Open Information Systems

Project Specification 2019-2020

Christophe Debruyne, WISE Lab, Vrije Universiteit Brussel
Maxim Van de Wynckel, WISE Lab, Vrije Universiteit Brussel

Introduction

The goal of the Open Information Systems *project* is to design and implement an ontology, use that ontology to annotate an existing system, and use the resulting knowledge base in a non-trivial manner.. The assignment will be carried out in teams of 4 people. Additionally, a report describing the design and choices will have to be produced.

Topic

The topics for this year's project are:

- Recipes: ingredients, steps, nutritional values, …
- Exercises: exercises, workouts, repetitions, calories burnt, …

We expect the groups to be equally divided across groups. You will provide a name for your group and system, as well as a one-paragraph description of your system. You may draw your inspiration from existing systems; WeightWatchers, Tasty, BasicFit,... What (nifty, yet feasible) feature distinguishes your application from the others?

Groups

Students will be assigned to a team of four members. Information about team members can be found on Canvas. We expect a fair partition of the workload between the members of the group. It will not be accepted that a student lets the teamwork for them without being involved in the project. Every problem with a team member should be reported as soon as possible.

Milestone 1: Conceptual schema

By November 8 you should provide the conceptual schema of your database modeled using (E)ER diagram. Use the (E)ER notation as taught at the VUB in the course Databases (Slides of this course can be found here). At this stage, the conceptual schema should be complete enough to capture every object that needs to be stored by your application. You should, therefore, provide use cases to describe the usage of the system and verify that the conceptual schema indeed covers all required objects. You should pay attention to the completeness and correctness of the model, and specifically the following aspects:

1. Weak entities should be correctly identified as such.
2. Upper and lower bounds for cardinalities should be correctly identified.
3. All entity types should have at least one natural key, i.e., a key that consists of one or more attributes. Artificial identifiers are only acceptable when unavoidable and motivated.

Milestone 2: Ontologies and Rules

We ask you to design and implement the ontologies and rules of your system. You will need to proceed in three steps:

1. Design your own ontologies and provide them as OWL files.
2. Design rules that will allow you to infer relevant information from your data.
3. You may adapt your conceptual schemas -- not to facilitate mappings, but to correct any conceptual errors you may encounter.

We encourage you to take into account existing ontologies and to refer to existing bodies of knowledge whilst developing the ontologies (e.g., base definitions on standards, references, etc.). The deliverable of this phase is a small report describing the ontologies. That report includes a WebVOWL visualization of your ontologies, a description of the rules you designed as well as a justification of why those rules are relevant in your system.

The deadline for this deliverable is the 22nd of November.

Milestone 3: Intermediate Presentations

The intermediate presentation will take place on the 27th of November from 9.00 to 13.00. This presentation will be the occasion for your group to present your application. Each presentation should last a maximum of 12 minutes. During that time you should be able to

- Give a small overview of the functionalities offered by your system.
- Present and explain the design choices you made for your ontologies.
- Present one scenario showing how your system could be used through the SPARQL endpoint in order to access information. You must show which part of the conceptual schema and the ontologies are involved. It is sufficient to present a scenario, no demonstration of a fully working SPARQL endpoint is required at this stage.

Milestone 4: Using Ontologies and Final Report

SPARQL Endpoint

You will have to provide a SPARQL endpoint accessible within the application that allows external users to query your system. You can either do it using Ontop's Protégé Plugin[1], create an ETL pipeline with R2RML and a triplestore of your preference, or any other libraries of your choice, *as long as you motivate*

[1] https://ontop.inf.unibz.it/

its usage. Your final report should detail and explain the *mapping* between your relational database and the ontologies.

During the final presentation, you will have to execute relevant queries on your SPARQL endpoint in order to demonstrate how you used ontologies and rules in order to infer useful information from your information system. We expect your endpoint to be able to at least some simple *competency questions*.

Applying the Ontology

This aspect of the project is twofold. First, teams from different groups are highly encouraged to develop, together, a scenario in which two SPARQL endpoints are combined. For example, what recipes correspond with a person's workout. Groups who worked together may share and include the same section in their reports.

Secondly, you should demonstrate that you can use the ontology and/or endpoint in a non-trivial manner. Groups can choose to either investigate a specialized topic from the class or look (e.g., ontology alignment, automatic linking,...) or propose the development of a minimal viable product (e.g., a script) that demonstrates how the ontology and/or endpoint can be used to implement a particular task.

Final Report

The final report should at least go through the following points:

1. How did you divide and manage the workload inside your team?
2. Motivate the design of your conceptual schema for the database.
3. Motivate the design of your ontology.
4. What are the rules to infer additional information from your database? Why are they relevant to your information system?
5. Present and explain the final ontologies, rules, and specifications of the mapping including the extensions that were required by the additional functionality.
6. Describe the non-trivial demonstrator
7. Provide a link to a repository or ZIP containing all artifacts (with detailed steps on how to set up, compile, … your application).

Dot not forget to provide the number of your team, the names, email addresses and enrollment numbers of all members of your team as well as the current academic year in the title page. The deadline for the final report is 21st December 2019.

Milestone 5: Final Presentation (exam)

The final presentation is to be scheduled (details will follow via Canvas) and will be a private defense of your project. We expect you to make good use of the feedback you received from the intermediate

presentation. The final application should be complete enough to illustrate different use case scenarios of your system, allowing you to justify the different design choices for ontologies and rules as well as demonstrate the database mapping.

Administrative Requirements

Every deliverable has to be sent by mail to the assistant Maxim.Van.de.Wynckel@vub.be with the following subject: **OIS-2019-*subject-student1-student2-student3-student4***

Here "student1" is the last name of the first team member and "subject" is the subject of the deliverable (see below). Mails that are not sent according to this header will be ignored! Make sure to send only one mail per team. For each deliverable and report, we expect a report worthy of an MSc student. Reports should be well structured (title, abstract, introduction – middle – conclusion, references). Plagiarism is not tolerated and we expect that resources are properly cited/attributed.

Deadlines Recap

The following gives an overview of all the relevant deliverables and dates within the project:

- November 8 at 23:59 - Conceptual schema (by email with subject OIS-2019-SCHEMA-student1-student2-student3-student4)
- November 22 at 23:59 - Ontologies and rules (by email with subject OIS-2019-ONTO-student1-student2-student3-student4)
- November 27 - Intermediate presentation (public)
- December 20 at 23:59 - Final Report (by email with subject OIS-2019-REPORT-student1-student2-student3-student4)
- Final presentation (dates will be published on the faculty's website)

Contact Information

If you have any questions, please feel free to contact the assistant Maxim.Van.de.Wynckel@vub.be or the professor at christophe.debruyne@vub.be.