Building Clinical Knowledge Graphs in FHIR RDF for Explainable AI Applications in Healthcare

Guoqian Jiang, MD, PhD[1], Emily Pfaff, PhD, MS[2], Christopher G. Chute, MD, Dr.PH[3], Guohui Xiao, PhD[4]
[1]Mayo Clinic, Rochester, MN, USA; [2]University of North Carolina, Chapel Hill, NC, USA; [3]Johns Hopkins University, Baltimore, MD, USA; [4]Free University of Bozen-Bolzano, Italy

Abstract

HL7 Fast Healthcare Interoperability Resources (FHIR) is rapidly becoming the standards framework for the exchange of electronic health record (EHR) data. FHIR Resource Description Framework (RDF) has become the first main-stream clinical data standard to incorporate the Semantic Web vision. The combination of FHIR, knowledge graphs and the Semantic Web enables a new paradigm to build classification and explainable artificial intelligence (AI) applications in healthcare. However, there is a critical need to build the FHIR-based data access and query on existing relational data sources to generate clinical knowledge graphs (CKGs) for facilitating standards-based semantic data integration, sharing and discovery in broader scientific research communities. The objective of the study is to develop and evaluate the methods and tools that expose the OMOP CDM-based clinical data repositories into virtual clinical knowledge graphs in FHIR RDF. We developed a FHIR-Ontop-OMOP system to generate virtual clinical knowledge graphs against the OMOP relational databases. The system consists of the following modules: 1) an input module including FHIR model ontology, OMOP CDM-based data repository, and OMOP-FHIR mappings represented by a mapping template; 2) a CKG generation module. We used an ontology-based data access tool known as Ontop as an engine; 3) a SPARQL endpoint and validation module; 4) a Semantic Web and AI application module. We evaluate the system in terms of the portability of the system across different institutions, the faithfulness of data transformation, and the conformance of generated RDF graphs to the FHIR RDF specification. This study is supported in part by the NIH FHIRCat R01 grant (R01 EB030529).