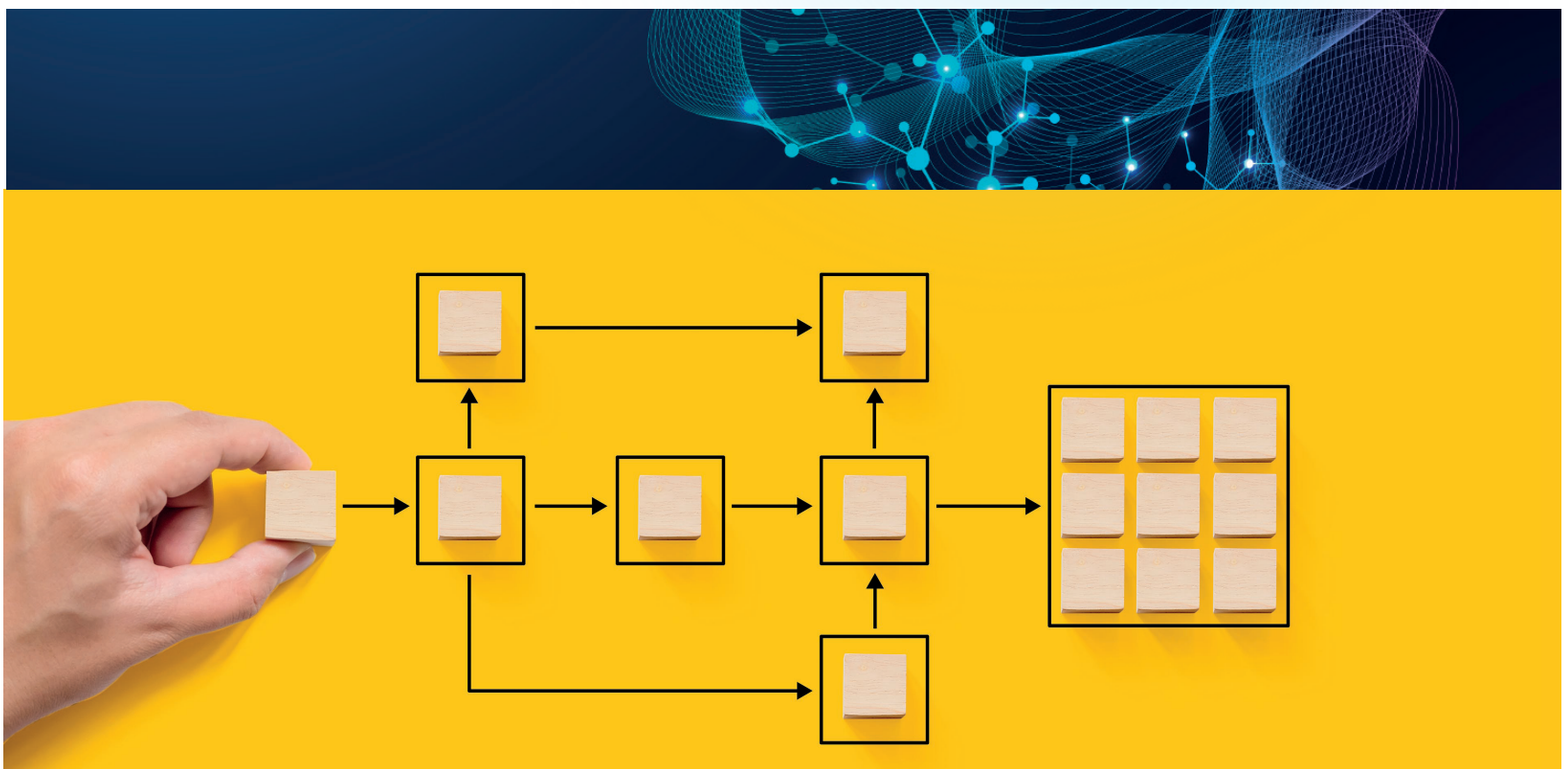
eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

FAIRSFAIR
Fostering Fair Data Practices in Europe

FAIRness assessment in the Life Cycle Assessment domain



IMPLEMENTATION STORY

Authors names, affiliations and ORCIDs:

- Agneta Ghose, Department of Planning, Danish Centre for Environmental Assessment (DCEA), Aalborg University, https://orcid.org/0000-0003-1972-1433
- 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.

Keywords:

F-UJI, FOOPS!, FAIRsFAIR Semantic Recommendations

Summary:

A researcher from the Life Cycle Assessment (LCA) domain assessed the FAIRness of one of their datasets using F-UJI and an ontology using FOOPS! and FAIRsFAIR Semantic Recommendations, as part of broader ongoing efforts to understand the level of FAIRness in the LCA domain and gain insights to define good practice workflows for data sharing.

Introduction

Life cycle assessment (LCA), the analysis of the environmental impact of a product over its entire life cycle, is a very multidisciplinary field of study. We get data from industries about product processing, but also environmental data from other sources, for example, data on emissions.

There are many challenges to achieve FAIRness in this domain. The data is coming from many sources and there are no standard formats, therefore, the way in which each piece of data is structured varies. There have been some efforts towards building repositories for research data in the domain, but they are mainly specific to a research group or to a country, and they can't be accessed by external research teams without being given special permissions first. Still, I have noticed that people in this domain, at least those from academia, are willing to share data. We depend on proprietary databases and that is data that we cannot share, but there is willingness to share the data that we develop. The problem is that there are no guidelines, and data is usually shared within the papers or as a pdf in the supplementary materials. Increasingly, maybe because the funding institutions are asking for that, the data or the preprint is also published on a generic repository. But there is still much to improve, such as providing interoperable formats and complete metadata.

My goal for the support action was twofold. First, I wanted to check how FAIR is our current approach to dataset sharing and gain insights to define good practice workflows for data sharing in the LCA domain. Then, I was also hoping to assess how well we have shared an ontology (BONSAI ontology) in which I have been working with some colleagues to give more structure to LCA or environmental assessment data.

Implementation:

I used the F-UJI tool[1] to assess one of our datasets, whilst for the ontology I tested FOOPS![2] and the FAIRsFAIR Semantic Recommendations[3]. F-UJI provided a moderate score for the dataset shared on Zenodo, which improved as I added more metadata information.

Our ontology is shared on GitHub and deployed on a website[4]. I first tried the website link in FOOPS! and the score was very low (4%). That was puzzling since we had followed good practice recommendations and made sure it had a licence linked to it. But then I learned from the mentors that the webpage was not accessible to others. I tried the tool again using the link to the GitHub page and the score was much higher (50%). It was interesting to see that the score can't be taken at face value, that there is more behind it.

After that, I tried FAIRsFAIR Semantic Recommendations and I found that method was better suited for my needs as it gives indications on what you should have. It also provides a value linked to each of the points and indicates how much this value could increase if improvements are implemented. I would also like to try SHARC and the FDMM tools, which I think could be very useful for me too.

1 https://www.f-uji.net/
2 https://foops.linkeddata.es/FAIR_validator.html
3 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
4 https://ontology.bonsai.uno

Challenges encountered and addressed:

The main challenges we encountered during the support action were related to ontology sharing, as described above. In that regard, I would be interested to have examples of use cases or workflows from other domains on how they have built and shared their ontologies.

On a generic level, our approach to share data through Zenodo seems good. It gets a good score in F-UJI, and Zenodo facilitates that the data gets a persistent identifier and versioning, that a metadata standard is followed, that the standard is semantically linked, etc. However, this approach presents some limitations at a domain specific level. The data in Zenodo can be shared in any format, for example pdf or word documents, and with any information; as a result, even if it has metadata, it is not necessarily well defined and interoperable. Unfortunately, these things cannot be tested using the F-UJI tool. Automatic tools can help identify if a machine can find the data or ontology, but at this stage, in our domain we are also missing more basic things like widespread use of persistent identifiers. There is a lack of guidance and awareness in the domain around using persistent identifiers when sharing data, and the databases available don't provide them. I think that a tool like F-UJI can help highlight what is lacking in the field.

Although the tools and the advice we can get from university data managers are very helpful, I miss more domain specific information, which will only arrive when there is more awareness about the possibilities and benefits of FAIRness in the domain. Recognition for good data sharing would also be an incentive, as otherwise it can feel to people like just more bureaucracy.

Impact:

This experience has been really valuable, because it helped validate my efforts towards FAIR sharing in the field of LCA. I was not aware of these assessment tools, but they certainly helped to benchmark the current data sharing practice and consider opportunities for improvements. The tools have given me an indication of where I am standing, and although the score cannot be taken at face value, the recommendations highlight points that need further consideration. One thing I learned for example is all the metadata information that should be available on the landing page.

During the support action, I was also writing a paper on sharing practices in my domain[5], linked to FAIR sharing practices, which got very positive reviews. I looked at current sharing practice in our domain, and I saw that, most of the time, the data is shared without persistent identifiers, and not only ontologies are not linked, but even the domain standard nomenclature is not necessarily used, or the data is not shared in the standard formats. A lot of things need to be improved. We also need more awareness about other aspects of FAIR data sharing, such as the different types of licenses and when to use them and data management plans, which have a lot of potential to guide people.

I would like to keep building on this knowledge, work on building workflows for data sharing, and increase awareness on how to best share data. I'm in conversation with colleagues that are also interested in this, and I have been presenting in domain specific conferences. I was also contacted by DeiC, the Danish e-Infrastructure Consortium, to give a presentation on my experience with the tool and we would probably have a follow up meeting to see what more we can do.

Key messages:

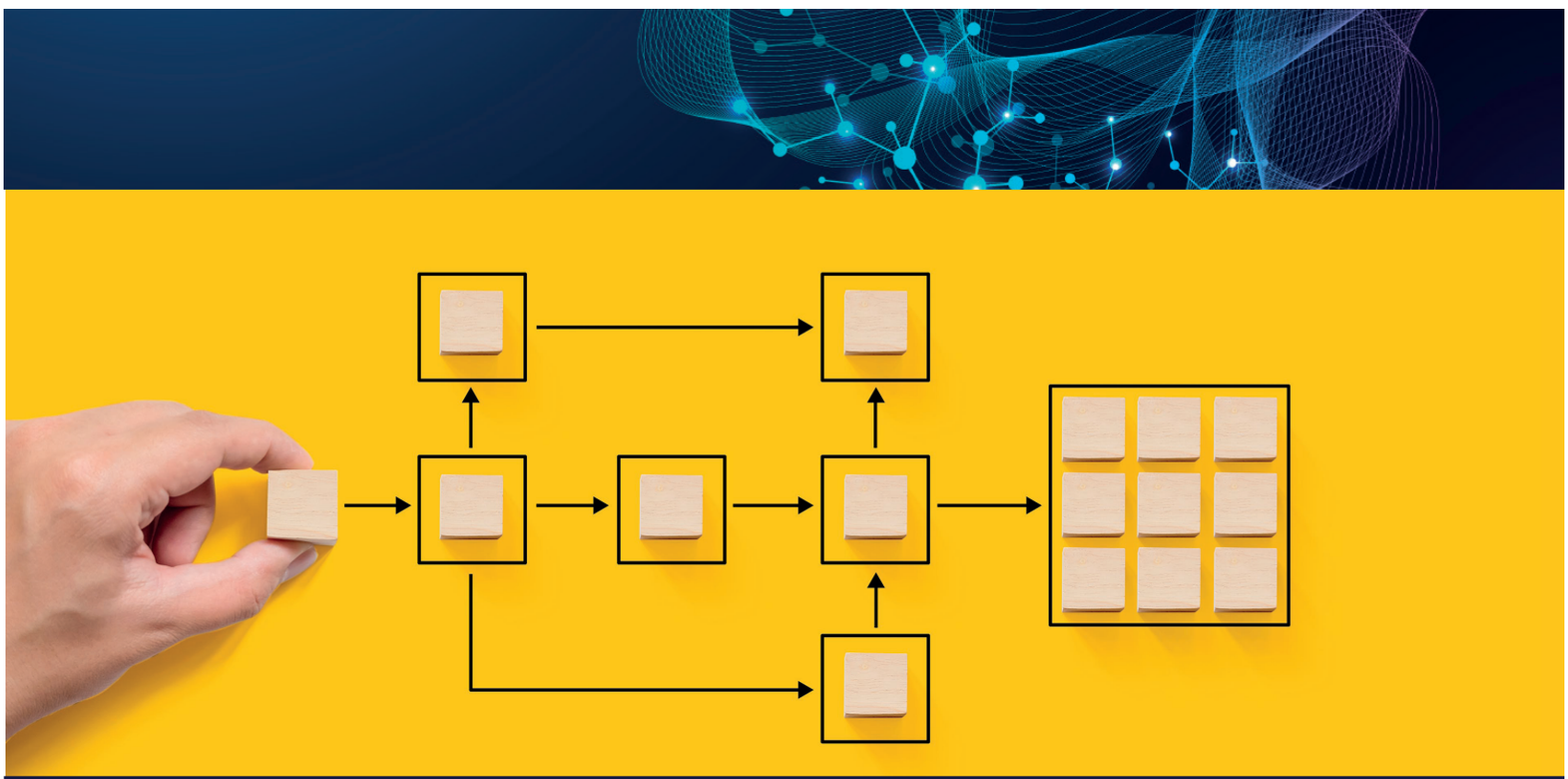
The metrics from most of the tools are indicative of issues but I can recommend not to take these metrics on their face value and see what is driving the score. I would also recommend using multiple tools.

5 Ghose, A. Can LCA be FAIR? Assessing the status quo and opportunities for FAIR data sharing. Int J Life Cycle Assess (2024). https://doi.org/10.1007/s11367-024-02280-3

@fairimpact_eu

company/fair-impact-eu-project/

fair-impact.eu



Funded by the European Union