

GraphScope - smart data access for the life sciences

Daniel M. Herzig and Günter Ladwig

SearchHaus GmbH, Karlsruhe, Germany

Abstract. GraphScope is the smart data access engine. GraphScope allows the user to access structured data, in particular RDF data, in a simple way by just entering keyword queries. The keyword queries are interpreted by GraphScope and precise results matching the information need are shown to the user. All technical aspects are hidden from the user. The user needs no knowledge about the data schema nor about the formal query language SPARQL. GraphScope recognizes the schema and builds an in-memory data synopsis, which is used to interpret the keyword queries. GraphScope is in particular suited for domain experts, who are familiar with the domain modeled by the knowledge base and who need fast and simple data access, but do not care much about the technical aspects of the IT infrastructure. GraphScope can be deployed on top of triple stores and provides a web interface.

We demonstrate GraphScope with a use case for the life sciences. The Gene Expression Atlas (<http://www.ebi.ac.uk/rdf/services/atlas/>) provided by the The European Bioinformatics Institute (EMBL-EBI) will be used to show common retrieval tasks in a life science scenario, such as 'Get experiments where the description contains diabetes' or 'Get samples derived from lung cells'. We will also give a view under the hood and discuss the main techniques of GraphScope. More information about GraphScope will be available at <http://graphscone.io> .