Enhancing Cultural Heritage Data Retrieval by Mapping COURAGE to CIDOC Conceptual Reference Model

Authors : Ghazal Faraj, Andras Micsik

Abstract : The CIDOC Conceptual Reference Model (CRM) is an extensible ontology that provides integrated access to heterogeneous and digital datasets. The CIDOC-CRM offers a "semantic glue" intended to promote accessibility to several diverse and dispersed sources of cultural heritage data. That is achieved by providing a formal structure for the implicit and explicit concepts and their relationships in the cultural heritage field. The COURAGE ("Cultural Opposition - Understanding the CultuRal HeritAGE of Dissent in the Former Socialist Countries") project aimed to explore methods about socialist-era cultural resistance during 1950-1990 and planned to serve as a basis for further narratives and digital humanities (DH) research. This project highlights the diversity of flourished alternative cultural scenes in Eastern Europe before 1989. Moreover, the dataset of COURAGE is an online RDF-based registry that consists of historical people, organizations, collections, and featured items. For increasing the inter-links between different datasets and retrieving more relevant data from various data silos, a shared federated ontology for reconciled data is needed. As a first step towards these goals, a full understanding of the CIDOC CRM ontology (target ontology), as well as the COURAGE dataset, was required to start the work. Subsequently, the queries toward the ontology were determined, and a table of equivalent properties from COURAGE and CIDOC CRM was created. The structural diagrams that clarify the mapping process and construct queries are on progress to map person, organization, and collection entities to the ontology. Through mapping the COURAGE dataset to CIDOC-CRM ontology, the dataset will have a common ontological foundation with several other datasets. Therefore, the expected results are: 1) retrieving more detailed data about existing entities, 2) retrieving new entities' data, 3) aligning COURAGE dataset to a standard vocabulary, 4) running distributed SPARQL queries over several CIDOC-CRM datasets and testing the potentials of distributed query answering using SPARQL. The next plan is to map CIDOC-CRM to other upper-level ontologies or large datasets (e.g., DBpedia, Wikidata), and address similar questions on a wide variety of knowledge bases.

Keywords : CIDOC CRM, cultural heritage data, COURAGE dataset, ontology alignment

Conference Title : ICCSIE 2020 : International Conference on Computer Science and Information Engineering

Conference Location : Sydney, Australia

Conference Dates : December 03-04, 2020

1