

# Report on "FAIR Signposting" and its uptake by the community

Mark D Wilkinson<sup>1,3</sup>, Susanna-Assunta Sansone<sup>2,4</sup>, Marjan Grootveld<sup>2,5</sup>, Richard Dennis<sup>2,6</sup>, David Hecker<sup>2,7</sup>, Robert Huber<sup>8</sup>, Stian Soiland-Reyes<sup>9</sup>, Herbert Van de Sompel<sup>5</sup>, Andreas Czerniak<sup>10</sup>, Milo Thurston<sup>4</sup>, Allyson L. Lister<sup>4</sup>, Alban Gaignard<sup>11</sup>

<sup>1</sup> Co-Chair, EOSC Task Force on FAIR Metrics and Data Quality

<sup>2</sup> Member, EOSC Task Force on FAIR Metrics and Data Quality

<sup>3</sup> Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Centro de Biotecnología y Genómica de Plantas (CBGP), Universidad Politécnica de Madrid (UPM), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) & Consejo Superior de Investigaciones Científicas (CSIC), Pozuelo de Alarcón, Madrid, ES, Spain. 0000-0001-6960-357X

<sup>4</sup> Oxford e-Research Centre, Department of Engineering Science, University of Oxford, UK. 0000-0001-5306-5690 (Susanna-Assunta Sansone), 0000-0002-7702-4495 (Allyson L. Lister), 0000-0002-6468-9260 (Milo Thurston)

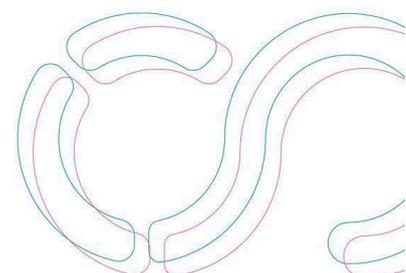
<sup>5</sup> Data Archiving and Networked Services (KNAW-DANS), The Netherlands. 0000-0002-2789-322X (Marjan Grootveld), 0000-0002-0715-6126 (Herbert Van de Sompel)

<sup>6</sup> Novo Nordisk Foundation Center for Stem Cell Medicine - reNEW, the University of Copenhagen, Copenhagen, Denmark, 0000-0002-4472-7194

<sup>7</sup> German Aerospace Centre, Research Data Management, Germany, 0000-0003-3836-2800

**EOSC Association AISBL**

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
+32 2 537 73 18 | [info@eosc.eu](mailto:info@eosc.eu) | [www.eosc.eu](http://www.eosc.eu)  
Reg. number: 0755 723 931 | VAT number: BE0755 723 931



<sup>8</sup> MARUM - Center for Marine Environmental Sciences, University of Bremen,  
Germany, 0000-0003-3000-0020

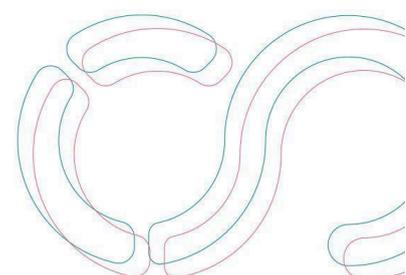
<sup>9</sup> Department of Computer Science, The University of Manchester, UK;  
Informatics Institute, University of Amsterdam, NL  
0000-0001-9842-9718

<sup>10</sup> Bielefeld University Library, Germany, 0000-0003-3883-4169

<sup>11</sup> Nantes Université, CNRS, INSERM, L'Institut du Thorax, F-44000 Nantes,  
France, 0000-0002-3597-8557

\* Corresponding author

\*\* All TF members have had the opportunity to review and edit this document;  
invitations to be on the authorship list were open to all hackathon participants

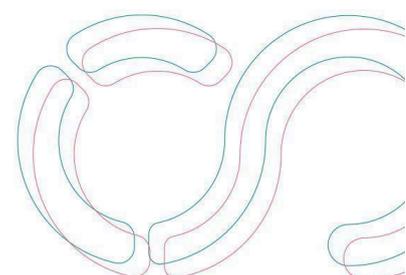


## Executive Summary

The FAIR Metrics subgroup of the EOSC Task Force on FAIR Metrics and Data Quality is dedicated to scrutinizing the adoption and impact of FAIRness metrics and the practicality of FAIRness evaluations. Previous reports have highlighted the inconsistency in results when the same digital object is evaluated by different tools, attributing these disparities to varied interpretations of the Metrics, metadata publishing practices, and the fundamental objectives of FAIR principles. The authors of this report, representing the FAIR Metrics subgroup, have facilitated six workshops and hackathon-style gatherings, convening diverse FAIR assessment stakeholders—including tool developers, standards and repository experts, and interoperability specialists.

The initial workshops, as outlined in an earlier report endorsed by EOSC, pinpointed a metadata publishing design pattern known as “FAIR Signposting.” This approach offers a transparent, compliant, and straightforward mechanism for guiding automated agents through metadata spaces to locate three essential FAIR elements: the globally unique identifier (GUID), the data records, and the corresponding metadata records. Furthermore, these sessions led to the development of a paradigm for creating reference environments. These environments serve as benchmarks for evaluating the compliance of metadata harvesters with FAIR Signposting criteria and for standardizing the metadata harvesting process of FAIR assessment tools.

This report provides an updated overview of the recent progress achieved by the FAIR Metrics subgroup, encapsulating the advances from the latest two hackathon events. We outline the progress towards developing an integrated suite of FAIR assessment tests, which are now becoming standardized across various FAIR



assessment instruments. Additionally, we present preliminary insights into the community's adoption of these practices and report on the ongoing effort to establish a FAIR testing governance body, along with strategies to ensure its enduring viability.

We, the authors, advocate for the EOSC Association and the EOSC Task Force on Long-term Data Preservation to endorse FAIR Signposting, along with the outcomes of our endeavours, as formal recommendations for EOSC-linked resources and EOSC-funded initiatives. We aim to promote an EOSC infrastructure where FAIR-enabling services can be assessed with clarity and uniformity. This will also ensure that tools designed for the EOSC will function with a standardized set of expectations applicable to all stakeholders - providers and users.

## Background

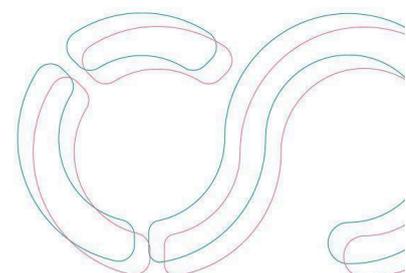
In our earlier reports<sup>1,2</sup>, we made several observations about the landscape of FAIR assessment, including that:

- FAIR compliance is currently “stuck” between being an increasingly common research and publishing requirement while remaining an unmeasurable set of ideals.
- Assessment tools are increasing, with at least 28 assessment platforms<sup>3</sup> without overarching guidance. This leads to more confusion in the community as the assessment platforms continue giving divergent results.

<sup>1</sup> Mark D. Wilkinson et al. (2022): **FAIR Assessment Tools: Towards an "Apples to Apples" Comparisons**. EOSC Metrics and Data Quality Task Force. <https://doi.org/10.5281/zenodo.7463421>

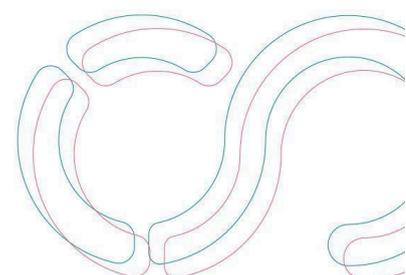
<sup>2</sup> Mark D. Wilkinson et al. (2022): **Community-driven Governance of FAIRness Assessment: An Open Issue, an Open Discussion**. EOSC Metrics and Data Quality Task Force. <https://doi.org/10.5281/zenodo.7390482>

<sup>3</sup> <https://fairassist.org/>



- Although FAIR assessment tools should consider different aspects of FAIR implementations to facilitate comparability of results, they should be able to support a selection of standard baseline metrics.
- The latest FAIR-IMPACT FAIR Assessment Challenges have shown that developers and publishers can be motivated very well with gamification methods to improve their FAIRness. At the same time, however, this can also lead to selecting the tool that generates the most favorable score, which can frustrate the ability to compare resources to one another consistently.
- The problem of metadata harvesting is a clear source of inconsistency, given the wide range of “acceptable” metadata publishing practices. This includes HTML-embedded metadata in various syntaxes and using varying technologies, content negotiation, typed links, link headers, and a wide range of vocabularies and schemas and inconsistent guidelines around these.
- In particular, “boutique” repositories (versus well-established specialist or generalist repositories), where the resource authors often have less experience in metadata publishing norms and less experience in correctly executing existing standards or paradigms, need clear, authoritative guidance on achieving high FAIR compliance quickly.

The problem of inconsistency between FAIR assessment tools is a wicked problem, as it requires that they behave identically when faced with an unpredictable world of conflicting and often improperly implemented attempts to “be FAIR.” For this reason, the FAIR Metrics subgroup of the EOSC Task Force (TF) concluded that it would be better for EOSC and of benefit to the broader community if there were a recommended approach to provisioning FAIR (meta)data for resources that wish to be recognized as participating in the Open Science Cloud. While the FAIR Principles



neither suggest nor recommend any technology or paradigm, this does not prevent individual communities from establishing recommendations, expectations, norms, or even policies that harmonise their participants and improve interoperability between them; in fact, this is recognized as an expectation in FAIR Principle R1.3, where FAIR publishers are encouraged to follow community standards.

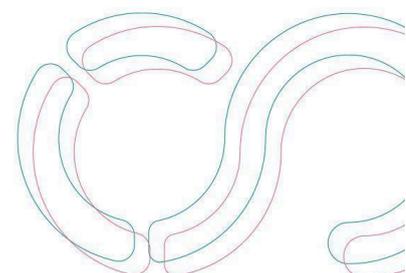
With this in mind, the final two hackathon events focused on utilising, exclusively, the FAIR Signposting<sup>4</sup> metadata publication design patterns (described below) for their metadata harvesting, using this metadata as the substrate upon which the participants jointly designed a set of generic FAIR tests. This included clearly defined expectations for what a data publisher should consider a success versus a failure. These expectations were then codified into a new set of benchmark Web pages; such reference environments can circumscribe and validate the behaviour of FAIR assessment tools authored by any stakeholder, both contemporary and future.

## Overview of FAIR Signposting

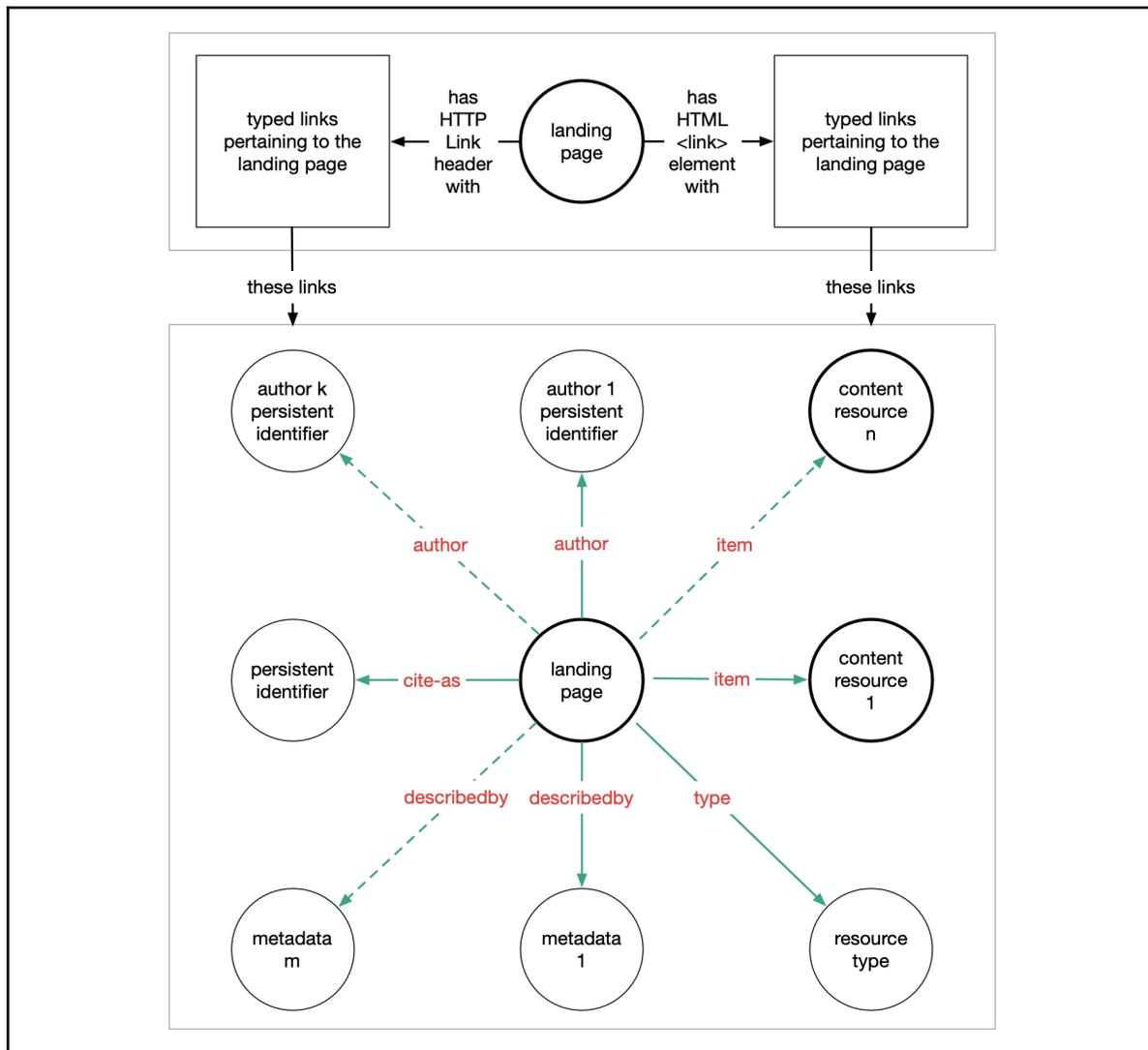
FAIR Signposting uses a well-established Web technology - typed links - to guide an automated agent through the digital object's record, including its canonical identifier, and (optimally) its metadata and data entities. It utilizes landing pages as the primary source of signposts, where the landing page contains pointers to the three critical information facets - GUID, data, metadata - an agent requires. Importantly, however, signposts can appear both as HTML Link elements, or as HTTP Link Headers; the latter allows them to be used to guide an agent from a data object (e.g. a video or workflow) to the landing page and/or directly to metadata about that

---

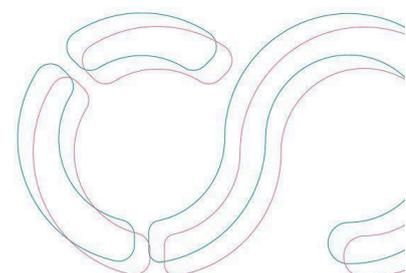
<sup>4</sup><https://signposting.org/FAIR>



digital object. A diagram showing how FAIR Signposts are used within a landing page is provided in Figure 1. The result is a set of pointers that can be unambiguously interpreted by a harvesting agent, and requiring no “guesswork” by that agent v.v., for example, content-negotiation.



**Figure 1: A FAIR Signposting landing page.** The HTTP Link Headers, or the HTML Link elements, contain a set of typed links pointing to the primary information features of a digital object, including its metadata records (via describedby links), its data records (via item links) and the canonical GUID of the landing page itself, for example, its DOI (via cite-as links). (Image taken from <https://signposting.org/FAIR>)



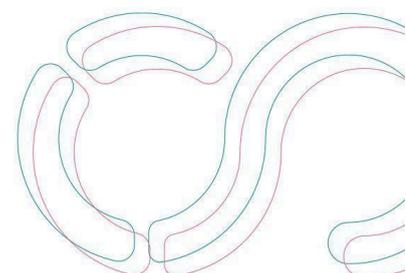
## Hackathons 3 and 4

The final two hackathon events were held virtually, on Sept 27 and 29, 2023, using the virtual conferencing software “Remo”. Participants were invited using the mailing list from the prior events, with an invitation to all recipients to forward the invitation to any other potentially interested parties. As expected, given that it was a more technically focused meeting, the overall attendance was lower than at previous events, where more open discussions were the focus. In attendance were:

Name	Affiliations	FAIR assessment project or tool
Mark Wilkinson	EOSC TF Co-chair; Universidad Politécnica de Madrid	FAIR Evaluator & FAIR Champion
Susanna-Assunta Sansone	EOSC TF Member; University of Oxford; ELIXIR Interoperability Platform	FAIRsharing
Stian Soiland-Reyes	The University of Manchester, University of Amsterdam, RO-Crate chair, ELIXIR	Signposting Reference Environment Signposting Python library
Milo Thurston	University of Oxford	FAIRsharing
Herbert Van de Sompel	DANS	FAIR Signposting
Andreas Czerniak	Bielefeld University Library, Germany	OpenAIRE FAIR validator
Wilko Steinhoff	DANS	EOSC SYNERGY & FAIRCORE4EOSC
Robert Huber	EOSC TF Member; Universität Bremen	F-UJI, FAIR-IMPACT

### EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
 +32 2 537 73 18 | info@eosc.eu | www.eosc.eu  
 Reg. number: 0755 723 931 | VAT number: BE0755 723 931

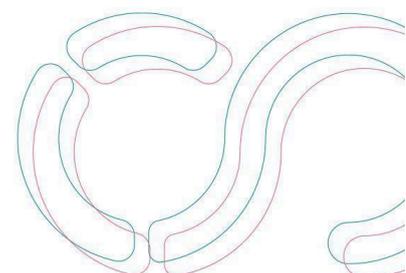


Allyson Lister	University of Oxford	FAIRsharing
Alban Gaignard	CNRS, Nantes University, ELIXIR-FR	FAIR Checker
Richard Dennis	EOSC TF Member	Observer

Each event lasted from 9 AM to 5 PM, with an opening discussion to set the tasks for the day and a final hour of updates and discussion. Participants would occasionally return to the virtual meeting space to discuss issues that required coordination. The remaining time the participants spent coding and running tests or reference environments. At this event, we focused on the “Findability” tests. We began exploring tests for “Reusability.” In each case, we examined the corresponding Maturity Indicators from the Research Data Alliance (RDA) FAIR Maturity Model<sup>5</sup>, as well as the implementations of the automated assessment platform authors. This included selecting standards for what should be considered a “pass” versus a “fail” for a given test. It also became clear that it was necessary to have a category of test result that was “undetermined” - this would be applicable in cases where, for example, a community claimed the existence of a community standard that they were following, but that standard could not be found in a standards registry such as the RDA-endorsed FAIRsharing<sup>6</sup>. An “undetermined” result would warn the community that they should find a suitable standards repository or registry (for example, FAIRsharing, or OBO), and register this as a standard; however, this result does not imply failure.

<sup>5</sup> FAIR Data Maturity Model Working Group (2020): **FAIR Data Maturity Model**. Specification and Guidelines. *Research Data Alliance*. <https://doi.org/10.15497/RDA00050>

<sup>6</sup> <https://fairsharing.org/>



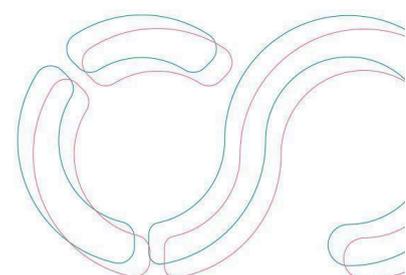
## The “F” Metrics

Considerable time was spent exploring the “Findability” Metrics. Issues included, for example, the selection of which GUID schema would be acceptable to the testing platform, agreement on the regular expression that would be used to match those schemas, discussion of the use of FAIRsharing<sup>7</sup> as a way to validate such schema, to avoid hard-coding GUID schemas into the assessment tools; and discussion of whether there should be secondary confirmation that a GUID was valid (e.g., via lookup).

Perhaps the most critical conversation regarding the “F” metrics was related to the meaning of “Globally Unique Identifier” in the context of the Web in general and FAIR Signposting in particular. Specifically, the FAIR Principles call for a globally unique and (usually) resolvable identifier for all data and metadata elements. In Signposting, identifiers resolve to pages containing typed links, pointing to metadata and data. The uniqueness of an identifier in the context of a Signposting-enabled network, therefore, is the combination of the identifier itself (generally the HTTP URI rendition of a PID) and the “path” that is followed by the agent to arrive at the metadata or data element it desires to resolve; the selection of the desired path is further refined by the “type” attribute and “profile” attribute that is included in the Signposting Link. This is a somewhat distinct interpretation of the meaning of an “identifier”; however, these Signposting paths are globally unique, transparent, unambiguous, and resolvable, thus consistent with the F Principles. Further work is however needed for recording HTTP-level provenance for a client following content negotiation using Signposting, although we note that RFC9264 Linksets are interpretable as JSON-LD.

---

<sup>7</sup> [https://fairsharing.org/search?recordType=identifier\\_schema](https://fairsharing.org/search?recordType=identifier_schema)



Here, we will provide just one example conclusion from these discussions; others will be published in the working document, associated reference environments in the benchmarks, and as records of FAIR metrics standards in FAIRsharing<sup>8</sup>.

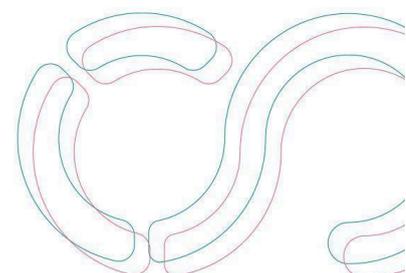
### Example: FAIR Signposting-based tests for FAIR Principle F1: (Meta) data are assigned globally unique and persistent identifiers

Hackathon participants decided that the four RDA MI Metrics for Principle F1 could be merged into a single assessment; this is a consequence of the Signposting path-based definition of a GUID (described earlier), thus providing the same GUID for both the metadata and the data. In addition, persistence required by F1 means that GUIDs also have to be PIDs, thus to comply with the EOSC PID policy<sup>9</sup>. The details of this decision were:

- Based on being both globally unique and persistent, the GUID/PID types that would be acceptable were DOI, Handle, w3id, Purl, Ark, and URNs (including LSID).
- These types would be matched using Regular Expressions shared between the FAIR assessment platforms.
- The participants decided that these tests were non-determinative in that the commitment to persistence cannot be objectively measured; persistence is a promise rather than an integral behaviour from using a specific technology. As such, the output of the tests will be “pass” when it is one of the six selected GUID types or “undetermined” if it follows any other schema.

<sup>8</sup> <https://fairsharing.org/search?fairsharingRegistry=Standard&recordType=metric&page=1>

<sup>9</sup> European Commission, Directorate-General for Research and Innovation, Hellström, M., Heughebaert, A., Kotarski, R. et al. (2020): **A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC)**, Publications Office, <https://doi.org/10.2777/926037>



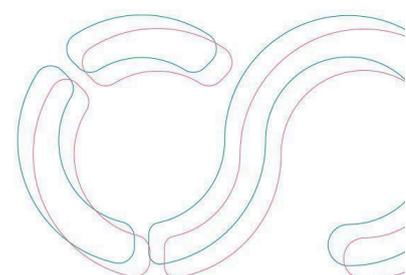
- Hard-coded lists of acceptable identifier types and their regular expressions were recognized as a sub-optimal approach. A solution was proposed by FAIRsharing participants who agreed to create facets for the above list of GUIDs in their registry’s metadata model to hold features such as persistence policies and GUID regular expressions, such that they can be looked up by the test rather than being hard-coded.
- All of these decisions were captured in a series of benchmark environments identified as:
  - <https://w3id.org/a2a-fair-metrics/50-rda-f1-01m-t1-metadata-pid/>
  - <https://w3id.org/a2a-fair-metrics/51-rda-f1-01m-t1-metadata-no-pid/>
  - <https://w3id.org/a2a-fair-metrics/60-rda-f1-01md-t1-citeas-pid//>
  - <https://w3id.org/a2a-fair-metrics/61-rda-f1-01md-t1-citeas-no-pid/>
  - <https://w3id.org/a2a-fair-metrics/52-rda-f1-01d-t1-data-pid/>
  - <https://w3id.org/a2a-fair-metrics/53-rda-f1-01d-t1-data-no-pid/>

The other F Principles and the R1 principle went through similar processes of detailed exploration, resulting in similar sets of decisions and the publication of associated benchmark environments (now totaling 75<sup>10</sup>) to ensure convergence of FAIR assessment tools into the future.

As a next step for the TF participants, in collaboration with other stakeholders such as FAIR-IMPACT, we wish to generate a detailed list of ‘Signposting FAIR Metrics’ that we can recommend to the EOSC-A based on the findings described above and the benchmark tests published at GitHub.

---

<sup>10</sup> <https://w3id.org/a2a-fair-metrics/>



## Comparison of Signposting with the FDO Initiative

The EOSC Interoperability Framework<sup>11</sup> has promoted the FAIR Digital Object (FDO) concept as a set of principles and requirements<sup>12</sup> to improve the machine-actionability of research outputs, with a strong focus on persistent identifiers and definitions. While the FDO concept can be realised in several ways<sup>13</sup>, including a “classic” Digital Object approach that is entirely based on the Handle system, Signposting has been highlighted<sup>14</sup> as a lightweight Web-based approach<sup>15</sup> to implement FDO<sup>16</sup>(Figure 2).

---

<sup>11</sup> European Commission, Directorate-General for Research and Innovation, Corcho, O., Eriksson, M., Kurowski, K. et al. (2021): **EOSC interoperability framework** – Report from the EOSC Executive Board Working Groups FAIR and Architecture, Publications Office <https://doi.org/10.2777/620649>

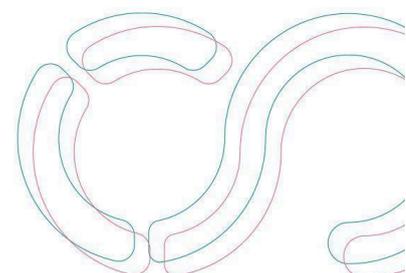
<sup>12</sup> Ivonne Anders et al. (2023): **FDO Forum FDO Requirement Specifications**. *FDO Forum*. PR-RequirementSpec-3.0. George Strawn, Peter Wittenburg (eds.) <https://doi.org/10.5281/zenodo.7781925>

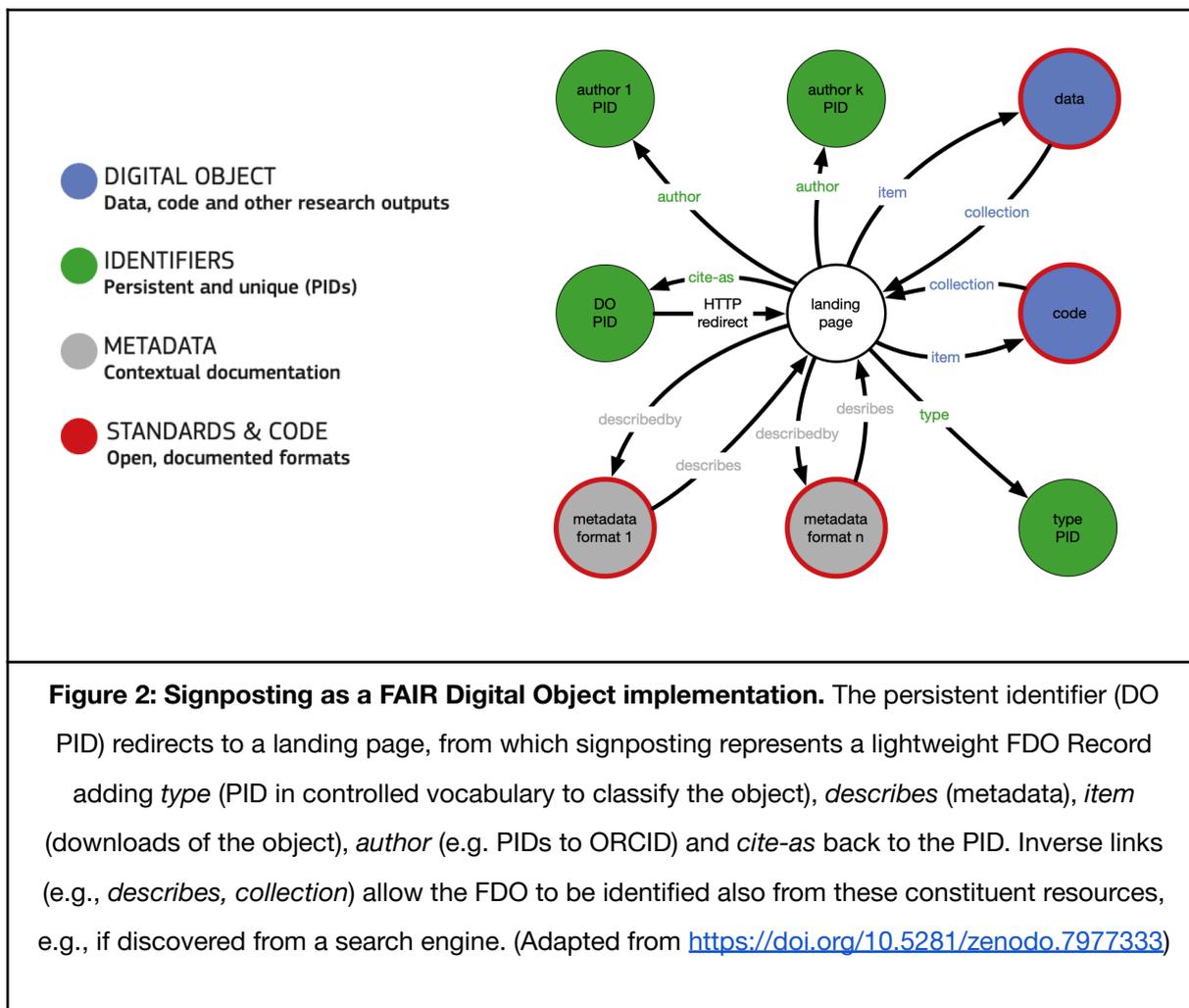
<sup>13</sup> European Commission, Directorate-General for Research and Innovation (2018): **Turning FAIR into reality** – Final report and action plan from the European Commission expert group on FAIR data, Publications Office <https://doi.org/10.2777/1524>

<sup>14</sup> Herbert van de Sompel (2022): **FAIR Digital Objects and FAIR Signposting**. *Zenodo* <https://doi.org/10.5281/zenodo.7977333>

<sup>15</sup> Stian Soiland-Reyes, Leyla Jael Castro, Daniel Garijo, Marc Portier, Carole Goble, Paul Groth (2022): **Updating Linked Data practices for FAIR Digital Object principles**. *Research Ideas and Outcomes* 8:e94501 <https://doi.org/10.3897/rio.8.e94501>

<sup>16</sup> Felix Bach et al. (2023): **Current Developments in the Research Data Repository RADAR**. *ARPHA Preprints*. <https://doi.org/10.3897/arphapreprints.e109429>



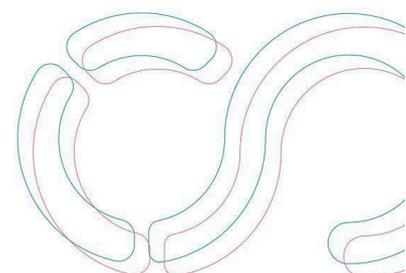


In this approach, the FDO’s “PID Record” is mainly a navigable data structure where the actual details of the digital object are described in separate metadata resources. It is, therefore, essential to not just have syntactic type information on the possible metadata (e.g., RDF in JSON-LD format) but also *profiles* of their use (e.g., RO-Crate<sup>17</sup> or Bioschemas<sup>18,19</sup>).

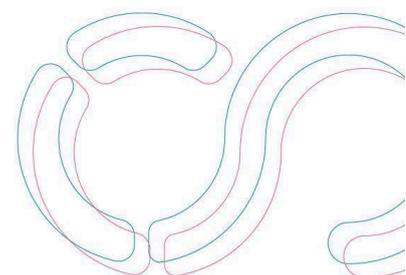
<sup>17</sup> <https://w3id.org/ro/crate>

<sup>18</sup> <https://bioschemas.org/profiles/>

<sup>19</sup> <https://doi.org/10.1371/journal.pcbi.1011120>



While being consistent with the FDO principles and requirements, there are several core differences between the FDO approach promoted by the FDO Forum (FDOF) and Signposting (SP). For example, SP is deliberately Web-based, thus it can leverage widely held skills and well-established and pervasive technologies, standards, tools, and registries (e.g. for media types), and these do not need to be designed from-scratch. SP, being a standards-based Web technology, is also decentralized, while the FDOF approach is centralized in the sense that it makes all interactions with DOs dependent on information (including regarding possible "operations" on DOs) that is stored in the central PID system. In contrast, SP is decentralized and expects the authoritative information to be provided by the custodian of the DO. This also has a consequence of FDOF being prone to single point of failure problems v.v. the PID system and, moreover, requires continuous synchronization of DO-related information from the custodian to the central node. SP limits itself to read-only interoperability and does not consider other types of operations because there are other Web standards (e.g. WebDAV, Linked Data Platform, Fedora API) that cover these aspects. In that sense SP adheres to a modular interoperability approach: rely on several compliant building blocks, each of which address a specific requirement. This is in contrast to FDOF, which covers the entire CRUD (create, read, update, delete) realm and in that sense is holistic. It is unclear, however, how other interoperability functionality regarding DOs, for which Web standards exist (e.g. annotation), would be addressed in the FDOF world view. By the FDO principles, structural elements such as type definitions and profiles should be FDOs. This reflects on the FAIR principle I2 (*meta*)data use vocabularies that follow FAIR principles. Nevertheless, the FDO community also considers how the existing FAIR vocabularies and persistent identifiers can co-exist with more stringent FDO representations. FAIR Signposting (a Signposting implementation guideline with precise instructions regarding link types, link cardinalities, etc.) could



be considered a non-intrusive overlay that can bridge the gap and make the otherwise informal FAIR conventions become explicit and give predictable machine actionability. Finally, typed links are already widely used within the global repository community, thus implementation of SP will generally require no new skills within the development team, and only modest code changes<sup>20</sup>

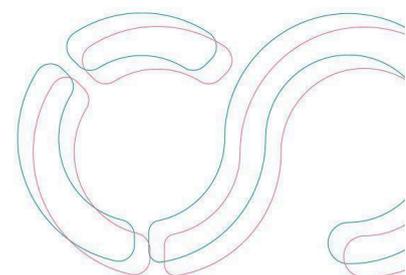
## Community Uptake of FAIR Signposting

The overarching mission of the FAIR principles and the EOSC is to streamline the accessibility and reusability of the vast and varied data pool. Given this diversity, providing detailed metadata by data sources enables effective data discovery and comprehension. Unlike specialized repositories, which define their data purpose, semantics, and structure, generalist repositories use depositors to characterize their data, making source-provided metadata even more pivotal. Yet, the motivation to furnish such metadata is minimal, as its utility is often only realized after downloading the data.

The FAIR Signposting framework offers a solution, presenting generalist repositories with a transparent methodology to reference source-provided metadata for the first time explicitly. This allows computational agents to facilitate data discovery and reuse more effectively. The widespread adoption of FAIR Signposting would thus serve the broader community and the EOSC substantially—not only enhancing the homogeneity of FAIR assessments but also empowering the discovery and interpretation of metadata before the data download.

---

<sup>20</sup> Personal communication from Dataverse developers who implemented FAIR Signposting in version 5.14 of the Dataverse software



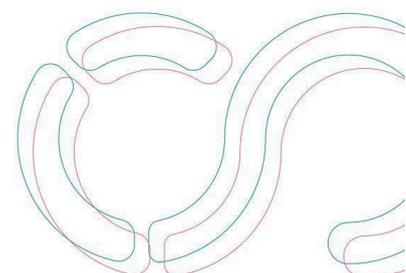
Shortly after the hackathon, TF members reached out to the US National Institute of Health (NIH) Generalist Repository Ecosystem Initiative (GREI)<sup>21</sup>, which represents the major global generalist repositories such as Dataverse, Zenodo, Figshare, OSF, Vivli, Mendeley Data and Dryad. We requested an opportunity to present the FAIR Signposting ideas to them, to explain the benefits, and to explore the costs. Concerning the costs, there seem to be two critical Signposting behaviours that will have different costs for implementation by the GREI community:

1. **Explicit references to the repository-provided metadata and the metadata provided by, e.g., DataCite, and detailed references to the records in the deposit.** In these cases, the links are already known by the repository. They can easily be added to the existing landing pages (in some cases, they are already provided by the repository as links, but using links for this purpose could be more consistent). This will likely be a low cost for implementation.
2. **Explicit references to user-provided metadata.** In most cases, the generalist repositories will need to provide a novel mechanism for distinguishing metadata records from data records in the deposit. Thus, while Signposting would allow those metadata to be discovered and used, it will likely require a more significant repository infrastructure re-design. Therefore, enabling this behaviour will come at a higher cost.

Concerning (1) above, the TF learned that Dataverse and Zenodo had already started implementing FAIR Signposting at this level. In the case of Dataverse, the currently deployed public codebase includes it while it is on the deployment path for Zenodo. The other repositories indicated it was also on their implementation pathway. Given adequate time and resources, they expressed considerable interest

---

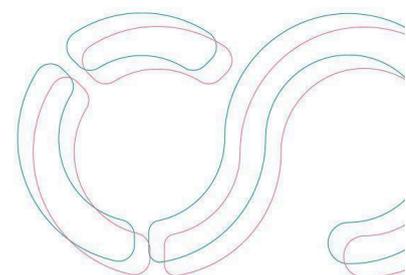
<sup>21</sup> <https://datascience.nih.gov/data-ecosystem/generalist-repository-ecosystem-initiative>



and enthusiasm to do so, which will depend on the repository developers' prioritisation. Concerning (2), several GREI members indicated that support for user-provided metadata was on their development pathway and that a clear standard for referring to it (i.e., via Signposting) was helpful in these efforts.

Knowing that Dataverse had implemented Signposting before the start of the two hackathon events, TF participants used the FAIR Signposting benchmarks created at the earlier events to test their implementation. Several “bugs” were discovered and reported to the Dataverse developers, which were rapidly corrected. This provides anecdotal evidence that the hackathon events are producing assessment tooling that is effective and useful. With this in mind, we have offered to provide GREI participants full (early) access to these benchmark tests and instructions on conducting self-assessment to assist their adoption of the specification.

The FAIR-IMPACT project launched a [support action](#) for *Enabling FAIR Signposting and RO-Crate for content/metadata discovery and consumption*, where 15 funded organisations and individuals across Europe are participating, including adding Signposting to eurac's [Environmental Data Platform](#), University of Novi Sad's [CRIS UNS](#), [InvenioRDM](#), DRI, FAIRPoints, FAIRDOM-SEEK, and CKAN. Several of these participants were Dataverse users and have now deployed and are testing the latest Signposting support; others were new to FAIR and are making their first structured metadata representations in RO-Crate while deploying Signposting. Overall, the feedback shows that it is relatively easy to retrofit Signposting if the data provider already has the other ingredients (e.g., RDF metadata or DOIs). However, for newcomers, the learning curve of such elements of the FAIR stack can still be daunting. It was tempting for them to add instead additional custom signposting (e.g., *rel="https://schema.org/affiliation"* from an author to an organisation), which we accept is valid but advise against.



Beyond the repository community, we also see the adoption of FAIR Signposting by tool-builders. For example, a Chrome extension has recently been published that will report when it discovers Signposting metadata in a Web page<sup>22</sup>. In addition, Signposting was a theme of a project at the BioHackathon Europe 2023 event<sup>23</sup>, where efforts were made to improve the tooling around Signposting and find convergence points between Signposting and the FAIR Digital Object specifications. A more complete list of Signposting adopters is available<sup>24</sup>.

## TF Outreach & Emergent Models for Governance

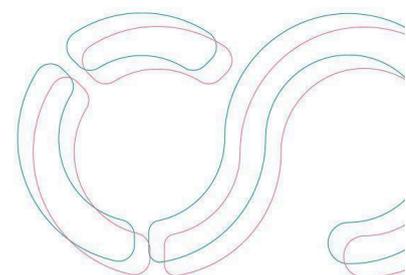
Shortly after the publication of our earlier report to EOSC on FAIR Assessment governance<sup>2</sup>, the FAIR IMPACT project authored a response document<sup>25</sup>, which was in part supportive but expressed some additional concerns and disagreements. Given that the objective of the latest two hackathon events was to engender harmonization between the assessment tools, the participants also took time to “observe the process.” In particular, the mechanisms for reaching agreement on the appropriate metrics, their associated testing code, quality assurance around these, and how this could be replicated in the future as more specialized (e.g., domain-specific) FAIR metrics are designed. The objective was to gather evidence

<sup>22</sup> Sandy Rogers (2023): **Signposting.org browser extension**. GitHub, <https://github.com/SandyRogers/signposting-chrome-extension>

<sup>23</sup> Stian Soiland-Reyes et al. (2023): **BioHackEU23 report: Enabling FAIR Digital Objects with RO-Crate, Signposting and Bioschemas**. *BioHackrXiv* (in preparation) <https://github.com/ResearchObject/bh2023-enabling-fair-digital-objects/>

<sup>24</sup> <https://signposting.org/adopters/>

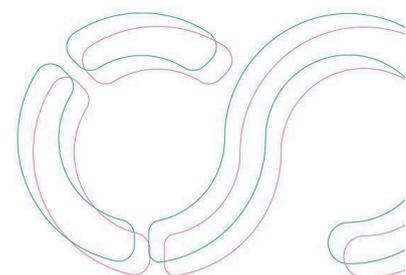
<sup>25</sup> Maaïke Verburg et al. (2023): **FAIR-IMPACT project response to "Community-driven governance of FAIRness assessment: an open issue, an open discussion"**. *FAIR-IMPACT*. <https://doi.org/10.5281/zenodo.7848127>



for how FAIR assessment governance could be established and maintained. Of note, we observed that creating a benchmark environment was a critical aspect of designing and harmonizing test software, and we anticipate that this will become a primary activity of any governance process. We have planned meetings around a joint TF/FAIR-IMPACT workshop where FAIR assessment governance will be discussed, and the various stakeholders will present their preferred models. The event is expected to take place in late November. In addition, we will continue discussing the role of standards registries such as FAIRsharing, where claims of use of community metadata standards can be validated via API calls against the records describing these standards instead of being hard-coded into the FAIR assessment tools; for example, FAIRsharing already follows the advice of the EOSC TF on Persistent Identifiers and includes registrations of the properties and schema for these entities.

## Conclusions & Recommendations

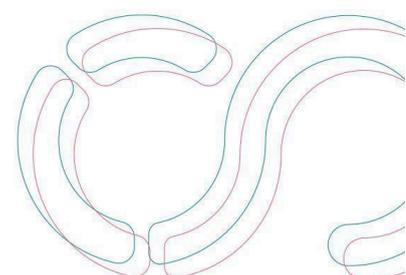
FAIR Signposting is a standards-compliant, straightforward, and unambiguous design pattern for (meta)data publication that fits well with existing Web publication behaviours and creates no new technology standard or infrastructure. Its predictability and consistency will dramatically improve the ability of agents to find and reuse data on the EOSC, providing much-needed harmonisation for EOSC tooling. The TF has emphasized professionalism in creating this specification and its associated compliance tooling to engender confidence in both the users of assessment tools and those being assessed. Implementation of Signposting by repositories and other EOSC resources and its use by FAIR assessment tools will help solve the problem of inconsistent FAIR tests; moreover, the suite of



benchmarks and a governance mechanism that similarly requires benchmarks for all future tests will help ensure new assessment tools do not recreate the current problem. Finally, we can expect additional benefits as repositories increase their support for user-provided metadata - the critical, rich, domain-specific metadata largely invisible to data discovery agents. FAIR Signposting provides a way to immediately maximize the benefit of this evolution in scholarly publishing, giving additional incentive to hasten its implementation.

Our recommendations to the EOSC as a result of these series of workshops and hackathons are as follows:

- 1) As a result of the professionalism of the documentation and the existence of an extensive reference environment for ensuring compliance of associated tools, the EOSC can be confident in their recommendation of Signposting as a FAIR-enabling design pattern for EOSC resources
- 2) Signposting is not intended to be used to the exclusion of other metadata publishing options - for example, embedded Schema.org metadata is still encouraged for improving search engine discovery
- 3) While Signposting offers some freedom in, for example, the minting of novel link-relation types, the EOSC should discourage this, since it interferes with the ability of agents to interpret the discovered metadata.
- 4) EOSC should encourage the registration of community standards in standards registries, such as FAIRsharing, as a means to both advertise the existence of (and commitment to) such a standard and as a mechanism for FAIR assessment tools to validate them.



## Acknowledgments

Participants of the six meetings who are not members of the FAIR Metrics and Data Quality Task Force but without whom none of this would have been achieved: Herbert Van de Sompel, Philippe Rocca-Serra, Avi Ma'ayan, Carole Goble, Ceilyn Boyd, Kristian Garza, Peter Doorn, Alban Gaignard, Thomas Rosnet, Stian Soiland-Reyes, Antonis Lempesis, Andreas Czerniak, Luiz Bonino, Michel Dumontier, Vincent Emonet, Barbara Magagna, Marie-Dominique Devignes, Wilko Steinhoff.

