

Semantic Web Project - 2023 -2024

1 Overview

The semantic web project will run until the end of the course. It will be carried out in groups of 4 (ideally). It is best to choose members as people who attend the same practical sessions.

Task 1: Go to

https://docs.google.com/spreadsheets/d/1NxmnLykbMIDJcNsjaBFTeRuI4l_6sOyjiPfV9Swn6uU/edit?usp=s_haring and enter the names of the members of your group.

The goal will be to build a knowledge graph (and related things) in a specific domain (see below). You will learn different aspects of doing this as we go through the lectures and the practical sessions. It is therefore strongly recommended that you work on the project progressively, every week. In other words, as described in the “Tasks” section below, try to apply the things we learn in the course as we learn them, practicing them on the project and do not leave things to the last minute, when you will be busy with several other projects, exams, etc. Many resources are also linked on Arche¹, which you might want to look through.

2 Target application

Your goal is to build a knowledge graph to explore regularities, patterns, and trends in UFO sightings. Many countries have a government-related department that keeps track of reports of UFO sightings, including different characteristics of what was seen, of the condition under which it was seen and, in many cases, of the conclusion of any investigation that might have followed.² As a result, multiple datasets exist describing such sighting in different ways, including:

- <https://nuforc.org/databank/> (US, but including data from other countries. CSV subset available at <https://github.com/rfordatascience/tidytuesday/tree/2e9bd5a67e09b14d01f616b00f7f7e0931515d24/data/2019/2019-06-25>)
- <https://www.cnes-geipan.fr/fr/recherche/cas> (France, downloadable as CSV)
- <https://www.gov.uk/government/publications/ufo-reports-in-the-uk> (UK, ministry of defense)

The reason to build an RDF graph for this is that it could help integrate all those different sources of information into something consistent, which will be easy to query and explore, and that could evolve as new data sources are integrated.

3 Tasks

As mentioned previously, you will progressively learn to carry out the different tasks required to build such a knowledge graph as the course progresses. In the end (by the last session, i.e. the **21st of June at 12noon**), you should submit on Arche a short report describing the different steps you followed, starting with the ones you can already do now (base modelisation and conceptualisation) and including: 1) a base conceptualisation and modelisation for your graph (S1, S2, S3), an ontology in OWL for you graph (S4), queries that could be used on your graph once built (S5), the graph itself in RDF converted from existing data (S6), some code demonstrating the use of your graph (S7), validation test done and constraints on the graph (S8), and ideas about how your application could benefits from learning techniques and other ways of exploring graphs (S9, S10)

¹ <https://arche.univ-lorraine.fr/course/view.php?id=70229>

² The TV series “OVNI(S)” ([https://en.wikipedia.org/wiki/UFOs_\(TV_series\)](https://en.wikipedia.org/wiki/UFOs_(TV_series))) relates to the French department GEIPAN (<https://en.wikipedia.org/wiki/GEIPAN>) which is one such department. It is kind of funny.