KGE - Knowledge Graph Engineering

KGE Projects Description

Existing examples of KGE projects

Fausto Giunchiglia

Contents

What is a KGE project

A Knowledge Graph Engineering (KGE) Project is a complete execution of process having the objective to build a **quality** and **reusable** KG able to support the requirements specified by a given **Purpose**.

- **Explicit goal**: A KGE project originates from specific needs of final users (the Purpose) that want to exploit the KGE results.
- **Implicit goal**: A KGE project aims to produce KGs (as well as resources composing KGs) with high level of reusability, with the objective to support a FAIR-based open data environment.

KGE project - Input

The **input** of a KGE project is composed by:

- **The Purpose** (functional requirement): the final user needs, to be satisfied.
- **Existing resources** (non-functional requirement): existing data and knowledge resources which have to be handled to built the KG. Often, this KGE input is provided as a data sources list from which the resources have to be collected.

KGE project - Output

The **output** of a KGE project is composed by:

- **Entity Graph** (EG): the KG produced which includes the entities required to satisfy the user's Purpose.
- **Project's documentation**: a project report which help external reader to understand the KGE process which led to the production of a specific KG.
- **Reusable resources**: all the (formalized) resources used to compose the final KG. They are provided as reusable resources in order to be exploited (singularly or as a set) in other KGE projects for different purposes.

Contents

KGE Project Documentation

The documentation aims to describe in details the KGE project, focusing on the following key points:

- Domain of Interest (DoI)
- Project Purpose
- Resources involved (phase by phase)
- Methodology phases development
- Final outcome
- Output evaluation
- Entity Graph exploitability (demo)
- Open issues

Thanks to the documentation produced, the project's output, as well as all the resources involved, become more understandable, clear and reusable by external users.

KGE Project Documentation - Objects

The project documentation objects that have to be produced are:

- **Project Report**: a textual document (defined following a specific template) describing the key points listed above (see previous slide).
- **Project Presentation**: a set of slides which aims to present the exploitability of the KGE final outcome.
- **EG Demo**: a demonstrator of the KGE final outcome. This demo should show how the EG developed during the project can be used to satisfy the user needs initially defined.

Contents

KGE Project's Artifacts

A KGE project, during its execution, produce several "artifacts" (not only data and knowledge resources):

- **Knowledge collected**: schemas, ontologies and other knowledge resources.
- **Data collected**: datasets.
- **Methodology intermediate outcomes**: this resource category includes all the artifacts (non included in the previous categories) that are used to build the final EG (such as requirement list, ER models, and others)
- **Final outcome**: the Entity Graph.
- **Metadata**: the meta information collected for both the knowledge and data resources handled in the project. The metadata allows to improve the resources reusability.
- **Code libraries**: scripts, small libraries and/or specific tools used to collect/filter/clean/format the data and knowledge resources.

Contents

KGE Project Planning

The efficiency of a KGE project is based on the **effort** and the **cooperation** among the different actors who play on it.

The roles covered by those actors are four:

- **Project Manage**r (PM): in charge of coordinating the whole projects, as well as the cooperation among the other roles.
- **Domain Expert** (DE): most of the time represented by the final user, she is the expert regarding the domain of interest (context in which the final K will be exploited).
- **Knowledge Engineer** (KE): responsible for the management of knowledge resources (KG's knowledge layer building).
- **Data Scientist** (DS): responsible for the management of data resources (KG's data layer building).

Project Planning - Roles

- PM, DE, DS and KE form a KGE project's *Team*.
- DS and KE are the most important roles along the process, and due to that, **they must be covered by, at least, two different people**.
- This means that the Team have to be composed by at least two actors, that in the worst case will cover all the four roles.

Project Planning - Gantt

Role	Task	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13
Project Manager	Coordination										
	Project report (general management)										
	Open project phase										
	Close project phase										
	Project presentation										
Domain Expert	Inception phase										
	Project report (Inception)										
	EG final evaluation										
	Project report (final)										
	Project Demo										
Knowledge Engineer	Kowledge metadata definiton										
	Inception phase										
	Project report (Inception)										
	Informal Modeling phase										
	Project report (Informal Modeling)										
	Formal Modeling phase										
	Project report (Formal Modeling)										
	Data Integration										
	Project report (Data Integration)										
Data Scientist	Data metadata definiton										
	Inception phase										
	Project report (Inception)										
	Informal Modeling phase										
	Project report (Informal Modeling)										
	Formal Modeling phase										
	Project report (Formal Modeling)										
	Data Integration										
	Project report (Data Integration)										

Contents

KGE Project Organization - Summary

KGE Project organization main aspects:

- Focus on **users needs** (project's Purpose).
- **Reusability** and **Shareability** of resources.
- Produce **quality resources** (using metadata) and **documentation**.
- **Effort** and **Cooperation** among the roles.
- Respect the project planning.

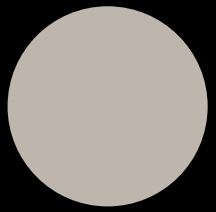
Contents

KGE Projects Examples

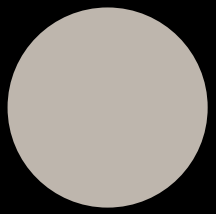
Projects form Knowledge and Data Integration course (2021-22):

- Events
- Transportation
- GeoParis
- Territory
- Education
- Facilities (University)
- Food Accomodation
- Health

KGE - Knowledge Graph Engineering



Fausto Giunchiglia



KGE Projects Description

Existing examples of KGE projects