



An Ontology for Discoverable and Interoperable Radiology AI Models and Datasets

Charles E. Kahn, Jr., MD, MS, FSIM, University of Pennsylvania; Abhinav Suri; Safwan Halabi, MD; Hari Trivedi, MD

Introduction/Background

"Model cards" and "datasheets for datasets" provide valuable metadata to detail the performance, intended use, and potential limitations of AI resources. However, their format as unstructured text limits the ability to search for relevant resources and to automate their analysis. We sought to create a formal description to increase transparency and interoperability, reduce bias, and promote reproducibility of radiology AI models and datasets.

Methods/Intervention

The Radiology Model and Dataset Ontology (RMDO) was created to define attributes for AI models and datasets in radiology. RMDO references external ontologies and vocabularies, including RadLex, the LOINC/RSNA Radiology Playbook, and radiology common data elements (CDEs). RMDO incorporates RSNA content codes, PapersWithCode.com classifications of machine-learning methods and tasks, and the Metrics Reloaded listing of model-performance metrics. A JavaScript Object Notation (JSON) Schema was defined to allow serialization of RMDO-based descriptions of radiology AI models and datasets.

Results/Outcome

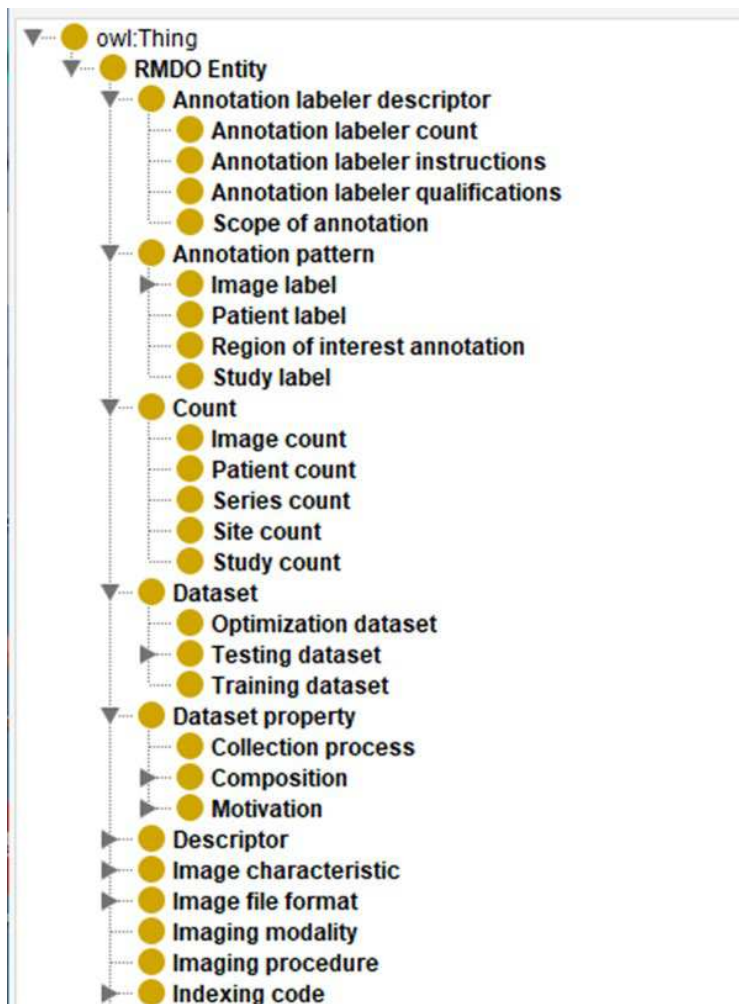
RMDO comprises 3,323 classes related by 4,403 logical axioms. The primary RMDO entity is a Project; its metadata describe authors, versioning, availability, licensing, and other features. Each Project consists of zero or more Models and zero or more Datasets. Model descriptions include architecture, intended uses, metrics, and ethical considerations. Dataset descriptions include imaging procedure, number of patients and images, image file format, output information, availability and licensing, partitions, annotation methods, and study cohort characteristics, such as demographics and disease prevalence. RMDO has been applied to datasets created for RSNA's AI competitions and for several published AI models.

Conclusion

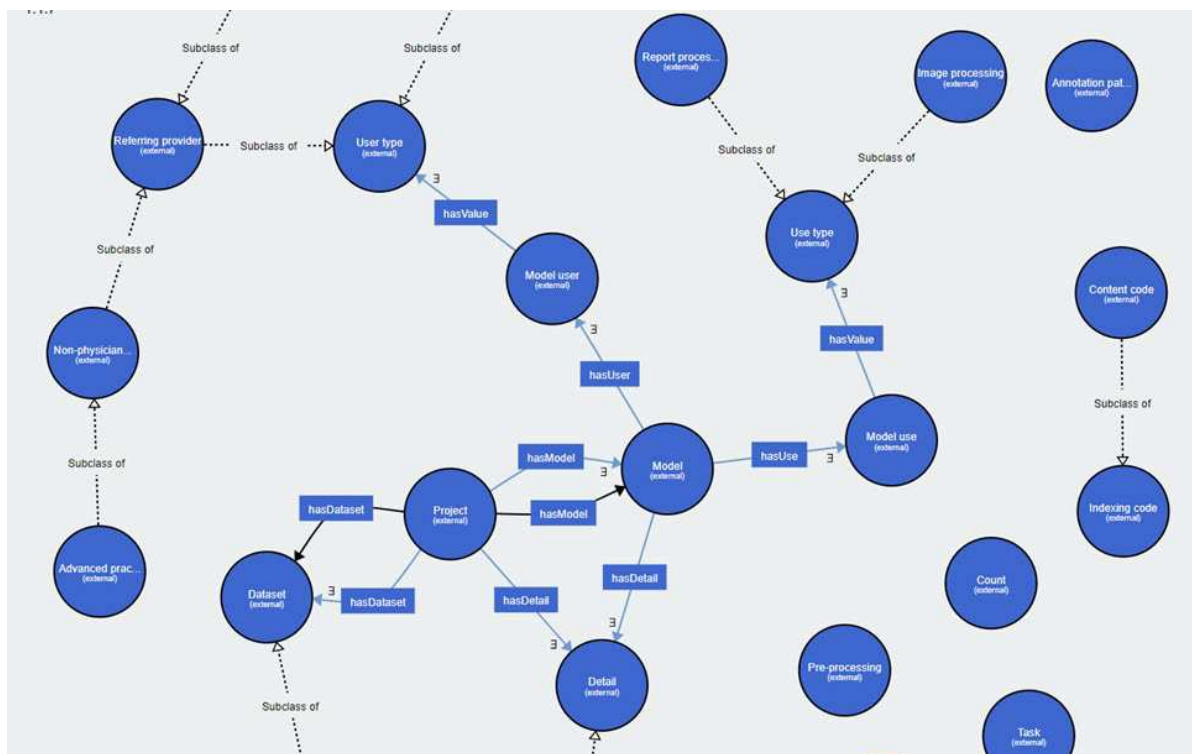
RMDO provides a standardized vocabulary that allows more effective classification and indexing of radiology AI models and datasets to make these resources more easily findable and accessible and to allow automated analysis of their underlying content.

Statement of Impact

An ontology to describe radiology AI models and datasets can make AI resources more findable and accessible. By allowing structured descriptions, the ontology helps promote reproducibility of AI models, and can aid in identifying and mitigating potential biases.



Portion of the RMDO class hierarchy, as shown in the Protege application.



Visualization of a subset of RMDO classes.

Keywords

Ontology; Model cards; Datasheets; Metadata; Interoperability; Radiology