Practical Session 7 – Building A Knowledge Graph

Today, we will continue to work on your project. Starting from the ontology that you have built already, the goal will be to import data from various sources in a way that fits your ontology.

You might not have finished your ontology yet, but make sure that you at least have an initial version and get on with the tasks here. You will work on this autonomously, but the practical session is an opportunity to ask questions and sort out some possible technical problems. Also, the work in the following might lead you to change your ontology, which is perfectly normal.

For the project, you are asked to build a knowledge graph of children stories. The idea is that your knowledge graph could be used by someone researching a particular theme, plot, type of character, or other aspects of children stories to find relevant ones, to compare stories with each other on those aspects, or to analyze trends in the way stories have evolved over time and cultures. As the lectures and practical sessions go, we will learn more about how that could be done and how we could use it.

At the end of all the practical sessions, you will have to submit:

1. The RDF code of the knowledge graph.

2. A short report briefly describing the steps you have gone through, the choices you have made, any SPARQL query you have used, and a description of any code you might have written.

So, keep notes of what you do and find!

T1: Bring back your ontology, make sure that it is open in Protégé and that it is at least a suitable basis to start populating data.

T2: From the results of the tasks in practical sessions 2 and 3, and others, identify which sources of RDF data could be used to add individuals to your ontolology. Write simple (to start with) construct queries to extract parts of RDF graphs to integrate into yours.

T3: Are there complicated values or ways to represent things in the imported data that require more than simple construct queries? Make your queries more complex or use code to change the data.

T4: You should have produced a number of RDF files at this point (the ontology, the results of the construct queries). Try loading them all into a new namespace in Blazegraph (with inference on). Write a few queries to check that it works as expected and test some of your competency questions.

T5: Can you find non-RDF data that you could also integrate?