FAIR and Open Science (FAIR/O) for Low Carbon Energy Research: Gap Analysis

(5) (Open-Source) Modellierung

Demet SUNA(1), Christopher BURGER-SCHEIDLIN(1), August Hubert WIERLING (2), Valeria Jana SCHWANITZ(2), Manfred PAIER(1)

(1) AIT Austrian Institute of Technology GmbH,

(2) Western Norway University of Applied Science

Motivation and research questions

Transparent and integrated energy data management with useful metadata and quality assurance provides the basis for society to choose, monitor and implement sustainable transition pathways. Open science research is of key relevance here because it enables the detection of conflicts across findings in different disciplines and facilitates the use of synergies in the first place. Therefore, databases need to adhere to principles of FAIR and open data (O), where FAIR stands for findable, accessible, interoperable, and reusable. This work aims to introduce identified community specific bottlenecks and gaps for the implementation of FAIR/O data, hindering the exploration of the research potential in the low-carbon energy field, conducted within the Horizon 2020 project EERAdata**[[1]](#footnote-1)**.

Methodical approach

For this analysis, three main approaches were followed: i) create a community-based discussion platform and reflect the community perspectives on FAIR/O and metadata concepts for the low carbon energy domain within two community workshops,ii)conduct human based assessment of the FAIR/O state of the art for a representative sample of energy databases and iii) conduct metadata assessment for nine selected databases by using fifteen “simple” Dublin Core metadata elements [1] which comprises fifteen metadata elements from descriptive (e.g. Title, Subject), administrative (e.g. Date, Rights) and structural (e.g. Format) categories.

For the FAIR principles, this work follows the criteria of Wilkinson et al [2]. For the human assessment we selected the online tool ARDC [3], as this mainly follows these criteria closely and provides a semi-automated scoring algorithm.

To determine the current state of the FAIR/O principles, a representative sample of databases from the energy domain was compiled. About thirty relevant databases were pre-identified by experienced researchers from different energy domains within the EERAdata consortium and extended by workshop participants. The scope was augmented by adding databases recommended for their FAIR implementations such as “Sharewind” [4], from the wind energy domain and general data repository initiatives closely complying with FAIR and metadata ideas, such as “Zenodo” [5], as best practise examples.

Results and conclusions

The most common gap relating to Findability was the lack of persistent identifiers for data and metadata. This may lead to broken links to data and metadata in the future. In terms of Reusability, the main gaps identified were partially recorded provenance and the lack of license information. While missing links to other metadata hinders the Interoperability, the uncertainty whether metadata would still be available when the data would no longer be make difficult the Accessibility.

When it comes to metadata gaps, a prominent gap for the assessment is the non-availability of the metadata information in a compact form, such as a summary displayed on the website together with the data, or in the form of xml files or equivalent formats. Metadata elements, that appeared to be used scarcely were source, publishers/contributors, language, format, and coverage. It also seemed that the distinction between creators, publisher, and contributors, was not always made rigorously.

Most databases selected for FAIR/O assessment were publicly accessible. Therefore, the question was not whether the data was open or not, but rather what the level of openness was. The assessment showed that most of the databases require human intervention. This implies a lack of machine-actionability even though data are publicly accessible.

Literature

[1] DCMI, 2012. DCMI: Dublin CoreTM Metadata Element Set, Version 1.1: Reference Description [WWW Document]. URL <https://www.dublincore.org/specifications/dublin-core/dces/>

[2] Wilkinson, M.D. et al, 2016: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

[3] ARDC, FAIR self-assessment tool – Australian Research Data Commons (ARDC) [WWW Document]. URL <https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/>

[4] Sharewind Dataset in Zenodo: <https://doi.org/10.5281/zenodo.1162738>

[5] Zenodo: <https://zenodo.org/>

1. This project has received funding from the **European Union’s Horizon 2020** research and innovation programme under **grant agreement no. 883823**. [↑](#footnote-ref-1)