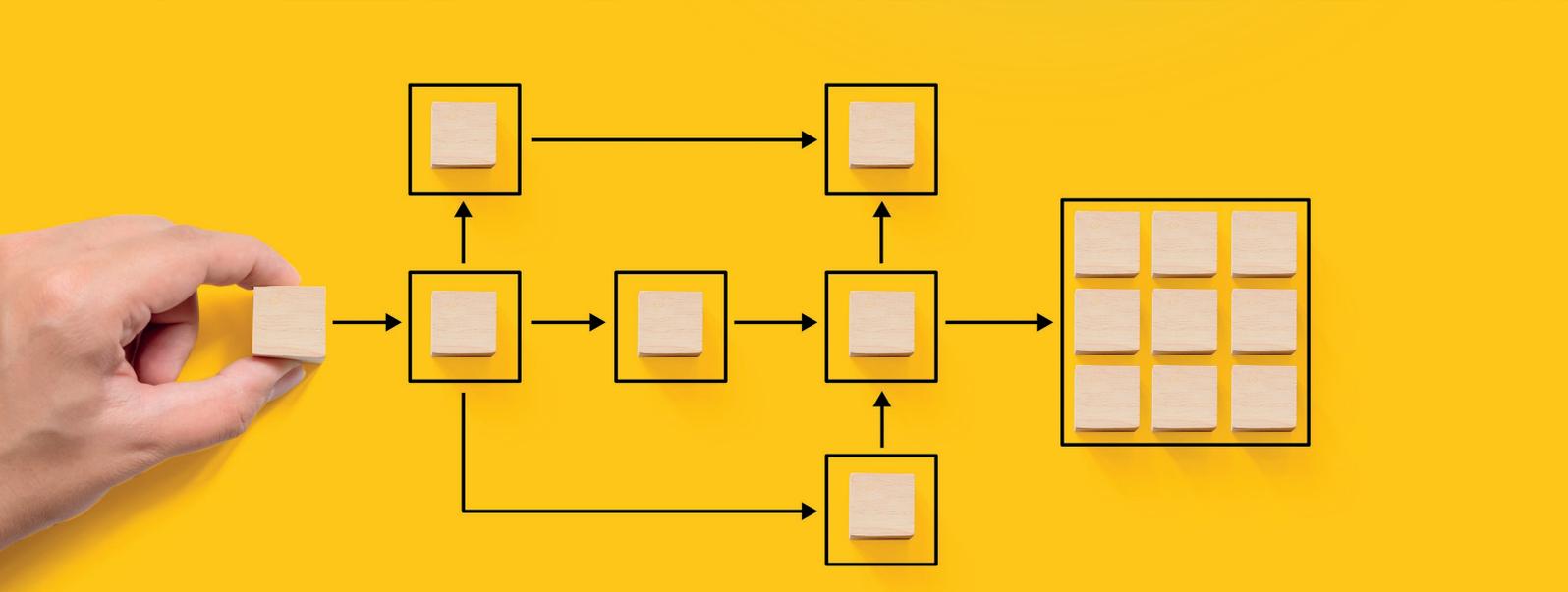# FAIRness assessment of vocabularies: Insights from the British Oceanographic Data Centre team

Authors names, affiliations and ORCIDs:

☐ Alexandra Kokkinaki, British Oceanographic Data Centre

☐ Gwenaëlle Moncoiffé, British Oceanographic Data Centre,
https://orcid.org/0000-0001-6559-4178

☐ Clara Linés, Digital Curation Centre (DCC),
https://orcid.org/0000-0003-3437-5145

Support action:

FAIRness Assessment Challenge. During the 3 month challenge participants took part in three virtual workshops to self-assess and incrementally improve the FAIRness of their selected outputs. During the support action, participants benefited from interacting with a group of mentors representing the various FAIRness assessment tools and methods.

Summary:

A team from the British Oceanographic Data Centre tested different methods and tools (10 simple rules, recommendations for FAIR semantics, O'FAIRe, FOOPS!, and F-UJI) in the task of assessing the FAIRness of their controlled vocabulary.

Introduction

The British Oceanographic Data Centre (BODC) is part of the UK's National Oceanography Centre. BODC operates the Natural Environment Research Council's (NERC) Vocabulary Server (NVS)[1], which provides access to semantic artefacts to the marine and wider community.

BODC has been providing expertise in the management of semantic artefacts for over 50 years, adopting linked data technology in the early 00s. We are actively involved in international initiatives in this domain. Both authors are co-chairs of the Research Data Alliance (RDA) Interest and Working Groups: the InteroperAble Descriptions of Observable Property Terminologies (I-ADOPT) Working Group[2] and the Vocabulary Services Interest Group[3].

We also had prior knowledge of the FAIRness assessment, having worked with the recommendations for FAIR semantics[4] in an RDA task group. We knew about the 10 simple rules[5] and O'FAIRe[6], although O'FAIRe was only developed recently and only available for users of the OntoPortal[7] technology so we had not had an opportunity to test it before. F-UJI[8] is well known but it is not specific to semantics, and therefore can only work superficially for vocabularies.

Approach taken:

We tested all those tools, including FOOPS![9], in assessing the FAIRness of our vocabulary. This exercise provided valuable insights into both our service and the tools themselves, each offering unique advantages based on specific needs:

☐ **10 Simple Rules:** These are a good starting point, especially for inexperienced users looking to publish vocabularies in a FAIR manner or to simply understand the minimum requirements needed. It is the most approachable and least intimidating method, also beneficial for those transitioning from a PDF-based to a digital vocabulary.

☐ **Recommendations for Making a Vocabulary FAIR**: These are more technical and detailed, suitable for those who have already implemented the 10 Simple Rules and want to progress towards more advanced technical interoperability.

☐ **O'FAIRe**: This tool is very detailed with helpful suggestions. It enhances the FAIRness of vocabularies published in OntoPortal-based repositories by improving associated metadata. It could be particularly useful for individuals wanting to publish a single vocabulary within the OntoPortal Alliance. While it doesn't directly increase the FAIRness of vocabularies published outside OntoPortal, it offers valuable general insights, such as properties checked for versioning and documentation. Importing a vocabulary to OntoPortal can be a beneficial FAIRness exercise even for those publishing elsewhere, as any issues encountered may indicate areas for improvement in interoperability and reuse. A suggested improvement for O'FAIRe is to enable metadata application to multiple vocabularies simultaneously.

1   https://vocab.nerc.ac.uk/
2   https://www.rd-alliance.org/group_output/interoperable-descriptions-of-observable-property-terminologies-i-adopt-wg-outputs-and-recommendations/
3   https://www.rd-alliance.org/rationale/vocabulary-services-interest-group/
4   Yann Le Franc, Luiz Bonino, Hanna Koivula, Jessica Parland-von Essen, & Robert Pergl. (2022). D2.8 FAIR Semantics Recommendations Third Iteration (V1.0). Zenodo. https://doi.org/10.5281/zenodo.6675295
5   https://doi.org/10.1371/journal.pcbi.1009041
6   https://github.com/agroportal/fairness/blob/master/README.md
7   https://ontoportal.org/
8   https://www.f-uji.net/
9   https://foops.linkeddata.es/FAIR_validator.html

☐ **FOOPS!**: This tool is designed for ontologies and does not work for other semantic artefacts like vocabularies. This highlights the need for a tool that can automatically assess and improve vocabularies, thesauri, and taxonomies independently of OntoPortal. We discussed this with the FOOPS! developers, who indicated they could extend FOOPS! to include vocabulary assessment. This would be a positive development; otherwise, a new tool would be necessary to fill this gap.

Challenges encountered and addressed:

Persistent URLs and Handles: One issue we encountered was how persistent URLs and handles are evaluated in the FAIRness assessment. For instance, the FOOPS! tool suggests that if you do not have a handle, then you do not have a PID. We believe this should be reassessed, possibly by considering additional metadata and the sustainability of the service. Our service provides very stable URLs, and it would be beneficial if there were a way to evaluate that our identifiers meet the FAIR principles for globally unique and persistent identifiers. Relying on an external PID service could potentially slow down our services or make them less available.

Levels of Exposure of Semantic Artefact: Another area which we believe would need attention is the level at which tools expect to find information for the FAIRness assessment. Currently, FOOPS! requires all information about the collection and the repository to be available at the concept level. However this approach is impractical and leads to a large amount of superfluous information. Since we use linked data, which the FAIR principles are based on, it is crucial that FOOPS!, or a new tool designed to assess the FAIRness of vocabularies, can evaluate semantic artefacts at different levels. This approach aligns with linked data principles, allowing movement between various levels through links.

Impact:

The support action was useful because it compelled us to conduct this analysis and evaluation, leading to valuable insights.

Even if our FAIRness score was not improved as a result, testing these tools has been useful and provided information that can help shape our strategy towards achieving greater FAIRness.

Actions identified include:

1. improving the publication of concept schemes in the NVS

2. fix the versioning of the collections: One of the FAIR principles is for vocabularies to have versions, we do have them at the level of concept but after some changes we did, some issues have arisen that need to be fixed.

3. Another action we have identified is to ensure that we are able to be imported automatically by tools like the OntoPortal Alliance or the OLS (Ontology Lookup Service). We keep publishing vocabularies through the EarthPortal, which is part of the Ontoportal and developed within FAIR-IMPACT. We started wondering if we need to publish NVS as a set of ontologies if many users request it. For us, it is all focused on our users, and if we are going to put the effort into improving any aspect of NVS, we need to make sure that it benefits our users. In this case, it benefits them through their ability to use different portals to access our vocabulary. We have users from a wide range of organisations and roles, and having different versions helps enable more people with different technical skills or requirements to use the vocabulary. For example, people that are more ontology focused often require a vocabulary based on classes, while ours is based on instances, and they are incompatible.

We are also participating in the FAIR-Impact Synchronisation Force sessions and would take that opportunity to continue providing feedback on the tools we tested and on FAIRness assessment of semantic artefacts in general.

Key message:

Our solutions may sometimes not be fully aligned with the rules and the aspects that lead to a high score but if they help us achieve interoperability and reuse in practice, that is what matters for FAIRness.

There is a gap in the automatic assessment of the FAIRness of vocabularies - either FOOPS! can be extended or a new tool is needed.

@fairimpact_eu

in

company/fair-impact-eu-project/

fair-impact.eu