

# Semantic Web

(KEN3140)

## Assignment 1:

### Resource Description Framework (RDF)

07-09-2022

**Deadline:** 20-09-2022

### Assessed learning objectives

This assignment is going to assess if the following learning objectives have been achieved.

Students are able to:

1. Conceptualize domain knowledge using the RDF abstract data model
2. Formulate RDF triples that accurately capture knowledge in a specific domain
3. Identify, distinguish between and select the appropriate RDF features for capturing a piece of information
4. Distinguish between constructs in RDF that are meant to capture the *meaning* of the data vs. constructs that are meant to capture the data directly
5. Assess the syntactic validity of an RDF graph
6. Assess the *basic* quality of an RDF graph. In particular, the degree to which the graph makes use of RDF and Linked Data “best practices” for representing information in a complete and concise way

## Assignment task description

For this assignment you will:

1. Construct a diagram of a **high school community**. This can be either your own high school or a fictional high school.
2. Construct an RDF graph (a set of RDF triples) providing descriptive information about the community members in your diagram created in Task 1 and the relationships they share with each other.

## RDF graph requirements

These are requirements that make it easier for instructors to evaluate your assignments. You will lose points if you do not meet these requirements. However, this is **not** the main or only criteria for scoring points in the assignment (see Grading criteria below):

1. The graph should be formulated using Turtle syntax
2. Only these school-related relationships should be captured in the graph:
  - a. Teacher and student relationships
  - b. Classmate relationship
  - c. Colleague relationship (e.g. Person A is a colleague of Person B)
  - d. Gender relationship (e.g. Person A has gender non-binary, Person B has gender unspecified, or Person C has gender female). Gender should be modeled as an **object property** in your graph and **not a type**
  - e. **Do not** include any other kinds of relations between school members in this assignment
3. A **minimum of 5** different **data properties (that are not of type xsd:string)** should be captured to describe each school member. Some example data properties you could include are first name, last name, age, height, date of birth etc. but you are not limited only to these. Of course, you are welcome to create artificial data to preserve privacy if you would like.
4. The graph must have a **minimum of 7** distinct school members and a **maximum of 15** distinct school members
5. Every school member in your graph should be related to another school member in the graph in some way (there can be no school members with no relation to another school member) **except one**. Why except one? It is not possible to specify all **teachers and students** for every entity in a **finite** school community graph! :)

## Best practice requirements of the graph

For objects and subjects in your triples you are welcome to create your own **valid** IRIs and namespaces. However, for **predicates and types** in your graph, you must **reuse** terms from [Schema.org](https://schema.org) or either [DBpedia's](https://dbpedia.org) or [Wikidata's](https://wikidata.org) vocabulary (or an alternative vocabulary that you can find). Additionally, if you are creating a school community of historical figures or celebrities, we expect that you determine whether these entities already have an assigned IRI on the Linked Open Data web first and **reuse** these.

1. What is Schema.org? It is a vocabulary defining general terms about people, places, events etc. that the major search engines (Google, Yahoo, Yandex etc.) maintain to help annotate Web content to make them easier to index by the search engines.
2. What is Wikidata? It is a large-scale RDF graph (called a knowledge graph) of structured information from Wikipedia, the Wikidata community also maintains a vocabulary (called the [Wikidata ontology](https://wikidata.org/wiki/Wikidata:Ontology)) alongside the knowledge graph to define types and relations that it uses.
3. What is DBpedia? It is a knowledge graph similar in goal to Wikidata. One major difference between them is that Wikidata information pages can be edited and submitted by users with accounts. This is similar to how Wikipedia pages are managed. Whereas, DBpedia information is almost exclusively automatically mined from [Wikipedia infoboxes](https://en.wikipedia.org/wiki/Wikipedia:Infoboxes).

## Grading criteria

1. Graph requirements are met (see previous two sections)
2. The syntactic validity of your RDF graph and the IRIs mentioned in it
3. The degree to which RDF best practices are followed
4. Correct usage of RDF features
5. Conciseness in the syntactic representation of the graph
6. Appropriate use of vocabulary (ontology) terms and the degree to which you reuse terms and IRIs
7. The degree to which the content in your school community diagram matches the content in your RDF graph

## Helpful information / FAQs

1. "I don't know what a school community looks like". Then search for school community diagrams online and it should become clear what they look like.  
<https://study.com/academy/lesson/the-school-community-definition-members.html>
2. you **MUST** use a software diagram tool to draw your school community. The entities and relationships in your diagram should be clearly marked and it should accurately reflect your RDF graph. Please also include the data properties (attributes) of the school members in the diagram. Here is a helpful video to start your research about school communities: [https://www.youtube.com/watch?v=PM79Epw\\_cp8&ab\\_channel=CGPGrey](https://www.youtube.com/watch?v=PM79Epw_cp8&ab_channel=CGPGrey)

3. Terms found in external vocabularies (e.g. <https://schema.org/>) do not necessarily have to be named/spelled exactly the same as what you expect. The focus is on the **meaning** of a term, not necessarily its string or textual representation. E.g., maybe you are looking to reuse a *relation* that captures the meaning that x is “teacherOf” y where x and y represent school members in your graph. You may find that *Schema.org* only contains a relation called “<https://schema.org/instructor>” and does not contain a relation called “<http://schema.org/teacherOf>”. This does not necessarily mean that the former relation is unsuitable. **Read the text description** of “<http://schema.org/instructor>” on the Schema.org website and decide from this description if the relation captures what you intend to capture **semantically**. If it is, you can reuse it! So do not get distracted by slight differences in spelling or variations in the form of the word.

## Deadline & submission instructions

The deadline for your assignment is **20-09-2022**. You should upload the following files separately on Canvas under Assignments>Assignment 1:

1. **Mandatory:** a single **.png** image file containing a clear, legible diagram of your school relations. It should include only class and property types ( do not include data properties! ). Name this file: “your-name\_your-student-ID\_ken3140\_assignment1.png”  
**Note:** You must use tools to create your diagram ( handwriting is not acceptable! )
2. **Mandatory:** a single RDF document with your completed school graph in **Turtle** syntax with **.ttl** extension. The filename should be:  
“your-name\_your-student-ID\_ken3140\_assignment1.ttl”
3. **Optional:** if you created your RDF document using a Jupyter notebook, upload your notebook as well with the filename:  
“your-name\_your-student-ID\_ken3140\_assignment1.ipynb”. **Ensure** that it is the same notebook which generated your uploaded RDF graph i.e., the file that is named:  
“your-name\_your-student-ID\_ken3140\_assignment1.ttl”.

**Please make sure to follow the assignment instructions carefully and meet all the requirements!** We will provide an example solution and opportunity to discuss this during the course. You will receive a grade out of 10 points for this assignment.

## Some relevant learning materials:

1. KEN3140 Lecture 1 & 2 slides (Canvas)
2. KEN3140 Lab 2 notebook and materials (Canvas)
3. [RFC specification for IRIs](#)
4. [Linked Open Vocabularies](#)
5. [RDF abstract syntax specification](#)
6. [Turtle RDF syntax specification](#)
7. [Schema.org vocabulary](#)
8. [Wikidata RDF graph](#)

9. [XML data types](#)

## Questions

Christopher Brewster ( [christopher.brewster@maastrichtuniversity.nl](mailto:christopher.brewster@maastrichtuniversity.nl) )

Remzi Celebi ( [remzi.celebi@maastrichtuniversity.nl](mailto:remzi.celebi@maastrichtuniversity.nl) )