FAIRSFAIR

Fostering Fair Data Practices in Europe

Assessment framework for FAIR data services

21 February 2022

FAIRsFAIR "Fostering FAIR Data Practices in Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558

Agenda

- Welcome & introduction
- Methodology & outputs
- The FAIRsFAIR Service Assessment Framework
- Using the framework
- Q&A
- Wrap up

Patricia Herterich, DCC

Sara Ramezani, SURF

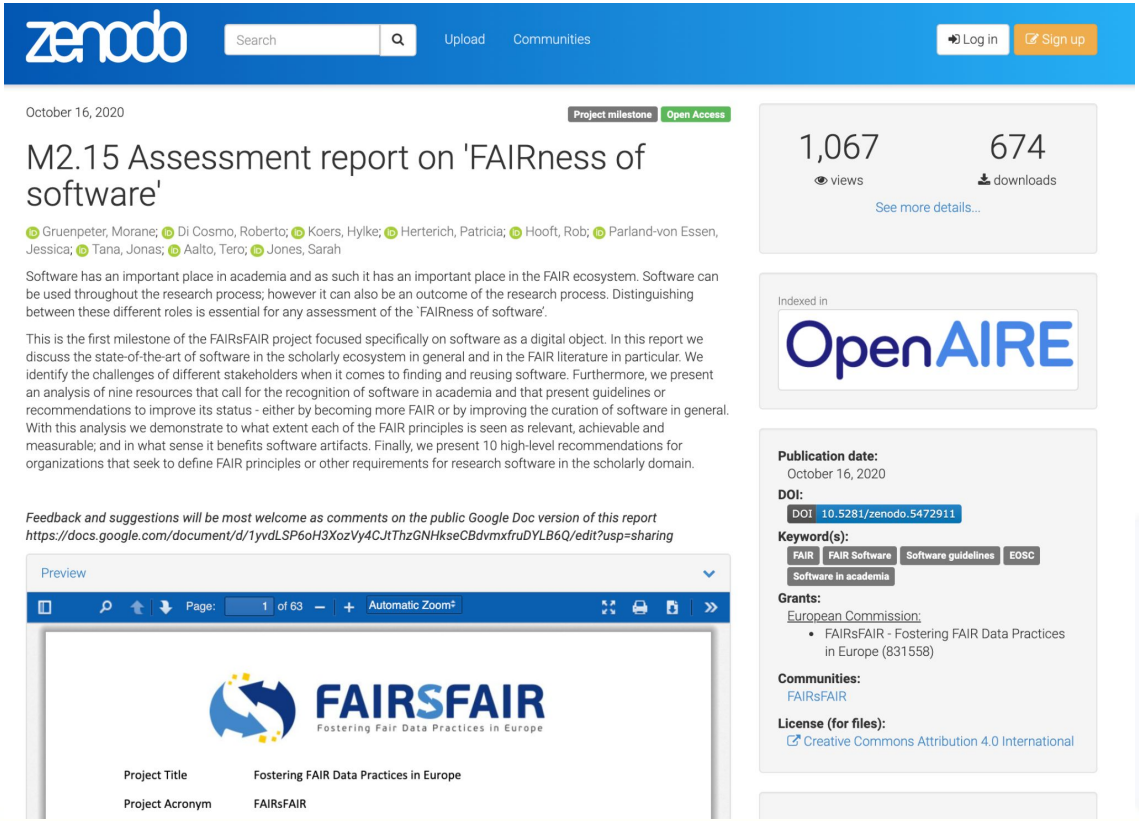
Objective

To deliver an assessment framework for data services that will help service owners to incrementally improve their services

→ *stimulating an optimal interplay between digital objects and services*

→ *help realize the full potential of a truly FAIR ecosystem*

M2.15 Assessment report on 'FAIRness of software’



zenodo

Search | Upload | Communities | Log in | Sign up

October 16, 2020

Project milestone | Open Access

M2.15 Assessment report on 'FAIRness of software'

Gruenpeter, Morane; Di Cosmo, Roberto; Koers, Hylke; Herterich, Patricia; Hooft, Rob; Parland-von Essen, Jessica; Tana, Jonas; Aalto, Tero; Jones, Sarah

Software has an important place in academia and as such it has an important place in the FAIR ecosystem. Software can be used throughout the research process; however it can also be an outcome of the research process. Distinguishing between these different roles is essential for any assessment of the 'FAIRness of software'.

This is the first milestone of the FAIRsFAIR project focused specifically on software as a digital object. In this report we discuss the state-of-the-art of software in the scholarly ecosystem in general and in the FAIR literature in particular. We identify the challenges of different stakeholders when it comes to finding and reusing software. Furthermore, we present an analysis of nine resources that call for the recognition of software in academia and that present guidelines or recommendations to improve its status - either by becoming more FAIR or by improving the curation of software in general. With this analysis we demonstrate to what extent each of the FAIR principles is seen as relevant, achievable and measurable; and in what sense it benefits software artifacts. Finally, we present 10 high-level recommendations for organizations that seek to define FAIR principles or other requirements for research software in the scholarly domain.

Feedback and suggestions will be most welcome as comments on the public Google Doc version of this report
https://docs.google.com/document/d/1yvdLSP6oH3XozVy4CJtThzGNHkseCBdvmxfruDYLB6Q/edit?usp=sharing

Preview

Page: 1 of 63 | Automatic Zoom

FAIRSFAIR
Fostering Fair Data Practices in Europe

Project Title	Fostering FAIR Data Practices in Europe
Project Acronym	FAIRsFAIR

1,067 views | 674 downloads

See more details...

Indexed in OpenAIRE

Publication date:
October 16, 2020

DOI:
DOI 10.5281/zenodo.5472911

Keyword(s):
FAIR | FAIR Software | Software guidelines | EOSC | Software in academia

Grants:
European Commission:
- FAIRsFAIR - Fostering FAIR Data Practices in Europe (831558)

Communities:
FAIRsFAIR

License (for files):
Creative Commons Attribution 4.0 International

M2.15 Assessment report on 'FAIRness of software: https://doi.org/10.5281/zenodo.4095091

How we reached the Assessment Framework

- Review of FAIR assessment frameworks for data (14 in total)
- Review existing assessment frameworks for services (not necessarily ‘FAIR’)
- Case studies: ‘How is this service enabling FAIR?’
- Formulate guiding principles for the assessment framework (so, still ‘meta’)

And, of course, lots of interactions and discussions with stakeholders and related working groups & projects

M2.7: Case studies and methodology for ‘FAIR enablement’

4.2. Case Study 1: B2FIND

Service Summary

B2FIND[21] is a metadata aggregator. The service harvests metadata from different community repositories and harmonises them such that users and services can search through the combined metadata. B2 FIND offers a rich faceted graphical search interface and a HTTP REST API that has been implemented in python for EUDAT’s B2FIND Training[22]

URL: http://b2find.eudat.eu/ **EOSC:** https://marketplace.eosc-portal.eu/services/b2find

Users

The service targets two types of user groups:

- Scientific communities that can provide their metadata and integrate via the B2FIND service with other metadata
- Scientists who can employ the service to search for interesting research data across different communities simultaneously.

Services

- Metadata harvesting and harmonisation to communities with a tool to search across the metadata for scientists.
- The relevant metadata of a DO is shown and a link to the metadata provenance is provided.

Target Digital Objects

- Metadata entries

Examples

- B2FIND entry (KONTROL 1984[23])
- OAI-PMH dataset’s metadata[24]

Purpose

B2FIND is a metadata aggregator. It gathers metadata from communities and repositories and integrates the different types of metadata. It provides a graphical user interface and an API to present the metadata and allows faceted searches across the metadata corpus.

Adoption

By now B2FIND hosts 824566 metadata entries harvested from 22 communities. We were unable to establish from the documentation how many users use B2FIND.

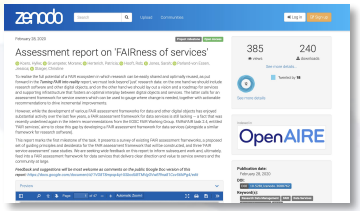
Documentation

EUDAT provides guidelines on how to use the B2FIND services[25] as well as detailed guidelines for harvesting and mapping metadata[26]

FAIR enablement mapping (see Annex C for details)

F1	F2	F3	F4	A1	A1.1	A1.2	A2	I1	I2	I3	R1	R1.1	R1.2	R1.3

FAIR enablement mapping:
Enable / Respect / Reduce

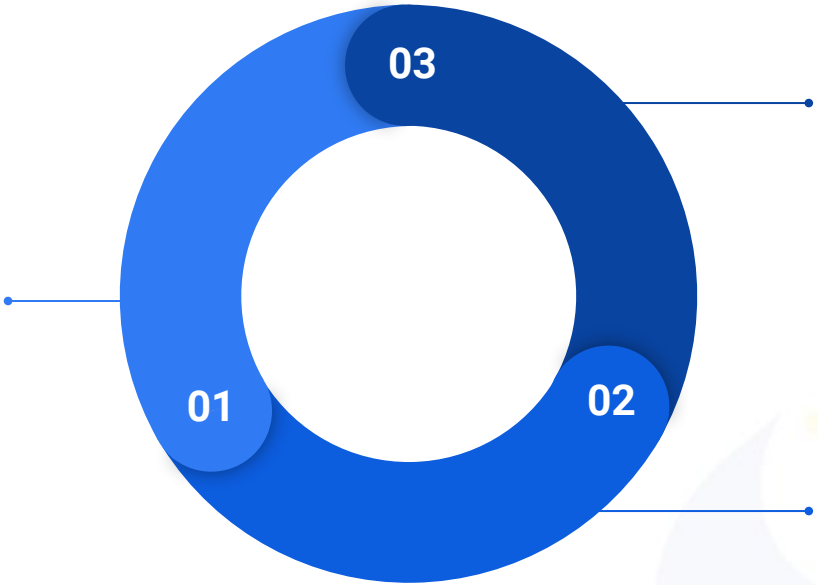


“M2.7 Assessment report on 'FAIRness of services'”, available at: https://doi.org/10.5281/zenodo.3688762

FAIRness of Services: Methodology

Structured literature review

Covering earlier work on service assessment such as CoreTrustSeal, the TRUST principles - but also recent work from EOSC-synergy, EOSC-nordic and other related projects.



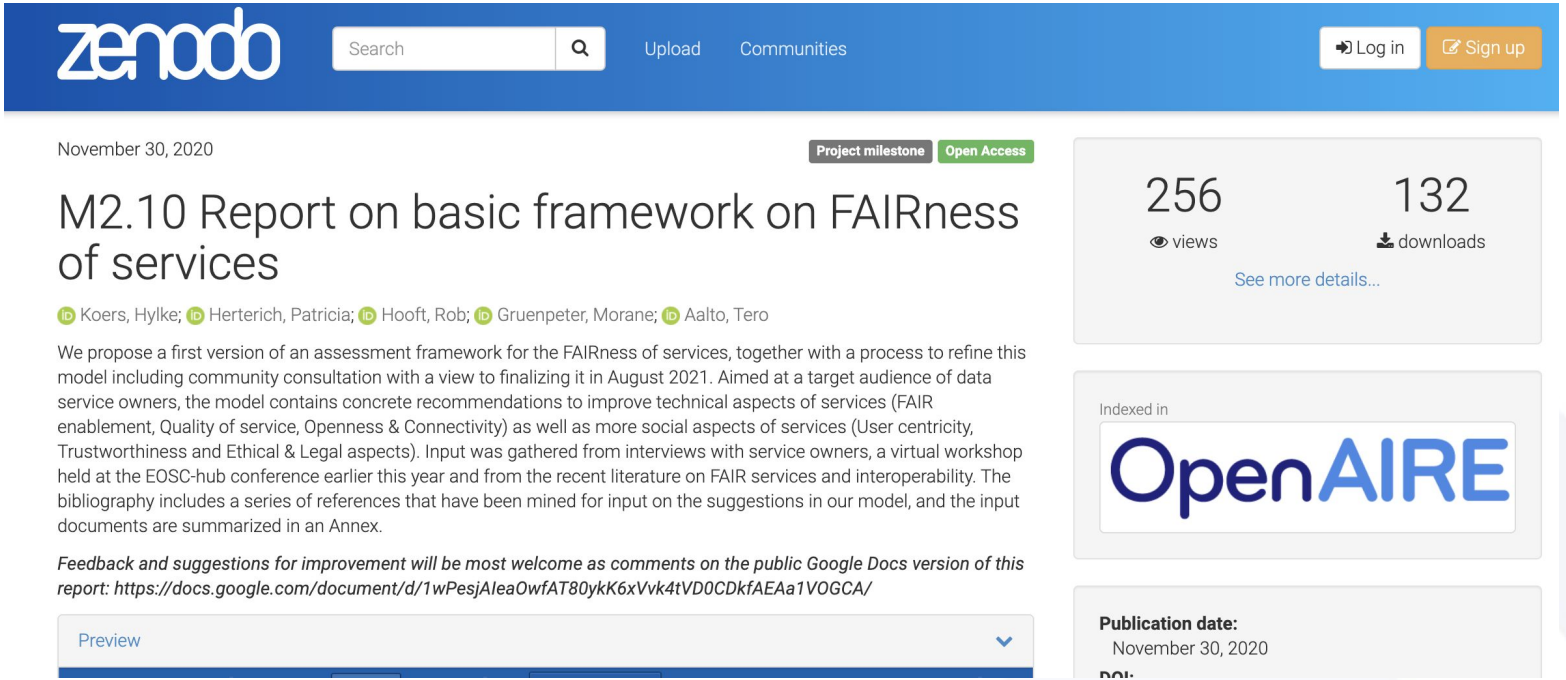
EOSC-hub week workshop

Feedback from FAIRsFAIR workshop on FAIR certification of repositories and other data services during EOSC-hub week in May 2020

Interviews with service owners

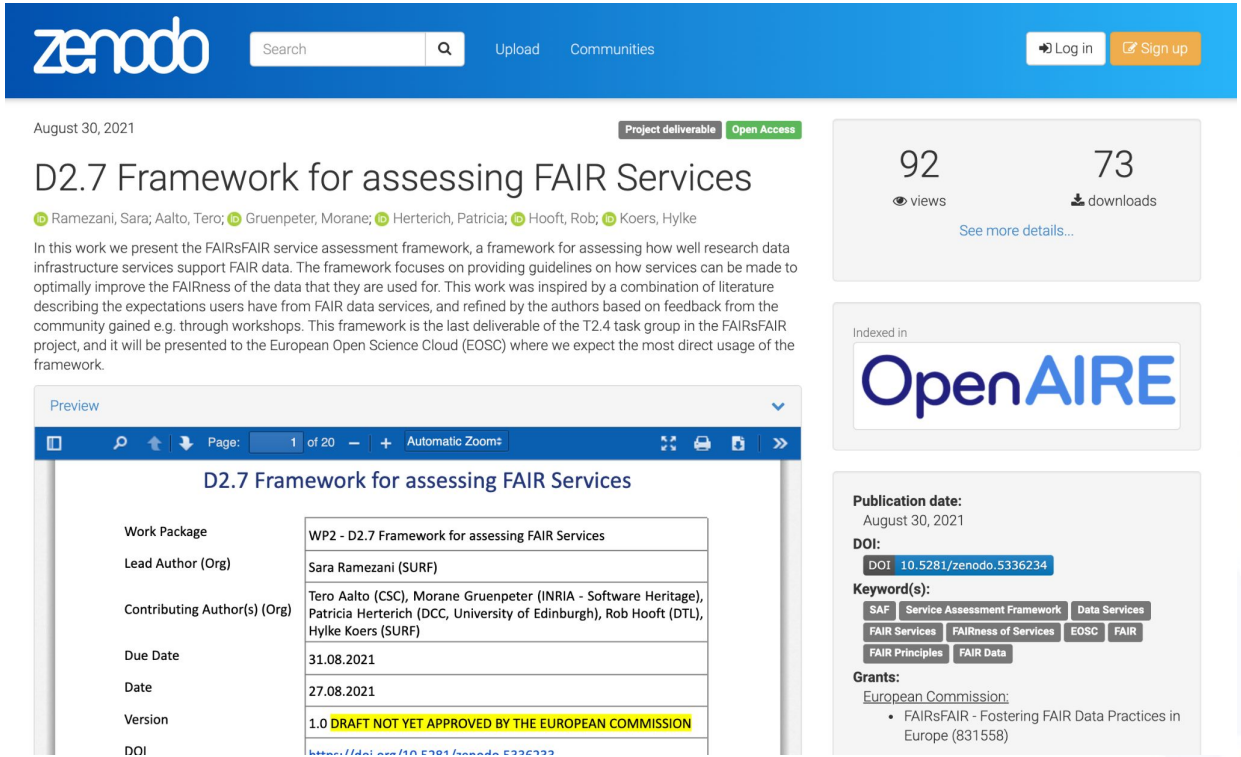
Semi-structured interviews with a range of service owners

M2.10: Basic framework for FAIR service assessment



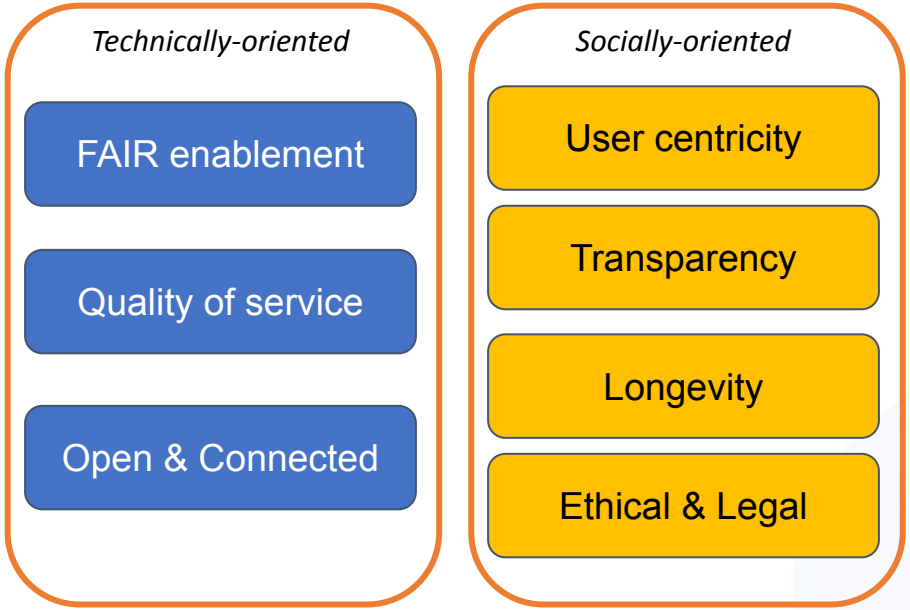
M2.10 Report on basic framework on FAIRness of services", available at: https://doi.org/10.5281/zenodo.4292599

D2.7 Framework for assessing FAIR Services



D2.7 Framework for assessing FAIR Services: https://doi.org/10.5281/zenodo.5336234

7 Aspects for FAIR service assessment



Each aspect has a high-level objective with a series of recommendations

High-level objective

Actionable, detailed recommendations

Aspect: SAF-O Open & Connected		
Objective: The service is operated in a low-barrier and inclusive way, seeking integrations and connections with other services and championing principles of openness consistent with Open Science and Open Research.		
Identifier	**Recommendation**	**Priority**
SAF-O-1	Publish clear, inclusive and non-discriminatory licences and/or terms of use. Enable wide access to the service.	Essential ☆☆☆
SAF-O-2	Seek integrations with other services rather than replicating functionalities, especially for common reusable infrastructure components. Provide documentation to ensure better sustainability for the network of integrations. Adopt the EOSC architectural components and standards as enablers for deep interoperability with other services in the EOSC portfolio[6].	Essential ☆☆☆

Each recommendation has a priority level

SAF-U-4	Strive for continual improvements to the user experience. In addition to making use of data and service usage statistics, actively work with the community to understand and improve usability, for example through user tests or design studios.	Essential ☆☆☆
SAF-U-5	Determine and monitor your target user community to understand how the service fits within its data management norms and expectations.	Important ☆☆
SAF-U-6	Ensure that there is an ongoing, consistent dialogue between the service and its user community, such that users can optimally make use of the service and influence its development.	Important ☆☆
SAF-U-7	Include multi-lingual support and accessibility features[9], both for the service and its documentation, to the extent relevant for the service's (potential) user base. Key information must be available in English if the service is intended to be included within EOSC.	Useful ☆

FAIRSFAIR
Fostering Fair Data Practices in Europe

Using the framework

Use cases

KIELIPANKKI
The Language Bank of Finland

Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

Applying SAF-F-2 on SURF PID Service

Aspect's objective	The service enables FAIR data by elevating the FAIRness of digital objects and/or supporting the FAIRification process. FAIR enablement is actively driven through the implementation of community-supported standards and interoperability frameworks.
Recommendation SAF-F-2	Engage with both the user community and other service providers to improve interoperability between services. Of particular attention here are authentication and authorization infrastructure (AAI), PIDs, and data and metadata encoding specifications. Seek alignment with existing or emerging data type registries and interoperability frameworks, e.g. the EOSC interoperability framework.
Service	The SURF Persistent Identifier Service, provides and hosts prefixes and PIDs for users
Implementation	The SURF EPIC PIDs are based on the services and software provided by the European Persistent Identifier Consortium, EPIC (http://www.pidconsortium.eu). The EPIC PID system is in turn based on another handle system, known simply as the "Handle System" (www.handle.net). This system is used globally on the internet and is what makes your data accessible via the world wide web.
Documentation (if available)	

SAF-F-2- Essential ☆☆☆

Technically-oriented

Open & Connected

Applying SAF-O-1 on The Language Bank of Finland

Aspect's objective	The service is operated in a low-barrier and inclusive way, seeking integrations and connections with other services and championing principles of openness consistent with Open Science and Open Research.
Recommendation SAF-0-1	Publish clear, inclusive and non-discriminatory licences and/or terms of use. Enable wide access to the service.
Service	The Language Bank of Finland is a service for researchers using language resources across digital humanities and social sciences.
Implementation	The licences, terms of use etc. are all published in the Language Bank Portal.
Documentation (if available)	https://www.kielipankki.fi/corpora/

SAF-O-1- Essential ☆☆☆

Technically-oriented

Open & Connected

Applying SAF-O-3 on The Language Bank of Finland

Aspect's objective	The service is operated in a low-barrier and inclusive way, seeking integrations and connections with other services and championing principles of openness consistent with Open Science and Open Research.
Recommendation SAF-O-3	Make the service and all documentation available online through URLs that are fully qualified domain names and assign PIDs where applicable.
Service	The Language Bank of Finland is a service for researchers using language resources across digital humanities and social sciences.
Implementation	The Portal contains all of the Language Bank's documentation, including user guides and other support material.
Documentation (if available)	https://www.kielipankki.fi/support/

SAF-O-3- Essential ☆☆☆

Applying SAF-U-1 on DMPonline

Aspect's objective	The service is managed so that it serves the (possibly evolving) goals of the user community and maximises usability while minimizing burden.
Recommendation SAF-U-1	Ensure the service provider organization has adequate support staff available to assist users where needed.
Service	DMPonline is a web-based tool that supports researchers to develop data management and sharing plans.
Implementation	DMPonline has a dedicated member of staff to deal with customer relations and client queries. The DMPonline helpdesk is covered 9 - 17 British time on working days.
Documentation (if available)	See team at: https://www.dcc.ac.uk/dmponline

SAF-U-1- Essential ☆☆☆

Socially-oriented

User centricity

Applying SAF-U-4 on DMPonline

Aspect's objective	The service is managed so that it serves the (possibly evolving) goals of the user community and maximises usability while minimizing burden.
Recommendation SAF-U-4	Strive for continual improvements to the user experience. In addition to making use of data and service usage statistics, actively work with the community to understand and improve usability, for example through user tests or design studios.
Service	DMPonline is a web-based tool that supports researchers to develop data management and sharing plans.
Implementation	DMPonline runs usability testing to gather feedback on existing features and how the system could be improved. New features are run past users during the development phase to gather input and in some cases, new designs are developed in collaboration with CDL's UX team to ensure best practice.
Documentation (if available)	Poster reporting latest usability testing results: https://doi.org/10.5281/zenodo.5795390

SAF-U-4- Essential ☆☆☆

Applying SAF-T-1 on the SWH archive

Aspect's objective	The service provider communicates with its stakeholders in a transparent manner.
Recommendation SAF-T-1	Clearly communicate the service's core value proposition and any pertinent (technical or non-technical) features, as well as its limitations.
Service	Software Heritage (SWH) is the software source code universal archive with several features (software source code save code now, deposit, download, intrinsic PIDs and more)
Implementation	SWH is clearly communicating through its website the core values of the universal source code archive mission. The approach of SWH is communicated on the website, it is an open source project which was founded by the french research centre Inria and is endorsed by the open source community and many sponsors from different domains
Documentation (if available)	Values: https://www.softwareheritage.org/mission/ Features: https://www.softwareheritage.org/features/

SAF-T-1- Essential ☆☆☆

Applying SAF-T-4 on the SWH archive

Aspect's objective	The service provider communicates with its stakeholders in a transparent manner.
Recommendation SAF-T-4	Be clear about how the service implements community standards.
Service	Software Heritage (SWH) is the software source code universal archive with several features (software source code save code now, deposit, download, intrinsic PIDs and more)
Implementation	The different communities are identified on the SWH website and for each there is a dedicated page. Also for researchers there is a dedicated howto guide about using SWH to archive and reference code.
Documentation (if available)	https://www.softwareheritage.org/community/developers/

SAF-T-4- Important ☆☆

Socially-oriented

Longevity

Applying SAF-L-1 on SURF Data Archive

Aspect's objective	The service provider designs the service with a timeframe for the maintenance and sustainability of the service in mind and implements measures accordingly, considering the researchers' need for reproducible research.
Recommendation SAF-L-1	Take reasonable measures to ensure sustainable long-term operation — including both financial and organisational aspects. Aim to reduce long-term operational dependencies on short-lived project funding. If available, provide clear information to indicate how long the service will minimally be available and maintained.
Service	SURF Data Archive is a tape-based solution for long-term storage of research data
Implementation	The underlying infrastructure of the Data Archive is, for the most part, structurally subsidized by SURF through NWO, the national Research Council of the Netherlands as part of their budget for support for infrastructure. The sustainability of the financial funding is ensured NWO funding of the infrastructure and support and a certain amount of usage and charging the marginal price of adding extra data for the rest of the volume. The contracts for the Data Archive are for at least 10TB and at least 5 years, reflecting this.
Documentation (if available)	

SAF-L-1- Essential ☆☆☆

FAIRSFAIR
Fostering Fair Data Practices in Europe

Recommendations and lessons learned

The framework is useful to

- Check your own documentation and processes
  - Start discussions if more of your documentation could be made publicly available

- Getting better at becoming FAIR, it’s for self-improvement

The framework is not

- Something you should use for externally assessing a service
- A show-off tool
  - You could publish a blog post or roadmap for any plans to become more FAIR-enabling to be transparent with the users of your service
- Something that can be turned into a machine-actionable assessment

Sustainability

- We highlighted the framework to the following EOSC Association task forces for their attention:
  - Rules of Participation and Compliance Monitoring TF
  - FAIR Metrics and Data Quality TF
  - Semantic Interoperability TF

- The work is openly licenced and can be taken forward by anyone interested.

Acknowledgements

- Sara Ramezani (SURF; task lead)
- Morane Gruenpeter (INRIA)
- Patricia Herterich (DCC)
- Rob Hooft (DTL)
- Tero Aalto (CSC)

And earlier contributions from:

- Hylke Koers (SURF; former task lead)
- Christine Staiger (DTL)
- Roberto Di Cosmo (INRIA)
- Sarah Jones (DCC)
- Jessica Parland-von Essen (CSC; work package lead)
- Jonas Tana (CSC)
- Hanna Koivula (CSC)