

EGU21-12560

<https://doi.org/10.5194/egusphere-egu21-12560>

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Applying FAIRness evaluation approaches to (meta)data preserved at the World Data Center for Climate (WDCC): results, lessons learned, recommendations

Karsten Peters-von Gehlen^{1,2}, Andrej Fast¹, Daniel Heydebreck¹, Heinke Höck¹, Amandine Kaiser¹, Andrea Lammert¹, and Hannes Thiemann¹

¹Deutsches Klimarechenzentrum GmbH (DKRZ), Datamanagement, Hamburg, Germany

²contact: peters@dkrz.de

The perceived community demand for research data repositories to provide services ensuring that stored data comply with the FAIR principles requires transparent evaluation of such services. In previous work, the long term archiving service WDCC¹ (World Data Centre for Climate) at DKRZ (German Climate Computing Center, Hamburg) underwent an even-handed self-assessment along the published FAIR principles and the results are published on the DKRZ homepage².

Here, we present results of an overhaul of the previous WDCC FAIRness-assessment by subjecting datasets archived in WDCC to a number of now available objective FAIR assessment approaches which are available as questionnaires or fully-automated web applications^{3,4,5}. In these approaches, FAIRness is assessed using so-called metrics or maturity indicators. While the terminology is more a choice of the test provider - e.g. the term 'metric' may be off-putting for some - both give quantitative results. First tests show that (meta)data archived in WDCC seem to attain a higher level of FAIRness when evaluated using questionnaires compared to the results obtained from fully-automated applications. Further work is needed to substantiate this finding.

We learn that while neither one of the two evaluation approaches is ideal, they both show merit. Questionnaires – answered by knowledgeable repository staff – capture domain- and repository-specific aspects of FAIRness, like the use of controlled vocabularies in the datasets, granularity of archived datasets, reuse documentation or clear assessment of local data access protocols. However, the human-performed evaluation does not capture machine-actionability in terms of FAIR. This aspect is – naturally – very well assessed by automatic evaluation approaches, but the results strongly depend on the way the tests for FAIR metrics/maturity indicators are implemented. However, automatic tests often only assess metadata FAIRness, lack domain-specific

FAIRness indicators or yield failed tests if a repositories' technical properties, e.g. the specification of authentication procedures for data access, are not compatible with what an automatic procedure is built to test for.

Therefore, since WDCC has an over 30 year long history of preserving climate-science related data

with a focus on reusability by the community (and beyond), FAIRness evaluations based on human-actionable questionnaires show a high degree of FAIRness. We further learn that there is an urgent need for specifically-designed automatic FAIR testing approaches taking into account domain-specific data standards and structures. Especially the availability of atmospheric and climate science related FAIR metrics/maturity indicators is very limited. We thus recommend compilations of the latter and we will aim at contributing to this effort.

In our contribution, we specifically showcase strong as well as weak aspects of the WDCC service in terms of FAIRness and report on our measures to increase domain-specific FAIRness of WDCC and present recommendations for establishing FAIR indicators for (meta)data common to the Earth System Science community. We will make the results of our assessment openly available on the WDCC homepage as well as produce a corresponding Open Access peer-reviewed publication.

References:

¹<https://cera-www.dkrz.de>

²<https://cera-www.dkrz.de/WDCC/ui/cersearch/info?site=fairness>

³<https://www.rd-alliance.org/node/60731/outputs>

⁴<https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#/>

⁵<https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>