use research globally. Based in Amsterdam, Matthew focusses on building DataCite into a sustainable global community. Previously Matthew was the Director of Engagement at ORCID where he played an important role in growing the community into an international-scale research effort. He completed a BA (Psychology) and Post-Graduate Diploma in Management at the University of the Witwatersrand. In addition, Matthew has also completed courses in Java, Flash, XML, html, Perl, and SQL.

BRITTA DREYER



Britta Dreyer, Business Manager, manages the DataCite business office. Member support, accounting, and managing DataCite's day-to-day business are her main activities. She holds a BSM from the Pepperdine University in Malibu and an MBA from the University of Applied Sciences and Arts in Hannover. Before coming to DataCite, Britta managed the family business. Her fields of work include human resources, marketing, and process optimization. She is passionate about business planning, project management, communication psychology, and her three children.

HELENA COUSIJN



Helena Cousijn, Community Engagement Director, is responsible for all DataCite's membership and community activities. She's committed to DataCite's mission of enabling data sharing and reuse and is especially passionate about data citation. It's important to her to communicate in a way that makes DataCite's services accessible to everyone. 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.

MARTIN FENNER



Martin Fenner, Technical Director, envisions, develops, implements and manages a robust technical architecture for DataCite. Until 2015 he was technical lead for the PLOS Article-Level Metrics project. Martin has a medical degree from the Free University of Berlin and is a Board-certified medical oncologist.

MIKE BENNETT



Mike Bennett, Application Developer, 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

Xiaoli Chen, FAIR Workflows Project Lead, is responsible for the outreach effort in the APAC region while co-chairing the DataCite APAC Experts 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, and contributed to the development of the various research data services there. Ukulele learner, bunny owner.

KRISTIAN GARZA

Kristian Garza, Product Designer, is responsible for determining information architecture, user research, and design to support an intuitive user journey for all DataCite services and products. Previously, he contributed to the development and implementation of a robust technical architecture DataCite. Before he held different research positions within the MSSL and the ELIXIR Interoperability platform. Kristian has a Ph.D. degree in Computer Science from the University of Manchester.

VERONICA GOMES FERREIRA

Veronica Gomes Ferreira, Intern, is our intern from Leiden Universiteit finishing her bachelor's in International Studies with a specialization in Latin America. 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.

RICHARD HALLETT

Richard Hallett, Software Engineering Manager, joined DataCite in late 2017 and contributes to the DataCite technical infrastructure. Having worked for years in software development for a variety of different organizations and industries (from large multinational companies, to small web agencies, to not for profits and then back into the commercial sector with e-commerce), Richard is pleased to be working in an open-source environment.

MARY HIRSCH

Mary Hirsch, Member Support Manager, takes care of making sure DataCite provides exceptional support for its community. Mary worked as a technical analyst at a global information company before transitioning to work as a documentalist at a research institute in Barcelona. She is passionate about Open Science.

PAUL VIERKANT

Paul Vierkant, Outreach Manager, contributes to the DFG-funded projects re3data COREF and ORCID DE. Prior to DataCite Paul worked for the Helmholtz Open Science Office and different universities where he was involved in building publication and data repositories. As a dedicated Open Science advocate Paul strives to spread the idea of openness in scholarly communication. Paul loves beautiful infographics and his crazy family even more.

LIZ KRZNARICH



Liz Krznarich, Adoption Manager, supports those integrating DataCite and ROR services into their workflows and systems, and collaborates with the community to identify and support new integration points. 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.

SUZANNE VOGT



Suzanne Vogt, Application Developer, joined DataCite in November 2020 as an application developer. She lives in the seacoast of New Hampshire with her family. Suzanne will contribute 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.

JOSEPH RHOADS



Joseph Rhoads works on the development team and contributes to the technical infrastructure. He lives in Providence, Rhode Island with his family. Prior to joining DataCite, Joseph led the repository team for nine and a half years at the Brown University Library to develop the Brown Digital Repository. Prior to that, Joseph studied mathematics and holds degrees in Computational Mathematics, Industrial and Applied Mathematics as well as Biomedical Mathematics. When not working, Joseph spends his time running around after his energetic toddler, traveling with his partner, playing board games, and reading as much science fiction and fantasy as he can get his hands on.

SARALA WIMALARATNE



Sarala Wimalaratne, Product Engineering Director, is responsible for the product vision and roadmap at DataCite. 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 10 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.

Research Publications

Cousijn, Helena et al. "Connected Research: The Potential of the PID Graph." Patterns Jan. 2021 : 100180. Crossref. Web. <https://doi.org/10.1016/j.patter.2020.100180>.

Cousijn, Helena, Ginny Hendricks, and Alice Meadows. "Why Openness Makes Research Infrastructure Resilient." Learned Publishing Jan. 2021 : 71-75. Crossref. Web. <https://doi.org/10.1002/leap.1361>.

Schrader, Antonia et al. "Die ORCID iD: Der persönliche Identifier in der Wissenschaft." DUZ Verlags- und Medienhaus GmbH (2021): n. pag. Web. <https://doi.org/10.48440/OS.HELMHOLTZ.032>



CONNECTING RESEARCH,
IDENTIFYING KNOWLEDGE

DataCite Annual Report 2021

www.datacite.org



Except for images and logos the DataCite Annual Report 2021 ©
2022 by DataCite is licensed under Attribution 4.0 International.
To view a copy of this license, visit
<http://creativecommons.org/licenses/by/4.0>