Table of Contents

1. Introduction

Knowledge graphs are information represented in the form of graphs where nodes are entities and the relationships between them are the edges. The fact that they help form better connections between data helps machines to process and connect loads of information for better results. Knowledge graphs are considered to be the foundation to artificial intelligence and are the main tool that is required to build intelligent agents such as chatbots that help answer questions based on the user's input. Each fact in a knowledge graph is represented as a triple (subject, predicate, object) which are interconnected with each other to form the whole graph.

In this project, we create a knowledge graph that stores the information about Universities and its academic details. Here, we concentrate on information about Concordia University alone. The goal of this project is to build a knowledge graph in such a way that it can answer questions about the University through SPARQL queries. This project is the first step of building an intelligent agent that is capable of replying to questions related to the university which would be further done in our next project. Knowledge graphs are created from the data extracted from the web. Here, we take information about the courses, topics related to it and further details regarding it from the Concordia University websites and its open databases.

We have constructed the University knowledge graph using RDF and RDFS standard in turtle format. The graph is built to store information about universities along with its dbpedia entry uri. But since we restrict our project to just one university, the graph stores information about all the courses offered by Concordia University and the topics that are covered in each course. It also stores student information such as name, email etc. For testing purposes, we have used dummy information about ten students and their course history.

2. Description

The University graph is stored in the form of RDF schema and contains mainly five components, namely, 'University', 'Course', "Topic' and 'Student'. These are stored in the form of the classes. Each of these components has attributes that store more information about them.

2.1.1 University

This contains the name and the dbpedia URI for the given university. Since we use only Concordia University in our project, this has only once instance.

2.1.2 Course

This contains the name of the course, course identifier and the description for the course.

2.1.3 Topic

This contains the name of the topic and the dbpedia URI entry for the topic.

2.1.4 Student

This contains the name of the student; first name and last name, ID number and email address of the student.

2.1.5 Course Grade

This contains a course class instance along with the grade that the student scored for that course.

3. Architecture

In this section, we describe the architecture of our RDF schema in details.

3.1 Vocabularies Used

3.1.1 Reused Vocabularies

3.1.1.1 RDFS Schema

http://www.w3.org/2000/01/rdf-schema#:

rdf:label - This is used to provide a human readable name for all the entities in our KG.

rdfs:comment - A description of the subject resource.

rdfs: domain

rdfs:range

3.1.1.2 RDF Schema

rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#

3.1.1.3 XML Namespace

xml: http://www.w3.org/XML/1998/namespace

3.1.1.4 XML Schema

xsd: http://www.w3.org/2001/XMLSchema#

3.1.1.5 FOAF Vocabulary

foaf: http://xmlns.com/foaf/0.1/

3.1.1.6 DBpedia Property

dbp: http://dbpedia.org/property/

3.1.1.7 DCMI Metadata Terms

dc: http://purl.org/dc/elements/1.1/

3.1.2 Our Schema

3.1.2.1 ISP

isp: http://intelligentsystemproj1.io/schema#

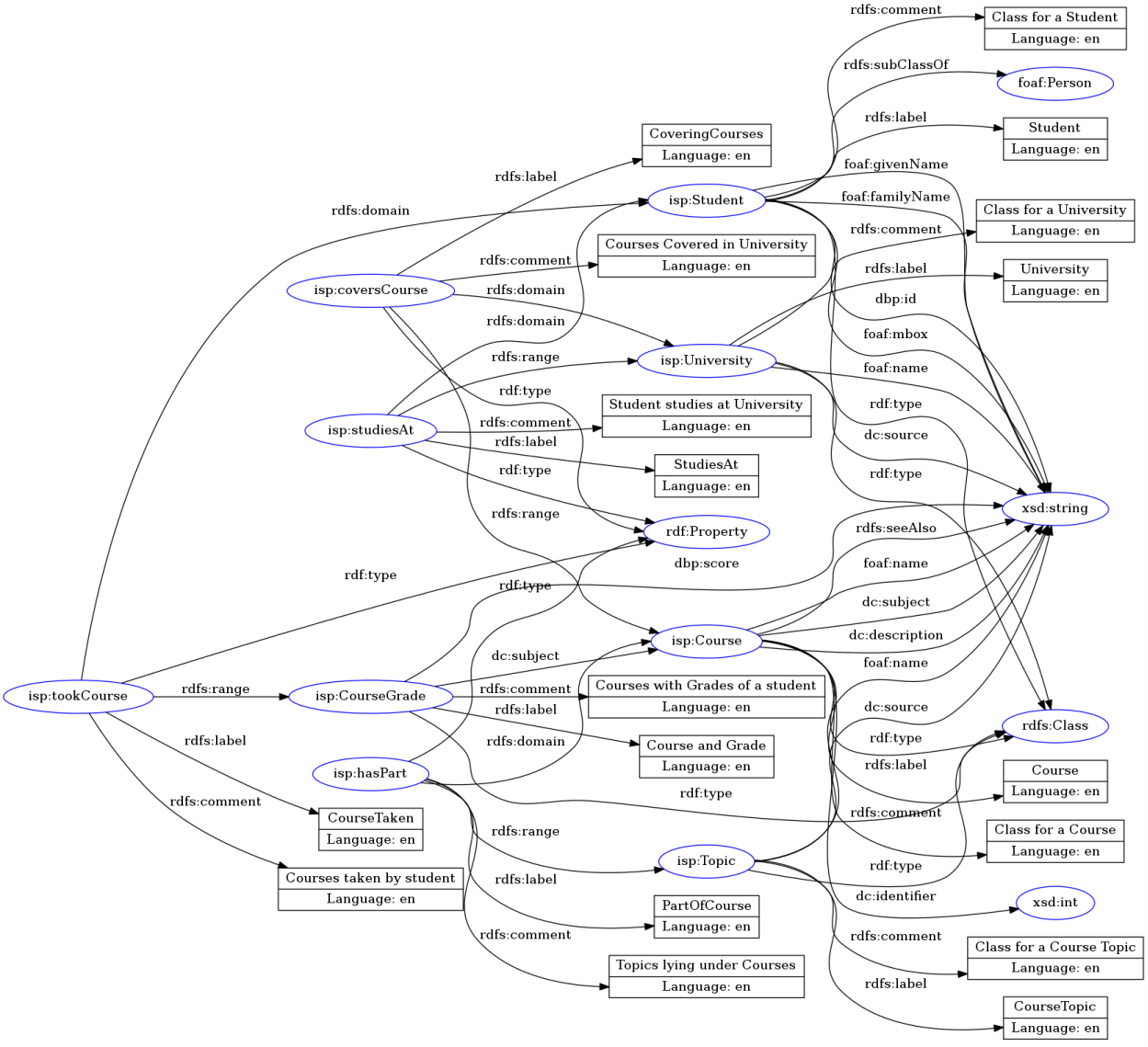
rdfs:comment
Class for a Student
Language: en
rdfs:subClassOf
foaf:Person
rdfs:label
Student
Language: en
foaf:givenName
foaf:familyName
rdfs:comment
Class for a University
Language: en
rdfs:label
University
Language: en
dbp:id
foaf:mbox
foaf:name
rdf:type
dc:source
rdf:type
xsd:string
rdfs:seeAlso
foaf:name
dc:subject
dc:description
foaf:name
dc:source
rdf:type
rdfs:label
rdfs:Class
Course
Language: en
rdfs:comment
Class for a Course
Language: en
rdf:type
dc:identifier
xsd:int
rdfs:comment
Class for a Course Topic
Language: en
rdfs:label
CourseTopic
Language: en
CoveringCourses
Language: en
rdfs:label
isp:Student
rdfs:domain
Courses Covered in University
Language: en
rdfs:comment
isp:coversCourse
rdfs:domain
rdfs:domain
rdfs:range
isp:University
rdf:type
Student studies at University
Language: en
rdfs:comment
isp:studiesAt
rdfs:label
StudiesAt
Language: en
rdf:type
rdfs:range
rdf:Property
dbp:score
rdf:type
rdf:type
isp:Course
dc:subject
isp:tookCourse
rdfs:range
isp:CourseGrade
rdfs:comment
Courses with Grades of a student
Language: en
rdfs:label
rdfs:label
rdfs:domain
Course and Grade
Language: en
isp:hasPart
rdf:type
rdfs:comment
CourseTaken
Language: en
rdfs:range
isp:Topic
Courses taken by student
Language: en
rdfs:label
PartOfCourse
Language: en
rdfs:comment
Topics lying under Courses
Language: en
Namespaces:
rdfs: http://www.w3.org/2000/01/rdf-schema#
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
xml: http://www.w3.org/XML/1998/namespace
xsd: http://www.w3.org/2001/XMLSchema#
foaf: http://xmlns.com/foaf/0.1/
isp: http://intelligentsystemproj1.io/schema#
dbp: http://dbpedia.org/property/
dc: http://purl.org/dc/elements/1.1/

Figure 1