
Data Repositories selected for FAIRsFAIR support

FAIRsFAIR is playing a key role in the contribution to policies and practices for broader adoption of FAIR practices and in the development of standards for FAIR certification of repositories.

Through [two Open Calls](#) run between July and August 2019, FAIRsFAIR selected 22 Data Repositories - out of 73 applications - which will get specific support for **two different initiatives**:

- Support for data repositories towards achieving CoreTrustSeal certification (www.coretrustseal.org)
- Support for the “FAIRification” of repositories and improvement of FAIRness and interoperability of data

10 Data Repositories selected for the [CoreTrustSeal certification support](#)

The submissions were assessed based on the repository’s focus on long-term preservation, on reuse and Designated Community as well as on feasibility for obtaining CoreTrustSeal certification within the given time frame. The Expert Committee did take into account a diverse geographical and disciplinary spread among selected repositories as well.

Apollo - United Kingdom

- Apollo is the University of Cambridge, institutional repository, established in 2003 as a service for sharing the outputs of Cambridge, research activity. Apollo repository underpins the commitment of the University of Cambridge to preserve for the long-term and provide access to its research as widely as possible in order to contribute to society as well as to academic advancement. Apollo is one of the most widely used institutional repositories in the UK.
- www.repository.cam.ac.uk

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"I came away with a better understanding of how CoreTrustSeal aligns with the FAIR data principles. It was also helpful to learn about challenges faced by other data repositories and discuss how these could be addressed". - Caylin Smith - Digital Preservation Manager at Cambridge University Library



DaSCH - Data and Service Center for the Humanities (DaSCH) - Switzerland

- The DaSCH repository is designed for complex, qualitative data systems including bitstream objects (like digital images, sound or movie data formats), within the national Swiss humanity research community. The main goals of the DaSCH are the long-term data preservation and permanent access to research data in order to keep them available for further research.
Swiss researcher mainly in the field of humanities with complex, qualitative data connected to bitstream objects.
- <http://dasch.swiss>

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"My expectations were fully satisfied by the workshop and the FairsFair team...We received lots of hints and tips, how things can be optimized and streamlined. We also learned that there is not only one correct way to fill in the certification forms, but there is flexibility in the way to outline the relevant processes and situations. I would like to thank the organizers and the FairsFair team for this rewarding event." - Gabriela Kuster Vettiger - Digital Humanities Lab

DASS-BiH (Data Archive for Social Sciences & The Humanities in Bosnia & Herzegovina) - Bosnia & Herzegovina

- DASS-BiH is the national service with a role of ensuring long-term preservation and dissemination of social science research data in Bosnia and Herzegovina (BiH). Data included in the archive are from social science research in the following disciplines: economics, political science, psychology, sociology, society and culture, social welfare policy and systems.
As a national service for Bosnia and Herzegovina, it is included in the list of CESSDA ERIC data archives. It is an organisational unit of CREDI, which is a non-profit think-tank organisation working on the research of migration, labour market, social policy and education.
- credi.ba/en/dass-bih/

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"The workshop was the opportunity to discuss in details all questions about the process of obtaining CoreTrustSeal certification from the experienced archives as DANS and UKDA are. It was also interesting to learn how FAIR principles overlap with CoreTrustSeal requirements and to get practical advice how to incorporate these into the everyday activities of our repository." - Amela Kurta - CREDI



DASSH - The Archive for Marine Species and Habitats Data - United Kingdom

- DASSH is the UK archive for marine species and habitats data. Marine data are expensive to collect, therefore we promote the 'collect once, use many times' paradigm. Part of a network of thematic marine UK data archive centres, operating under the Marine Environmental Data and Information Network (MEDIN) framework. Partner in the biological component of the European Marine Data and Information Network (EMODnet) and the UK node for the Ocean Biogeographic Information System (OBIS).
- www.dassh.ac.uk

ESRF Data Repository - France

- The ESRF data repository is intended to store and archive data from photon science experiments done at the ESRF and to store digital material like documents and scientific results which need a DOI and long term preservation. Data are made public after an embargo period of maximum 3 years.
- data.esrf.fr

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"Very useful to help us understand the certification process - sharing this with other repositories at a similar stage helped us deepen our understanding - exchanging with data managers from other domains is a great way to spend the day". - Alex di Maria - European Synchrotron Radiation Facility

IAGOS Data Center - France

- The IAGOS Data Center serves the IAGOS data to the Scientific Atmospheric Research community. IAGOS is part of the H2020 project ENVRI FAIR aiming to make interoperable all data from European environmental research. Indeed, the IAGOS community should be extended to other environmental domains such as Marine, Solid Earth, Biodiversity, etc IAGOS has more than 600 individual users. In addition to these academics, IAGOS also provide Near Real Time data to the Copernicus Atmosphere Monitoring Service (CAMS) for model validation and data assimilation. Different research groups take advantage of the IAGOS data:
 - Satellite community for satellite data validation
 - Model community for data assimilation and model validation
 - Scientists working on process studies, trends and climatologies
- iagos-data.fr



Feedback from the [First Certification Support Workshop for Data Repositories](#)

"Thanks to the panel of experts encountered, the FAIRsFAIR workshop allowed IAGOS to have a better understanding of the CTS requirements and to identify the ones that will be the most challenging to achieve". - Damien Boulanger - Research Ingenieur at CNRS

ICOS Data Portal - Sweden

- The community is the ICOS Carbon Portal community of station PIs that have their data processed and curated and users of the data worldwide. Data users are partly the same community of PIs for their science on the systems they study locally; but also modellers of Earth 's climate system, dynamic vegetation models, forestry and crop scientists, atmospheric inversion modellers and Earth Observation data providers.
- www.icos-cp.eu

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"It was good to meet representatives of other selected repositories and to see that there will be support for us in the process of application. My conclusion after the day was that it will take some effort to get the required documentation together for evaluation, and it will lead to a constant update of our data-workflow, but in the long term it will be certainly beneficial for the (re-)use of our data. With the support from FAIRsFAIR I expect that we can submit soon our application for the CTS". - Harry Lankreijer - Lund University

The Movebank Data Repository - Germany

- The Movebank Data Repository publishes datasets of animal-borne sensor data. All published datasets are standardied by first importing them to [Movebank](#), a free, global research platform and e-infrastructure hosted by the Max Planck Institute of Animal Behaviour that helps researchers manage, share, analyse and archive animal movement and bio-logging data.
- www.datarepository.movebank.org

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"The workshop was really helpful to us as it was a great opportunity to build a connection to other repositories with similar interests and challenges. The formulated roadmap is a great tool for tackling the future steps necessary in achieving the certification". - Gabriel Schneider - University of Konstanz



SOCIB - Balearic Islands Coastal Ocean Observing and Forecasting System Data Repository - Spain

- The Balearic Islands Coastal Ocean Observing and Forecasting System -SOCIB-, is a Marine Research Infrastructure, a multi-platform and integrated ocean observing and forecasting system that provides streams of data, added value products, and forecasting services from the coast to the open ocean (Tintor^v© et al., 2013). SOCIB activities are grouped around three main areas that follow the GOOS themes: climate, ocean health, and operational services.
- thredds.socib.es
- apps.socib.es/data-catalog

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"Now I have a clearer idea of all the follow-up actions in order to apply for the CoreTrustSeal Certification for the SOCIB data repository". - Inmaculada Ruiz - Data Manager at SOCIB

Tárki Data Archive - Hungary

- The Tárki Data Archive, the national social science archive is operated by Tárki Joint Research Center (Tárki JRC), which is a private not-for-profit association of academic and educational institutions. Over the past two decades, the Data Archive has collected and archived more than 650 empirical social research data collections that are suitable for secondary analysis
- www.tarki.hu/eng/adatbank

Feedback from the [First Certification Support Workshop for Data Repositories](#)

"Thank you for all of the organizers of the workshop. Good to meet the partners and the leaders of this project personally, good to realize we are not alone with our challenges". - Péter Hegedus - Coordinator at Tarki Databank

Two face-to-face training and consultancy workshops will be organised with the selected repositories in 2020, while an in-depth evaluation and certification support will be delivered for the repositories to obtain the CoreTrustSeal certification before October 2021.

The selected repositories will also be invited to provide feedback on FAIRSFair outputs in view to identify a series of repository requirements in the light of implementing the FAIR principles.



12 Data Repositories selected to improve their level of interoperability

Twelve data repositories will participate in designing implementation features that will enable and increase repository interoperability.

6 Developer Repositories

A "Developer Repository" is usually set as a central data storage location. The following six repositories have been chosen and will participate in the development.

EUDAT B2Share - Finland

- The EUDAT Collaborative Data Infrastructure is a defined data model and a set of technical standards and policies adopted by European research data centres and community data repositories to create a single European e-infrastructure of interoperable data services
- b2share.eudat.eu

Dryad Digital Repository - USA

- The Dryad Digital Repository is a curated resource that makes research data discoverable, freely reusable, and citable. Dryad provides a general-purpose home for a wide diversity of data types
- datadryad.org

Central University of Punjab Knowledge Repository - India

- Knowledge Repository is a service initiated by the Central University of Punjab Library that provides long-term access to the scholarly resources generated by the University. The primary objective of this service is to collect, organize, preserve and provide access and dissemination of scholarly resources produced by the faculty members, student, staff and other members of the University Community.
- kr.cup.edu.in

Digital Academic Archives and Repositories - Croatia

- DABAR (Digital Academic Archives and Repositories) is the key component of the



Croatian e-infrastructure's data layer. It provides technological solutions that facilitate maintenance of higher education and science institutions' digital assets, i.e., various digital objects produced by the institutions and their employees

- dabar.srce.hr/en

The Movebank Data Repository - Germany

- Through this repository, Movebank allows users to publish animal tracking datasets that have been uploaded to Movebank (www.movebank.org). Published datasets have gone through a submission and review process, and are typically associated with a written study published in an academic journal. All animal tracking data in this repository are available to the public
- <https://www.datarepository.movebank.org>

Data Inrae - France

- The Data Inra portal provides new services to facilitate the management, sharing and retrieval of Institute data.
- <https://data.inra.fr>

6 Tester Repositories

Six Tester Repositories have also been chosen with the aim to engage with a wide community of experts in testing the solution.

Phaidra - Italy

- Phaidra (Permanent Hosting, Archiving and Indexing of Digital Resources and Assests) is the University of Padova Library System's platform for long-term archiving of digital collections
- phaidra.cab.unipd.it

FRDR - Canada

- FRDR addresses a longstanding gap in Canada's research infrastructure by providing a single platform from which research data can be ingested, curated, preserved, discovered, cited and shared



- www.frdr.ca/repo

PANGAEA - Germany

- PANGAEA services are generally open for archiving, publishing, and re-usage of data. The World Data Center PANGAEA is member of the ICSU World Data System
- www.pangaea.de

data Sciences Po - France

Data Sciences Po represents the data repository for Sciences Po but also a repository open to the wider scientific community in the social sciences.

- data.sciencespo.fr

DataverseNO - Norway

- DataverseNO is a national, generic repository for open research data from researchers from Norwegian research institutions. Some collections within DataverseNO also accept data from researchers from other institutions
- dataverse.no

NBN Atlas - United Kingdom

- The NBN Atlas is an online tool that provides a platform to engage, educate and inform people about the natural world
- nbnatlas.org

All applicant repositories and interested experts are welcomed to join the cooperation. The first reference implementation of the data repository features is planned to be ready in **February 2021** and the second implementation **one year later**.

