Building an Application Ontology and Knowledge Graph for Rare Disease Patient-Level Data

Emily Hartley [1], Daniel Olson [1], Ian Braun [1], Nicolas Matentzoglu [2], James A. Overton [3] and Ramona L. Walls [1]

[1] Critical Path Institute, 1730 E. River Rd. #200, Tucson, AZ 85718, USA
[2] Semanticly Ltd., 71-75 Shelton Street, Covent Garden, London, United Kingdom WC2H 9JQ
[3] Knocean Inc., 2-107 Quebec Ave., Toronto, ONT M6P2T3 Canada

Keywords
Rare disease, application ontology, knowledge graph, patient data

1. Introduction

This poster describes the process of developing an application ontology, the Critical Path Ontology (CPONT), and building a knowledge graph from Observational Medical Outcomes Partnership (OMOP) structured data. The integrated knowledge graph is part of the Rare Disease Cures Accelerator-Data and Analytics Platform (RDCA-DAP®) [1] and will support clinical development in rare diseases.

2. Methods and Results

We used the Ontology Development Kit [2] to set up the CPONT development environment. We then used ROBOT [3] to import modules from multiple OBO ontologies. We imported publicly available curated mappings as well as a list of our own mappings of OMOP vocabulary terms from the patient-level data to terms in the OBO ontologies. To create a knowledge graph, we used a python script to load the patient-level records into a ROBOT template file and used ROBOT to convert the template into OWL instance-level triples. This OWL file is merged with cpont.owl, and lastly, ROBOT query is used to run SPARQL queries on the combined OWL file. CPONT is available at https://gitlab.c-path.org/c-pathontology/critical-path-ontology. The knowledge graph will be available to approved users through the RDCA-DAP portal.

3. Conclusion

To our knowledge, this will be the first widely available knowledge graph to encompass patient-level clinical trial and electronic health record data for rare diseases.

4. Acknowledgements

5. References

[1] Critical Path Institute, Rare Disease Cures Accelerator-Data and Analytics Platform. URL: https://c-path.org/programs/rdca-dap/.

[2] N. Matentzoglu, C. Mungall, D. Goutte-Gattat, Ontology Development Kit, version 1.2.29, 2021. URL: https://hub.docker.com/r/obolibrary/odkfull. doi:10.5281/zenodo.4662066.

[3] R. C. Jackson, J. P. Balhoff, E. Douglass, N. L. Harris, C. J. Mungall, and J. A. Overton, ROBOT: A tool for automating ontology workflows, BMC Bioinformatics, volume 20, July 2019.