Geophysical Research Abstracts
Vol. 21, EGU2019-15856, 2019
EGU General Assembly 2019

ADOnIS: Managing Critical Zone Research Data Using Semantic Web Technologies

Alsayed Algergawy (1), Bernd Kampe (2), Hamdi Hamed (1), Birgitta König-Ries (1), and Udo Hahn (2)

(1) Institute for Computer Science, Friedrich Schiller University, Jena, Germany (alsayed.algergawy@uni-jena.de), (2) Jena University Language & Information Engineering (JULIE) Lab, Jena, Germany

The Collaborative Research Centre AquaDiva is a large collaborative project spanning a variety of domains, such as biology, geology, chemistry, and computer science with the common goal to better understand the Earth's critical zone, in particular, how environmental conditions and surface properties shape the structure, properties, and functions of the subsurface. Within AquaDiva large volumes of heterogeneous observational data are being collected. Besides this structured data, knowledge is also encoded in an unstructured form in scientific publications. To deal with the necessary integration of these data, we are currently implementing ADOnIS, the AquaDiva Ontology-based Information System.

Encoding conceptual domain knowledge ontologies supports the interlinking between semantically annotated data and thus enables the integration, synthesis, and querying of heterogeneous data. ADOnIS provides the following services:

o a set of semantic services that facilitate reuse and tailoring of existing domain ontologies and as such are key for the efficient and effective creation of semantic-enabled data management solutions.

o Integrated and seamless access to structured data and unstructured publications by making use of a variety of semantic technologies such as ontologies and natural language processing (NLP) tools. This includes (1) a specifically tailored domain ontology providing the fundamental concepts for the description of observational research data and publications as relevant for the AquaDiva project, (2) a data annotation scheme linking the research data with the conceptual knowledge and finally (3) a search mechanism and a graphical user interface for the integrated retrieval of heterogeneous research data and publications. By this, we aim to reduce the cognitive burden put on searchers while, at the same time, increasing the coverage and quality of search results.