DataCite Annual Report 2020

DataCite

www.datacite.org

Index

Who We Are	3
Message from the Executive Director, Matt Buys	4
Message from the Executive Board President, John Chodacki	7
Message from the Executive Board Treasurer, Marco Marsella	8
DataCite Community	9
DataCite Services	14
DataCite Technology	16
Projects and Funded Initiatives	18
Financial Overview	21
Executive Board Roster	23
Staff Roster	26

Who we are

DataCite is a leading global non-profit organization that provides persistent identifiers (DOIs) that enable researchers to identify, locate, access, connect, and cite research datasets with confidence.

OUR VISION

Connecting research, identifying knowledge.

OUR MISSION

DataCite's mission is to be the world's leading provider of persistent identifiers for research. Through our portfolio of services, we provide the means to create, find, cite, connect, and use research. We seek to create value and develop community-driven, innovative, open, integrated, useable, and sustainable services for research.

OUR GUIDING VALUES

DataCite is trustworthy – we're fully dedicated to open research and all its related content.

We're community owned and driven – founded by the research community, we're a democratic organization that engages with and listens to the broad research community to meet the needs of researchers.

We provide timely support to our members – we respond quickly to the changing needs of our community, adding new services as needed.

We actively involve our membership in the building of new services – we're open to building services for all types of technologies.

We're a global leader – we have an international reputation for providing DOIs for traditional and non-traditional research outputs.

We're transparent – we have a structure that allows us to be proactive, agile, and responsive to the needs of the research community.

We advocate for the role of all research content in the research landscape – we engage in outreach that reflect the interests of our diverse community.

Message from the...

Executive Director, Matt Buys

As 2020 comes to a close, we reflect back on a year of global uncertainty with many challenges faced by our entire community. It can only be described as a year of the unexpected, and it demonstrated the importance and strength of communities working together. Despite the challenges faced, DataCite had a year with lots to celebrate and it is important to reflect on some of these highlights. It is important that we celebrate our collective success with our members and the broader community. At the end of 2020, our community consisted of members across 44 countries with over 2.300 repositories registering DOIs!

Earlier in 2020, we set out our Vision 2020 with some ambitious goals and objectives. In reflecting on the year, I would like to highlight a few moments that stood out:

In the **Community Engagement** team, we started the year with our open consultation process to develop and approve our new, scalable membership fee model. It was pleasing to have such an overwhelming approval from the DataCite General Assembly (97.3%) and the support we received from the community was fantastic. In particular, we want to thank the Membership Model Advisory Group members and the DataCite team members who consistently worked through the discussions with all community stakeholders. In addition to the new membership fees model, we continued to expand our global outreach with developing regions and gained over 50 new members this year. Our team set up regional expert groups and hosted virtual member meetings across three time zones (APAC, EMEA and Americas). In addition, we formalized our coordination with systems through our Service Providers program and we recently hosted our first registered Service Providers meeting.

Our **Business Office** had a busy year further improving our processes and systems as we mature as an organization. We implemented an automated billing system, which manages the complete billing workflow and are confident that this scalable solution supports the continued growth of our membership. We also established digitized and standardized accounting workflows with the German DATEV platform, which includes an interface to the German tax authorities to ensure continued compliant tax reporting, in collaboration with our tax consultant.

In August 2020, DataCite introduced a new **Product Engineering** team (through an internal team restructure) to work closely with our members and the community to drive the DataCite product strategy and ensure that our development closely aligns your needs. In the past few months, we have been conducting surveys and polling sessions through Open Hours and member meetings to gather feedback to incorporate in our product development goals for 2021. We are also working on improving our internal processes to deliver value-added, user-friendly, and stable services. Together with the Business team, the Community Engagement team, and the Development team, the Product Engineering team brings DataCite services to the center of our thinking to drive our decisions.

In our **Development team**, launching DataCite Commons was the most important activity of 2020. DataCite Commons is a discovery service launched in October 2020 that will replace DataCite Search during 2021, and will greatly expand its capabilities. We completed work to separate the staging and testing environments and formalized our technical infrastructure approach with improved monitoring and analytics of our member services. The Development team also implemented enhanced functionality across our services and tools, such as the updated DOI form in Fabrica.

We know it has been a challenging year for us all and it is during these times that we realize our collective strength. Despite these challenges we have continued to grow as a global community through working together to provide the means to create, find, cite, connect, and use research. The uptake and use of our services continues to increase.

RESOURCE TYPE (2020 DOIS REGISTERED)	
• Dataset	1,638,123
• Text	1,543,562
• Image	381,516
• Physical Object	268,199
• Collection	90,351
• Other	83,439
• Software	57,779
• Audiovisual	52,678
• Sound	13,583
• Interactive Resource	9,163
• Data Paper	8,218
• Event	694
• Workflow	500
• Model	207
• Service	97

DOIS REGISTERED BY YEAR	
• 2020	4,087,735
• 2019	3,789,954
• 2018	3,616,436
• 2017	3,095,152
• 2016	2,359,066
• 2015	2,570,439
• 2014	1,891,634
• 2013	713,814
• 2012	493,460
• 2011	158,338

I would like to congratulate the team on a successful year and all that they achieved. It is gratifying to close out a year with lots to celebrate. We thank our community for the continued support and look forward to working with you all during 2021.

Executive Board President,
John Chodacki

This year brought great challenges as we grappled with how to adjust to the pandemic and our 'new normal'. However, through this turmoil, our global DataCite community has also seen several opportunities for elevating our campaign for connecting the world's research. Like never before, governments, institutions, funders, and researchers not only see the value of networking research but also see the need to invest in the open infrastructure that makes that possible.

As DataCite continues to grow, it is important that we remain grounded in our values and collectively identify our strategic priorities for the future. During 2020, the team focussed on efforts to ensure improved stability, continued sustainability and scalability of the DataCite infrastructure.

Our transition to the updated membership model was an important milestone for DataCite as this supported scalability, sustainability, and inclusivity. The transition also means that our infrastructure is no longer significantly dependent on continued project funding, rather sustained through the cost recovery membership fee model. It was extremely positive to have received such strong approval of the model in the 2020 Business Meeting of the General Assembly. This ensures DataCite's open infrastructure is well positioned to help address the emerging needs of our communities.

The pandemic has created new complexities to our members. While the Board continues to provide strategic oversight to the operational activities of the organization as well as the strategic visioning, we also look for ways to help mitigate the challenges we are collectively encountering during the pandemic. During 2021, the Board will be working with Matt to lead an inclusive strategic visioning process that focuses on investment in our core services and risk mitigation for our members. We look forward to gathering input from our membership to propel DataCite into our next phase with robust technical infrastructure and an engaged community.

The pandemic has highlighted the value of networking research and the need for more investment in open infrastructure. With a strong foundation, DataCite is poised to play a key role in building a better future.

Executive Board Treasurer,
Marco Marsella

2020 was an unpredictable year: the pandemic has shattered our world and DataCite faced the daunting task of predicting our financial indicators in a world where nothing was predictable.

We made a conscious decision to adopt a very prudent approach, with detailed monthly financial indicator reports, reducing costs where possible, reducing income estimates based on the uncertainty and budget constraints.

The fruits of this approach are seen in the 2020 report and in the 2021 budget. Despite the challenging situation, we expect to maintain a largely positive trajectory with only a slight deficit in 2020 due to the timing of project funding and transitioning the new membership fees model. It should be noted that such a deficit was already forecast in the 2020 budget and that, due to mitigating actions, the deficit was indeed lower than expected. In 2021, we expect our bottom line to shift back into to black whilst continuing to adopt the same cautious approach. 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 2300 repositories managed by over 600 organizations in 44 countries.

2300+
Repositories

238
Members

44
Countries

20m +
DOIs

800+
Organizations

The DataCite membership and fee model

At the start of 2019, we embarked on a member consultation to ask members for their input on the fee structure. An advisory group looked at the outcomes and developed a proposal that was voted on at the 2020 business meeting of the General Assembly. It was pleasing to have overwhelming approval from the General Assembly (97.3%).

THE NEW FEE STRUCTURE ENSURES:

Predictability through graded and fixed tiers

Scalability as we continue to support the growing demands of our community

Sustainability by aligning with DataCite's core cost drivers

Inclusivity in providing low fees for small organizations

Simplicity by applying fees at the organization level

The updated member model and fee structure can be found on the DataCite website.

Throughout 2020, we worked with our members on the transition to the new model, which also included changes in Fabrica. We are very pleased that we now have 36 consortia that are all structured according to the new 3-layered model with Consortium Organizations and their repositories.

NEW MEMBERS

In 2020, DataCite again saw substantial member growth, with fifty members joining DataCite. Welcome to these new members!

1. Weizmann Institute of Science
2. PANGAEA
3. LifeWatch ERIC
4. State Archives of Belgium
5. Project Data Sphere
6. National Institute of Geophysics, Geodesy and Geography-BAS
7. Memorial Sloan Kettering Cancer Center
8. Auburn University
9. Institute of Oceanology Polish Academy of Sciences
10. University of Bern
11. World Bank - International Bank for Reconstruction and Development (IBRD)
12. Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq
13. RWTH Aachen
14. Underline Science
15. Pittsburgh Supercomputing Center
16. Cohen Veterans Bioscience, Inc.
17. Linguistics Data Consortium
18. European X-Ray Free-Electron Laser Facility

19. National Geological Archives of China
20. Méditerranée Infection Foundation
21. Georgia Institute of Technology Library
22. Schloss Dagstuhl – Leibniz-Zentrum für Informatik
23. Institute for Genome Sciences
24. Texas Medical Center Library
25. NASA Heliophysics Data Environment (NASA HPDE)
26. J. Craig Venter Institute (JCVI)
27. Computer-Aided Drug Design (CADD) Group, Center for Cancer Research, National Cancer Institute, National Institutes of Health
28. Institute of Smart Systems and Artificial Intelligence, Nazarbayev University
29. FORS
30. Crossref
31. ARTiFACTS
32. Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK)
33. Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz)
34. National Ecological Observatory Network (NEON)
35. South African Radio Astronomy Observatory
36. F1000
37. GFZ German Research Centre for Geosciences
38. University of Debrecen
39. Howard Hughes Medical Institute Janelia Research Campus
40. LYRASIS
41. Open Data Repository
42. Council for Scientific and Industrial Research (CSIR) South Africa
43. Perimeter Institute for Theoretical Physics
44. METADATAWORKS
45. arXiv
46. Karlsruhe Institute of Technology
47. Peking University
48. Georgian Integrated Library Information System Consortium 2017
49. Fraunhofer-Gesellschaft (Fraunhofer Society)
50. Chan Zuckerberg Initiative

Registered service providers program

We recognize that as a community-driven organization, our Members make use of various platforms or systems to help them manage their DOI workflow. Many of our Members make use of repository platforms, CRIS systems or other services. In order to better serve our Members, it is critical for us to have a close relationship with these Service Providers to ensure best practice adoption, better understand the needs of our Members, and support the workflows adopted by our Members.

Therefore in 2020, we launched the DataCite Registered Service Providers program. A DataCite Registered Service Provider is an organization that has integrated with one of the DataCite APIs in order to allow other DataCite Members and Consortium Organizations to register DOIs using the other organizations own log-in credentials. This program allows DataCite to improve communication between ourselves, our Members, and Service Providers such as repository platforms and CRIS systems. We want to provide guidance and best practices to the Service Providers that integrate with our APIs and to establish a consistent communication channel to hear from Service Providers about how we can address our Members' needs together.

IN 2020, THE FOLLOWING ORGANIZATIONS REGISTERED WITH US AS A SERVICE PROVIDER:

4Science (services for DSpace, DSpace-CRIS, Dataverse, OJS)

Atmire (services for DSpace)

eScire (services for DSpace, Dataverse OJS, with a focus on Mexico/Latin America)

Ex Libris Esploro

Figshare

Haplo Services

Poznan Supercomputing and Networking Center

Pure

The Library Code GmbH (services for DSpace)

See our Service providers listing for more details and the latest list of providers.

DataCite member survey

For the last 3 years, we have been sending an annual survey to our members to receive input which helps us reflect on how we did and how we can improve moving forward. The responses we received in 2020 play an important role in setting priorities for 2021. DOI registration remains the most important reason to join DataCite, with 89% of members indicating that this was their main reason to join DataCite. However, around one third of respondents also indicated that they joined DataCite to support the DataCite mission and participate in DataCite governance. Following our work on the DataCite value proposition, we asked members what value they are getting out of being a DataCite member. Discoverability of research outputs was mentioned most often, closely followed by being able to follow best practices and the ability to connect research outputs into the research ecosystem.

When asked what we do well, support was mentioned most often. In 2018, support was seen as the main thing we need to improve, so we are happy we succeeded in changing that! Registering DOIs, both through Fabrica and the APIs, was also mentioned by many members. Members appreciate that we are community driven, listen to our members, and really try to engage with the community. Open Hours was seen as a good example of this. In last year's member survey we heard that we should not get distracted by too many new things, so it was good to see that several members commented that we have a good balance between core services and innovation.

In previous years, one of the main areas for improvement was timely communication when changes were made, which was not mentioned as often this year. Members indicated that stability of services was their main concern in 2020, specifically DataCite Fabrica. It is therefore part of our 2021 vision to make improvements in this area. Several members also indicated they would appreciate more opportunity for interaction and the ability to influence DataCite's direction. This year, we also specifically asked whether members are part of a DataCite group or committee, and 18% responded they would like to join a group. We encourage interested members to contact us directly so we can find a suitable group or committee.

DataCite Services

Fabrica

Fabrica is our main service for DataCite Members to manage their account information, repositories, prefixes and DOI registration. In 2020, we added functionality to enable members to create and update more metadata through DOI Fabrica from. In addition to the required properties, users can now also add all of the recommended and optional properties.

To improve the member experience, we introduced a dedicated test system for existing members and potential members to test Fabrica that is not used by the DataCite development team. Members can use the test Fabrica system to explore features without real data.

In 2020, we also worked on improving the performance of Fabrica and underling services. Performance of the underlying APIs were improved to retrieve information more quickly. This intern reduced the load time and improved the user experience of Fabrica. This year we will continue to simplify and improve user experience of Fabrica.

DataCite Commons

In October 2020, as part of our work on the FREYA project, we released a frontend application to interact with the PID graph, DataCite Commons. It is a discovery service that enables simple searches while giving users a comprehensive overview of connections between entities in the research landscape. Users can search for connections between research outputs (DOI), people (ORCID) and organisations (ROR). At the moment, DataCite commons consists of all of DataCite DOIs, ORCIDs and RORs and part of Crossref DOIs (https://commons.datacite.org/statistics). We will continue to expand the PID graph by including DOIs from other RAs (CrossRef and JaLC) DOIs and their connections.

This year, we will be adding a repository search to DataCite Commons to enable users to find repositories and connected PIDs providing a comprehensive platform for searching the DataCite metadata store. Therefore, we will be retiring DataCite search and Repository Finder from December 2021.

APIs

Earlier in 2020, we separated our member API from our public API. This helps us route traffic to different servers based on whether the user is authenticated or not. This allows us to better handle server load and to better monitor response time for different sets of services.

We worked on improving the performance of REST API to enable fast querying and retrieval of information. The REST API is used by Fabrica and our other APIs (eg: OAI) thus improved performance of most DataCite services. This year, we will continue to improve performance of the APIs including GraphQL API which is used by DataCite Commons.

DataCite Technology

THE DATACITE DEVELOPMENT IN 2020 CONTINUED WORK IN THESE THREE AREAS:

Development of new and improvements of existing services

Improvements to the backend infrastructure to increase stability and performance, and to reduce costs

Technical support to members and users

Development of new and improvements of existing services

As part of the work in the FREYA grant, DataCite in May 2020 officially launched a GraphQL API – after making a prototype available in May 2019 – to query the PID Graph. In October 2020 we launched DataCite Commons, the web service that provides an easy to use interface for the GraphQL API. DataCite Commons provides a single search interface to content registered with DataCite and Crossref and we developed a workflow to automatically import Crossref metadata and convert them to the DataCite metadata format for a single search index of all DOIs. By the end of 2020 we had imported metadata for 10 million Crossref DOIs and will continue to do so in 2021. We also integrated ORCID (people) and ROR (organizations) metadata into the GraphQL API and DataCite Commons via federated search.

The development team worked on many other service improvements in 2020, most notably improvements to the DOI registration form in the Fabrica service that now supports all metadata properties in the DataCite schema.

Improvements to the backend infrastructure

API performance, and service stability continued to see improvements in 2020, mainly through work on our Elasticsearch search index, splitting our REST API into a public and member API, and adding auto-scaling to our Docker container infrastructure. We improved the monitoring of our services, including cost monitoring. We significantly reduced our total infrastructure costs by moving to an annual pricing plan with our cloud provider.

Technical support to member and users

The development team in 2020 continued to spend a significant proportion of its time with technical support. This includes bug fixes and small functionality improvements, but also second-line support on how to better use DataCite services in the form of addressing member questions via email, webinars and updated documentation. This work is done in close collaboration with the engagement team.

Projects and Funded Initiatives

maDMP

Project Title: maDMP

Funder: National Science Foundation, EAger Grants

DataCite Funds: 92272 €

Duration and Start Date: 24 mo., January 1, 2019

Partners: California Digital Library and DataCite

Objectives: maDMP focuses on creating metadata and DOIs for DMPs so they can be connected to other components of research.

FAIRsFAIR

Project Title: FAIRsFAIR

Funder: European Commission, Horizon2020 program

DataCite Funds: 242500 €

Duration and Start Date: 36 mo., January 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.

FREYA

Project Title: FREYA

Funder: European Commission, Horizon2020 program

DataCite Funds: 877500 €

Duration and Start Date: 36 mo, Dec. 1, 2017

Partners: Science and Technology Facilities Council (STFC) (lead), European Organization for Nuclear Research (CERN), European Bioinformatics Institute (EMBL-EBI), The British Library, PANGAEA, Data Archiving and Networked Services (DANS), Australian National Data Service (ANDS), ORCID, Crossref, Hindawi, Public Library of Science (PLOS).

Objectives: FREYA, the successor to THOR, got underway in 2018. The mission of FREYA is to foster a robust environment for a range of persistent identifiers as an essential component of the European Open Science Cloud (EOSC). FREYA partners are providing the essential building blocks for supporting changes in the way researchers work and the tools they use. This is all coming together in a new vision for how research is conducted, exploiting the full potential of Open Science and is a core to DataCite's strategic mission.

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: 177070 €

Duration and Start Date: 24 mo., May 1, 2019

Partners: American Geophysical Union, Ecological Society of America, European Geosciences Union, Japan Geoscience Union, Earth Science Information Partners, RDA, etc.), 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 divers 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.

COREF

Project Title: re3data COREF

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)

DataCite Funds: 130406 €

Duration and Start Date: 30 mo., January 1, 2020

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.

ROR

Project Title: Research Organization Registry (ROR)

Partners: Crossref, DataCite, Digital Science

Start Date: July 1, 2018

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.

ORCID-DE 2

Project Title: ORCID-DE 2

Partners: DataCite, Helmholtz Open Science Office at the German Research Centre for Geosciences GFZ, the German National Library, Bielefeld University Library and the German National Library of Science and Technology (TIB) Hannover

DataCite Funds: 74940 €

Duration and Start Date: 30 mo., January 1, 2020

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.

MAKE DATA COUNT 3

Project Title: Make Data Count 3

Funder/ Sponsors: Alfred P. Sloan Foundation

DataCite Funds: 193647 €

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

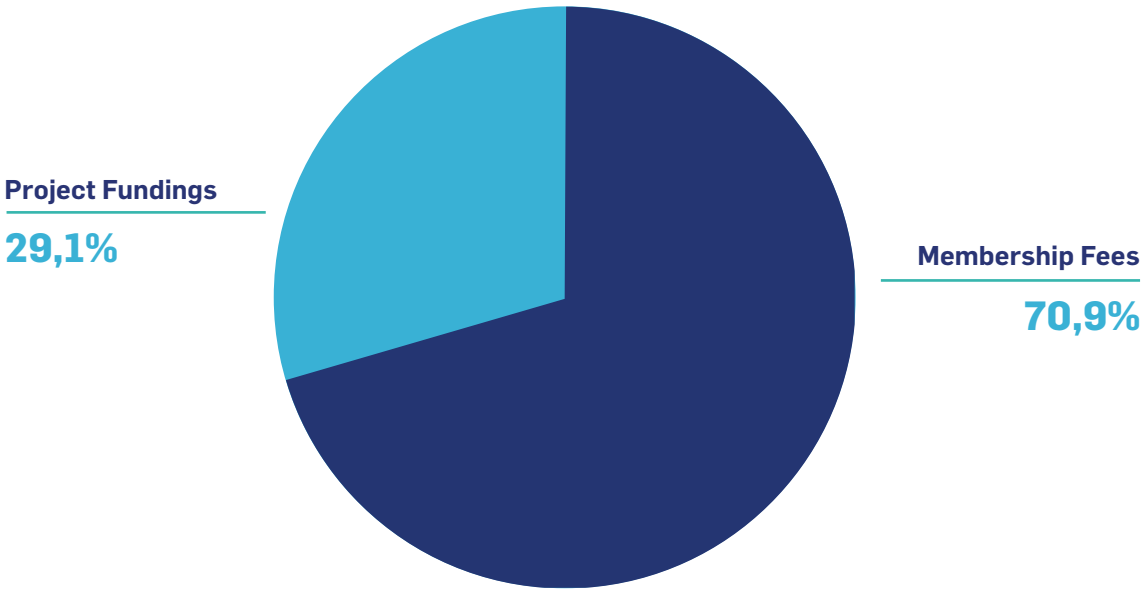
Objectives: Make Data Count is a global, community-led initiative focused on the development of open research data assessment metrics. The principles of our social and technical infrastructure are rooted in transparency and accessibility.

Financial Overview

Based on pre-audited financials, our 2020 projected deficit was €45.8K lower (29%) than budgeted. This can be attributed to reduced spending and dedicated cost management. Although we exceeded our membership target by adding 51 new members in 2020, the total income is €34.2K (3.5%) below budget due to accrual recognition of membership revenue and the timing of the membership start date. As planned, the 2019 surplus was used to counterbalance the 2020 deficit. Our 2020 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

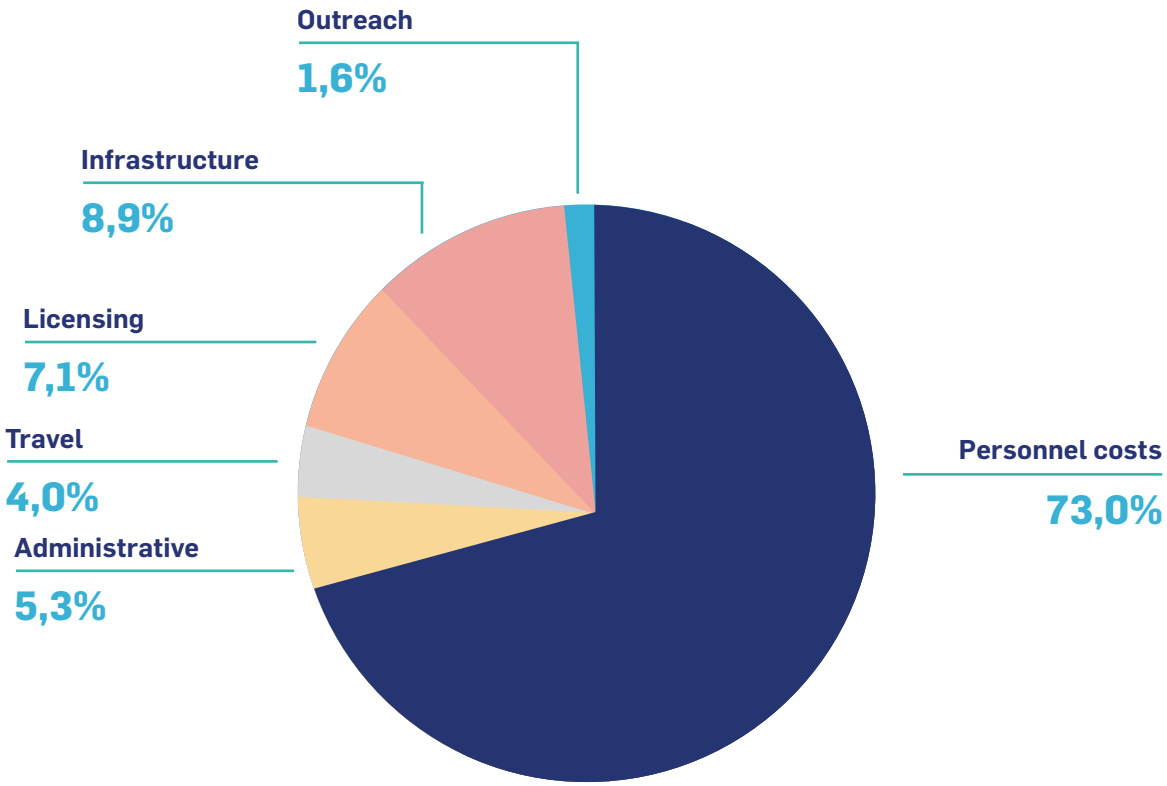
Revenue continues to track positively with 51 new members during 2020. Membership revenue was €953K (3.5 % below budget); and project funding €391K (14.4% below budget) due to deferred start of project work. Total revenue was 1,343.6K. €



Expenses

Total 2020 expenses were € 146K under Budget (9.1%). It can be noted that the lower expenses are due to travel reductions, membership outreach activities and adjusted resource allocation of project activities (existing staff as opposed to new additional team members). The latter was part of a risk mitigation strategy to minimize the potential impact of COVID-19 and the transition to the new membership model. This decision has proved to be sensible.

We recorded additional expenses for legal counsel and facilitated staff strategy development sessions. Administrative expenses are higher due to the payroll processing payments for US staff and one-time configuration costs of the billing management application JustOn.



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.

MOHAMED A. 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

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.

MARK HAHNEL

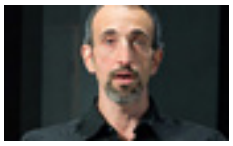
Mark Hahnel is the founder of figshare, a repository where users can make all of their research outputs citable, shareable and discoverable. He has a PhD in stem cell biology at Imperial College London.

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.

MARCO MARSELLA

Marco Marsella 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.



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.

IRINA SENS

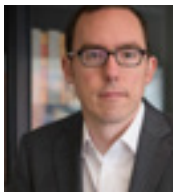
Irina Sens is the Deputy Director at the German National Library of Science and Technology (TIB). She led the German project "Registration of Scientific primary data" which was responsible for the introduction of DOIs for research data.

HERBERT VAN DE SOMPEL

Herbert Van de Sompel is Chief Innovation Officer at Data Archiving and Networked Services (DANS) in The Netherlands. He has a long-term interest in infrastructure to support scholarly communication in the digital era.

JAMIE WITTENBERG

Jamie Wittenberg is Head of Scholarly Communication at Indiana University Libraries, where she oversees research data, open access, and publishing services. Her team collaborates with faculty and students to curate, preserve, and share their work.



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.

REBECCA ROSS

Rebecca Ross 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.

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.

VERONICA GOMES FERREIRA

Veronica Gomes Ferreira, Intern, is our intern from Leiden Universitait 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 HALLET

Richard Hallet, Application Developer, joined DataCite in late 2017. He contributes to the DataCite technical infrastructure and the EU-funded FREYA project. 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.

KRISTIAN GARZA

Kristian Garza, Application Developer, 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.

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.

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.

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.

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.

DataCite Annual Report 2020

Identify, cite, and connect research with confidence.

www.datacite.org

DataCite
FIND, ACCESS, AND REUSE DATA