Building domain specific vocabularies for packaging and archiving research and cultural data

Authors

- Mike Lynch / University of Sydney [ORCID](#)
- Peter Sefton / University of Queensland. [ORCID](#)

Abstract

This presentation will cover a complete ecosystem showing tooling for creating vocabularies in the Schema.org Schema style, publishing them online so that URLs resolve to useful definitions for users (which is often NOT the case with linked data vocabularies, including quite widely used ones), generating user interfaces for editing metadata and data records that use the vocabularies via automatically generated Editor Profiles and show the resulting data in use in repository portals.

This work is built on the Research Object Crate (RO-Crate)(1) specification for linked-data object description which was designed to be as programmer and end-user friendly as possible.

Examples will be drawn from the Language Data Commons of Australia where the vocabulary ecosystem was largely developed and refined, as well as other projects that use this general approach including two ARDC co-investments in data integration, the OHRM and Heurist projects.

Acknowledgements

Reference

1. Soiland-Reyes S, Sefton P, Crosas M, Castro LJ, Coppens F, Fernández JM, et al. Packaging research artefacts with RO-Crate. Data Sci. 2022 Jan 1;5(2):97–138.