KDI **Knowledge and Data Integration**

Integration of medical data on Covid-19

KDI 2020 Project Proposal

Contents

Contents

Medical Domain

- This project aims to collect and manage medical data. The medical data domain is composed by a large number of aspects involving persons, structures, medicines, treatments and not only those.

- For this project the focus selected is the current pandemic situation of the Covid-19. A huge amount of data have been generated due to the virus, regarding infected people, symptoms, medicines and therapy adopted and so on.

- Due to the fact that the each country generated its personal statistics on the pandemic, starting from several different data sources, the representations of the data on this topic are a lot defined and structured in different ways.

Contents

Competency Queries

In order to understand which kind of data have to be considered in this project, the competency queries which drive the inception phase can be like:

- Which aspects of the pandemic the project has to report ?

- Which countries are involved in the project ? Does the project consider the pandemic globally or only in a sub set of countries ?

- Which kind of information can be extracted from the data available in the final KG ?

- How those information is structured, and which are the most relevant elements within it ?

- .....

Contents

Standards

Following some of the standards adopted in the datasets to be collected:

- ICD-9 & ICD-10

- LOINC

- SNOMED CT

- .....

Contents

Knowledge Resources

The reference knowledge resources for the task are:

- All the medical standards needed to manage the data
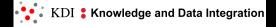
- OMOP

- .....

Contents

Data Resources

The data resources needed for this project are available on line from different sources. Most of the data can be downloaded or scraped by the web. A first examples can be:

- https://data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data

- https://www.cas.org/covid-19-antiviral-compounds-dataset

Contents

Outcomes

- The objective of this medical domain project is to collect and manage those data coming from different sources, in a Knowledge Graph (KG) which is able to maintain and provide data on various aspects of the pandemic.

- The project aims to collect data coming from different sources, in order to integrate them together generating a KG which cover heterogeneous scenarios.

- The project aims to improve the already available set of Knowledge and Language resources need to manage the medical data.

KDI : Knowledge and Data Integration

KDI 2020 Project Proposal

Integration of medical data on Covid-19