KGE - Knowledge Graph Engineering

A Metadata Knowledge Graph

KGE 2023 Project Proposal

Contents

Contents

Project Objective

- The context to be considered for this project, is an **Open Data environment**, concretely represented as a web portal through which several kind of information can be retrieved.
- The web portal allows the users to collects different kind of datasets that can be considered **singularly** or with meaningful **links** between them.
- In order to provide to the users the best way to retrieve the data they need, the datasets collected by the web portal have to be well described through **metadata**.
- The metadata is the main focus of this KGE project. With objective to build a KG able to represent the different kind of datasets available in the web portal, as well as the connection between them.

Contents

Context

- The web portal considered for this proposal, called "***Data Scientia***" collects all the resources produced by KGE processes following the iTelos methodology.
- As the KGE course will teach, there are several kinds of resources produced by the methodology, identified within 5 different inforamation layers:
  - L1: conceptual resources
  - L2: language resources
  - L3: teleology resources
  - L4: ontology resources
  - L5: data value resources
- In a single KGE process, resources representing instances of all the above layers are produced, **related to a single Purpose!**
- Resources relative to a single purpose have to be linked together by specific metadata.

Contents

Purpose

- The different resources, through the Data Scientia web portal, can be retrieved by the user **singularly**, but the user must be able to collect all the different kind of resources r**elated to a specific KGE process** (Purpose).

"A KG supporting the Data Scinetia web portal's users to find the most suitable resource for their needs, as well as all the resources linked by the one searched."

NOTE: this project will be extended after the KGE course. It is suitable for anyone who is interested in future collaboration with the KnowDive research group.

Contents

Data Resources

- To achieve the project objective the students have to analyze the different kind of resources collected within the Data Scientia web portal.
- For this reason the following resources and information will be provided to the students:
  - **Existing KGE projects** collecting instances for all the information layers.
  - **Data Scientia resources and projects** from which the most important information can be collected.

Contents

Schema Resources

The reference schemas provided to achieve the project's objective are:

- A **Data Scientia top level description and requirements**, in order define better the usage of the web portal, thus identifying the most important data descriptors to be transformed in metadata.
- A set of **resources specific requirements**, for each kind of resources collected within the web portal.
- **Schema.org**: for modelling general concepts https://schema.org/docs/full.html

Contents

Outcomes

- The KG produced as project's final outcome has to satisfy the requirements defined for the web portal users.
- The final KG structure will include the information provided by the resources metadata, **shaped as entities and relative attributes.**Such a KG will be **tested through a queryable endpoint**, through which the users will be able to find the information required to satisfy their purpose.
- The data used to populate and test the final KG will be those collected from existing KGE projects.
- The final KG, the metadata defined for each kind of resource, as well as the new purpose specific schema produced, will be collected as project's resources in order to be **reused by future projects**.

KDI : Knowledge and Data Integration

KGE 2023 Project Proposal

A Metadata Knowledge Graph