

Toward an implementable framework of FAIR principles for Earth science data management and stewardship

Abdullah Alowairdhi (alow4530@vandals.uidaho.edu), Xiaogang Ma (max@uidaho.edu)
Department of Computer Science, University of Idaho, Moscow, ID 83844

INTRODUCTION

- FAIR is an acronym of Findable, Accessible, Interoperable, and Reusable. The FAIR principles were first launched in 2014.
- Works on FAIR also emphasize that each of the four principles should be equally applicable to both humans and machines.
- The Earth science community is beginning to apply FAIR principles in research activities for better data stewardship (Stall et al., 2018).

FAIR Guiding Principles

- FAIR Principles (Wilkinson et al, 2016)** suggest that every data element should be:
- Findable:** Should have a globally-unique identifier, and that this identifier should be associated with contextual, searchable metadata.
- Accessible:** The globally-unique identifiers should all resolve to data or metadata using an open, standard protocol.
- Interoperable:** The data and metadata should use a formal, broadly applicable representation language, and utilize open and widely-accepted domain-relevant vocabularies and ontologies.
- Reusable:** The data should be richly described with an abundance of cross-references, and with a clearly-defined mechanism for accessing provenance and license information.

DISCUSSION

- FAIRtool.org** will be a useful framework to evaluate and improve the FAIRness of data outputs in Earth science.
- The tool can be used by both individual researchers and data managers.
- ESIP community will find it an adequate service for better management and stewardship of data.
- Benefits of FAIR data: 1) Increased citations, 2) Career recognition, 3) New collaborations, 4) Easier to find useful data, 5) New research opportunities.

Acknowledgements

We thank the support from the ESIP Lab Incubator Program and the insightful comments from ESIP community members on our research ideas.

REFERENCES

- Stall, S., Yarmey, L., Boehm, R., Cousijn, H., Cruse, P., Cutcher-Gershenfeld, J., Dasler, R., de Waard, A., Duerr, R., Elger, K., Fenner, M., Glaves, H., Hanson, B., Hausman, J., Heber, J., Hills, D., Hoebelheinrich, N., Hou, S., Kinkade, D., Koskela, R., Martin, R., Lehnert, K., Murphy, F., Nosek, B., Parsons, M., Petters, J., Plante, R., Robinson, E., Samors, R., Servilla, M., Ulrich, R., Witt, M., Wyborn, L., 2018. Advancing FAIR Data in Earth, Space, and Environmental Science. *Eos*, 99. DOI:10.1029/2018eo109301.
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. DOI:10.1038/sdata.2016.18.

OBJECTIVES of This Research

- Contribute to the technical framework of FAIR implementation to help researchers achieve better data stewardship.
- Automatically produce human and machine-readable FAIR-compliant metadata for digital data outputs of the Earth science community.
- Develop a metric system to evaluate the FAIRness level of the digital data output and to help improve the FAIRness.
- Apply semantic web technologies in the technical development of a portal FAIRtool.org to improve the open data environment for Earth science.

METHODS

- Building the FAIR-O ontology by utilizing a collection of ontology and vocabulary standards: DCAT, Dublin Core, PROV-O, FOAF, Schema.org.
- Using the FAIR-O ontology to build the FAIR semantic web application.
- The FAIRtool.org, a semantic web application, will use FAIR-O and other APIs to build FAIR-compliant metadata for various datasets.
- FAIRtool.org includes an evaluation system to calculate points assigned to each FAIR principle and generate a final FAIRness grade.
- The process in FAIRtool.org starts when a user login with his ORCID and upload a dataset. The dataset will be assigned a DOI and go to the other steps in the process to complete the 15 FAIR principles.

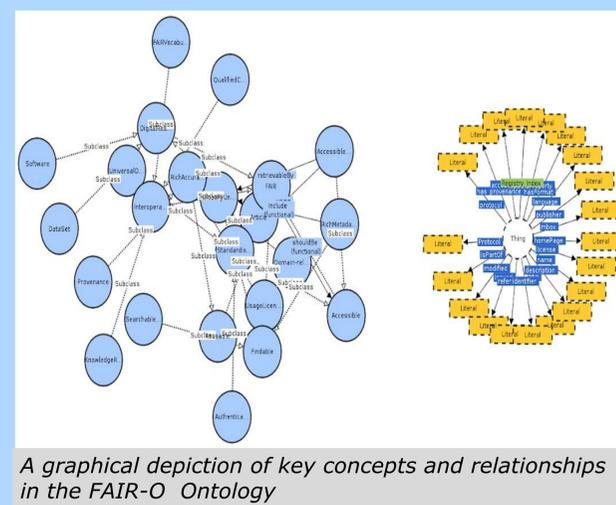
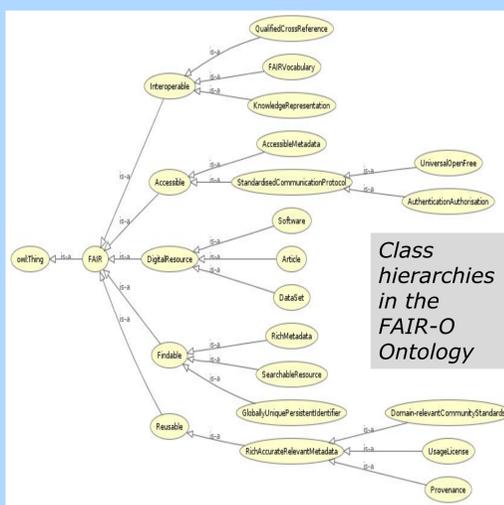


OPEN DATA



TECHNOLOGIES

- Protege (<https://protege.stanford.edu>) for ontology design and encoding.
- OWLviz and VOWL for visualized diagrams of FAIR-O.
- Vitro (<http://vitro.mannlib.cornell.edu>) for semantic web application.
- Apache-Tomcat web server, other APIs, and JavaScript programming.



ANTICIPATED RESULTS

- Leverage several existing standard ontologies and vocabularies to develop the FAIR-O ontology
- FAIRtool.org** to automate the production of metadata for dataset.
- Production of a RDF format for the metadata.
- The evaluation process in FAIRTool.org can be re-executed over time and can allow a user to monitor the improvement of the FAIRness of data.
- The FAIRness metric system and the technical process in FAIRtool.org may be extended to cover other objects, such as code and sample.