8-2024

Materials Data Science Ontology (MDS-Onto): Unifying Domain Knowledge in Materials and Applied Data Science

Van D. Tran
Case Western Reserve University, vxt101@case.edu

Jonathan E. Gordon
Case Western Reserve University, jeg165@case.edu

Alexander Harding Bradley
Case Western Reserve University, ach159@case.edu

Balashanmuga Priyan Rajamohan
Case Western Reserve University, bxr261@case.edu

Quynh D. Tran
Case Western Reserve University, qdt@case.edu

See next page for additional authors

Authors

Van D. Tran, Jonathan E. Gordon, Alexander Harding Bradley, Balashanmuga Priyan Rajamohan, Quynh D. Tran, Gabriel Ponón, Yinghui Wu, Laura S. Bruckman, Erika I. Barcelos, and Roger H. French

Materials Data Science for Stockpile Stewardship

COE: US-Department of Energy-NNSA Award

Materials Data Science Ontology (MDS-Onto): Unifying Domain Knowledge in Materials and Applied Data Science

Van Tran[1]*, Jonathan E. Gordon[1]*, Alexander Harding-Bradley[1], Priyan Rajamohan[1], Quynh Tran[1], Gabriel Ponon[1], Yinghui Wu[1], Laura S. Bruckman[1], Erika I. Barcelos[1], Roger H. French[1]

[1] Materials Data Science for Stockpile Stewardship Center of Excellence, Cleveland, OH 44106, USA, * Equal Contribution

Purpose

Ontologies are used in the scientific community as a means to standardize concepts and terminology to facilitate data reuse

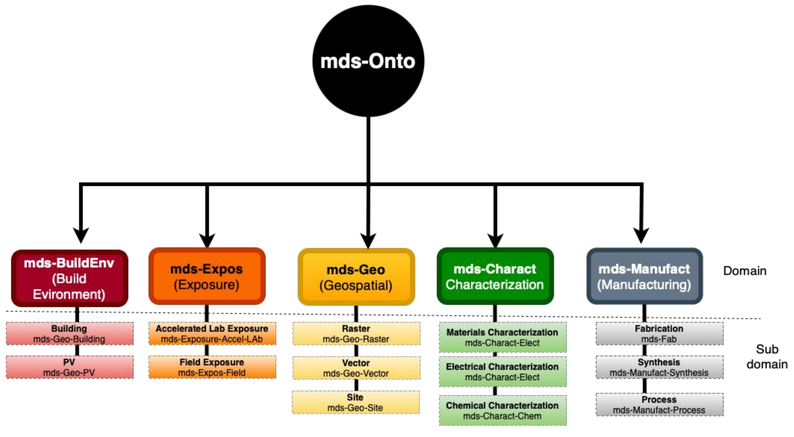
MDS-Onto Framework: Providing a comprehensive framework for ontology development, publishing, integration with the semantic web, and storage formats

R and Python packages: A user-friendly ontology creation and visualization tool that also allows for data and analysis to be annotated with robust metadata

MDS-Onto Framework

MDS-Onto Framework includes recommendations for ontology creation and development, how to FAIRify data, and platform for publishing created ontologies.

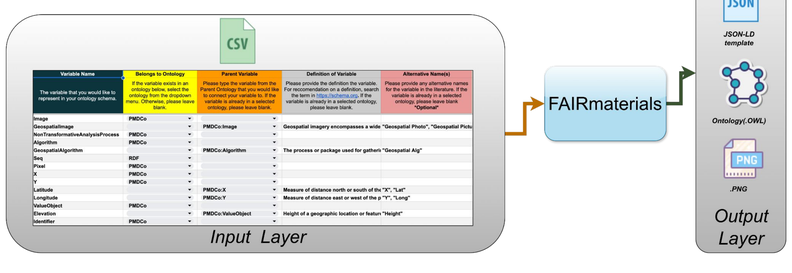
Application of MDS-Onto Framework: The creation of the MDS-Ontology, which encompasses 5 domains and over 30 subdomains in materials science and engineering.



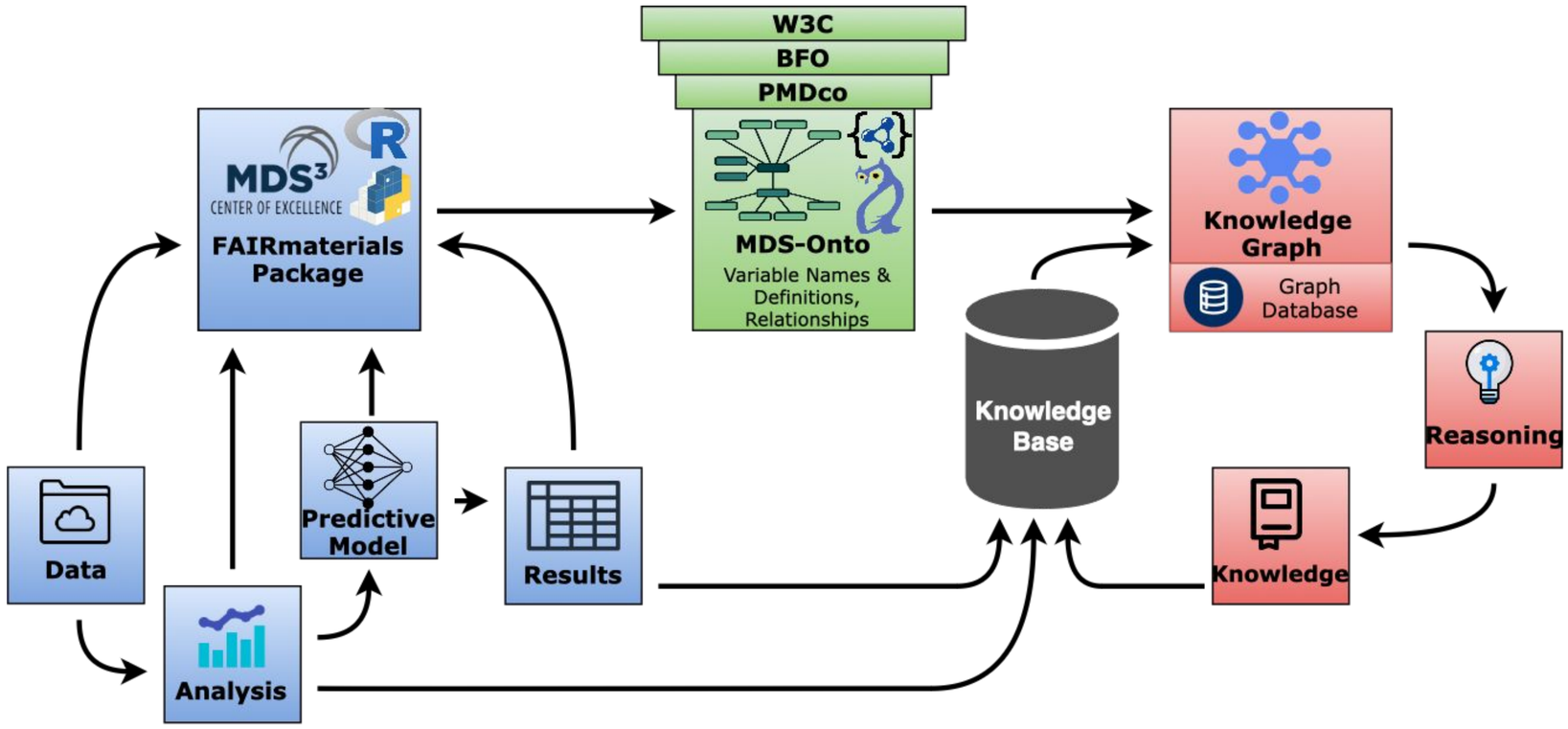
FAIRmaterials R and Python Package

`FAIRmaterials`[3] is a bilingual package simplifying the creation and visualization of ontologies.

The package converts simple and structured CSV inputs into rich, well-defined ontologies to support the process of data FAIRification[1].

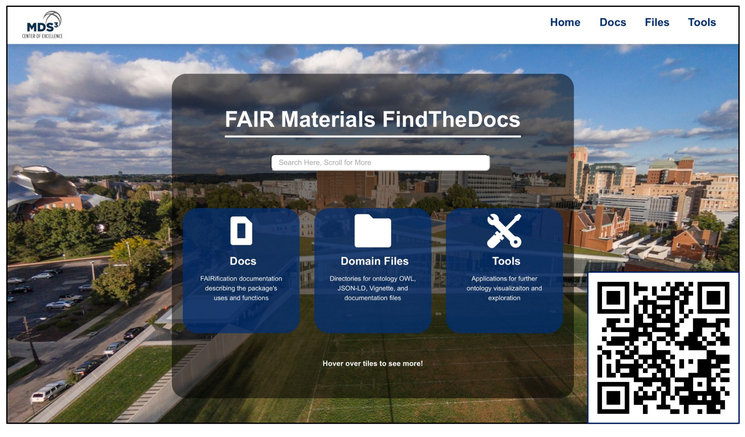


MDS-Onto Framework Workflow



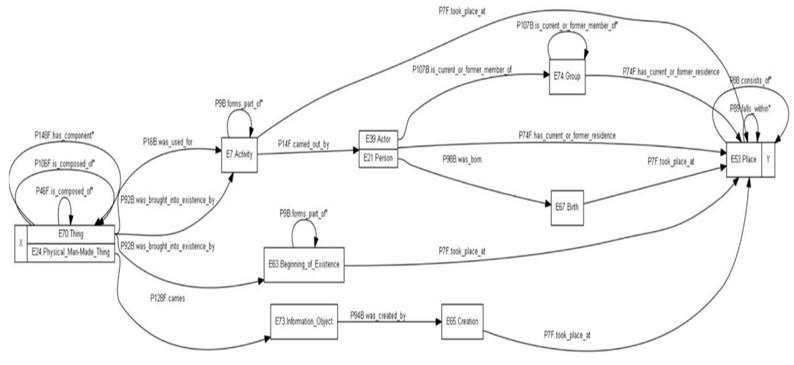
Find The Docs Website

Publishing ontologies is crucial for enabling new suggestions and integration into other researchers data. SDLE FAIR Materials FindTheDocs website, facilitates interoperability by offering an accessible interface for searching ontologies created via the MDS-Onto Framework.

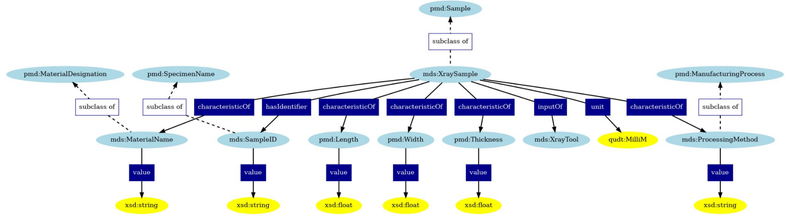


Knowledge Graphs

Knowledge graphs are networks where nodes represent concepts and entities and edges represent relationships between these concepts and entities. A neural-symbolic AI will act as a reasoner for query requests, and the knowledge base will be stored in graph database with ontologies as schema for organization.



MDS-Onto Framework Applied



References

[1] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3(1), 160018. doi:10.1038/sdata.2016.18.
[2] Zhang, J. et. al. Neural, symbolic and neural-symbolic reasoning on knowledge graphs. *AI Open*, Vol 2, 2021. doi:10.1016/j.aiopen.2021.03.001
[3] Gordon, J. E., et al. FAIRmaterials: Ontology Tools with Data FAIRification in Development. https://CRAN.R-project.org/package=FAIRmaterials, 2024. doi:10.32614/CRAN.package.FAIRmaterials.

Acknowledgement

This work was supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) Agreement Numbers DE-EE0009353 and DE-EE0009347, Department of Energy (National Nuclear Security Administration) under Award Number DE-NA0004104 and Contractnumber B647887, and U.S. National Science Foundation Award under Award Number 2133576.