Project Report On

"Studybot"

Submitted in Partial Fulfillment Of

COMP 6741: Intelligent Systems

To

Dr. René Witte

By

Divya Bhosale 40094565

Mitalee Naik 40115689

Studybot is an intelligent agent that can answer university course-related questions using a knowledge graph and natural language processing.

Competency Questions:

1) What is COMP 474 about?
2) What is the content for Lecture 4 of course COMP 6741 ?
3) What is the corresponding Lab for Lecture 10 of Machine Learning course?
4) What courses cover Machine Learning topic?
5) Which lecture is available for machine learning topic?
6) What is the lecture and lab content for Intelligent Systems?
7) What topics are related to Lecture 2 of COMP 6741 course?
8) What is the syllabus for COMP 6741?
9) What departments are available under Gina Cody School of Engineering & Computer Science?
10) What graduate courses are offered by Gina Cody School of Engineering & Computer Science?

Vocabulary:

We have reused the following vocabularies.

RDF	type	Define class
	property	Define properties
RDFS	class	Define class
	subClassOf	Define subclass of a class (similar properties)
	label	Provide label to entity
	comment	Provide additional information about the entity
	seeAlso	To associate entity with dbpedia/wikidata entries
FOAF	name	Provide a name to entity
OWL	sameAs	Provide similar instance of entity on dbpedia
NPG	number	Provide a number to entity
DC	identifier	Provide identifier to entity
	description	Provide description of an entity
SCHEMA	isPartOf	Associate two entities with each other
DBO	type	To define type of university
XSD	string	Datatype of literal
	integer	Datatype of literal

We have reused the focu schema from the lab.
focu: <*http://focu.io/schema#*>
focup: <*http://focu.io/schema/property#*>
Eg:

1. *focu:Course a rdfs:Class ;*
   rdfs:label "Courses"@en ;
   rdfs:comment "Represents the courses offered in the Universities included in the chatbot."@en .

Above is the entity defined to represent a Course as a class.

2. *focup:belongs_to*
   rdf:type rdfs:Property ;
   rdfs:label "Belongs to"@en ;
   rdfs:comment "Lecture/Topic belongs to the course."@en .

Above is the entity defined to represent a property which defines the relation between lecture/topic with course.

focu	University	Describe university as class
	Course	Describe course as class
	Lecture	Describe lecture as class
	Lab	Describe lab as class which is subclass of lecture
	Topic	Describe topic as class
focup	belongs_to	Describe relation between two classes
	has_outline	Include outline to course
	has_level	Specify level (Undergraduate/graduate)
	includes_slides	Include lecture slides
	includes_worksheets	Include worksheets
	includes_reading	Includes lecture contents such as reading materials
	includes_other_material	Includes lecture contents such as other materials -videos,images
	includes_content	Includes lecture contents

Knowledge Base Construction:

a) Dataset:

- Our dataset consists of information extracted from the below CSVs available on Concordia University's open data portal:-
    a. CU_SR_OPEN_DATA_CATALOG.csv
        i. Course ID
        ii. Course Subject
        iii. Course Number
        iv. Level
    b. CU_SR_OPEN_DATA_CATALOG_DESC.csv
        i. Course ID
        ii. Course Description
    c. CU_SR_OPEN_DATA_SCHED.csv
        i. Course ID
        ii. Faculty Description
- University

Property	**Description**
rdf:type	Is of type dbo:University
rdfs:label	Stores the University Name
rdfs:comment	University description in short
rdfs:seeAlso	Homepage URL
owl:sameAs	Dbpedia dbr:Concordia_University

- Course

Property	**Description**
rdf:type	Is of type focu:Course
focup:belongs_to	Course belongs to Course Subject
focup:has_level	GRAD/UGRD
npg:number	Course Number
dc:identifier	Stores the Course ID

schema:isPartOf	Department Information
foaf:name	Course Name
rdfs:seeAlso	Course information link (CSV)
rdfs:comment	Course description in short

- Lecture

Property	**Description**
rdf:type	Is of type focu:Lecture
focup:belongs_to	Lecture belongs to a Course
rdfs:comment	Lecture description in short
npg:number	Lecture Number
foaf:name	Lecture Name
focup:includes_slides	Includes lecture contents such as slides
focup:includes_worksheet	Includes lecture contents such as worksheet
focup:includes_reading	Includes lecture contents such as reading materials
includes_other_material	Includes lecture contents such as other materials -videos,images

- Lab

Property	**Description**
rdf:type	Is of type focu:Lab
focup:includes_slides	Includes lecture contents such as slides
rdfs:comment	Lab description in short
npg:number	Lab Number
foaf:name	Lab Name

rdfs:subClassOf	Lab is a subclass of Lecture

- Topic

Property	**Description**
rdf:type	Is of type focu:Topic
rdf:label	Topic Label
owl:sameAs	DBpedia link to the topic
schema:isPartOf	Topic could be part of a Lecture/Lab
foaf:name	Topic Name

b) Process and Developed Tools for KB Construction

- We have constructed the Knowledge Base using python pandas by reading the CSVs available on Concordia's open data portal. All required columns are extracted and stored in 'courses_list.csv'. Duplicate courses are dropped and triples are created by iterating over this file.
- Lecture and Lab content is generated by iterating over the folder/subfolder structure and creating triples for every Lecture/ Lab.
- Topics are added manually.
- Steps to reconstruct the KB :-.
    - Step 1 : Run the generate_university_courses.py file
    - Step 2 : Run the generate_lecture_content.py file
    - Step 3: topics.nt file is manually created.

Queries:

1) What is COMP 474 about?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX np: <http://www.nanopub.org/nschema#>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
SELECT ?course ?description
        WHERE
        {
                        ?course rdf:type focu:Course .
                        ?course dc:description ?description .
                        ?course npg:number 474 .
                        ?course focup:belongs_to "COMP" .
        }
```

Output:

	course	description
1	focu:COMP_474/	"Rule-based expert systems, blackboard architecture, and agent-based. Knowledge acquisition and representation. Uncertainty and conflict resolution. Reasoning and explanation. Design of intelligent systems. Project. Lectures: three hours per week. Laboratory: two hours per week. Prerequisite: COMP 352 or COEN 352. "

2) What is the content for Lecture 4 of course COMP 6741 ?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX np: <http://www.nanopub.org/nschema#>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?lecture ?slides ?worksheet
        WHERE
        {
                        ?lecture rdf:type focu:Lecture .
                        ?lecture focup:includes_slides ?slides .
                        ?lecture focup:includes_worksheet ?worksheet .
                        ?lecture npg:number 4 .
```

```
                    ?lecture focup:belongs_to ?course.
                    ?course rdf:type focu:Course.
                    ?course focup:belongs_to "COMP".
                    ?course npg:number 6741.
        }
```

Output:

	lecture	slides	worksheet
1	focu:COMP_6741Lecture4/	<file:///D:/CONCORDIA /COMP_6741_Intelligent_Systems/Project_Assignment /comp6741project/database/Concordia_University /COMP_6741/Lecture/4_Knowledge_Based_Queries /slides04.pdf>	<file:///D:/CONCORDIA /COMP_6741_Intelligent_Systems/Project_Assignment /comp6741project/database/Concordia_University /COMP_6741/Lecture/4_Knowledge_Based_Queries /worksheet04.pdf>

3) What is the corresponding Lab for Lecture 10 of Machine Learning course?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX np: <http://www.nanopub.org/nschema#>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT ?lab ?lab_name ?lab_slides
        WHERE
        {
                    ?lecture rdf:type focu:Lecture.
                    ?lecture npg:number 10 .
                    ?lecture focup:belongs_to ?course.
                    ?course foaf:name "MACHINE LEARNING".
                    ?lab rdf:type focu:Lab.
                    ?lab focup:includes_slides ?lab_slides.
                    ?lab rdfs:subClassOf ?lecture.
                    ?lab foaf:name ?lab_name.
        }
```

Output:

	lab	lab_name	lab_slides
1	focu:COMP_6321Lab10/	"10_Gans"	<file:///D:/CONCORDIA /COMP_6741_Intelligent_Systems/Project_Assignment /comp6741project/database/Concordia_University /COMP_6321/Lab/10_Gans/lab10.ipynb>

4) What courses cover Machine Learning topic?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX np: <http://www.nanopub.org/nschema#>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
SELECT ?topic ?course_name ?course_subject ?course_number
        WHERE
        {
                        ?topic rdf:type focu:Topic .
                        ?topic foaf:name "Machine Learning" .
                        ?topic schema:isPartOf ?course .
                        ?course rdf:type focu:Course .
                        ?course foaf:name ?course_name .
                        ?course npg:number ?course_number .
                        ?course focup:belongs_to ?course_subject .
        }
```

Output:

	topic	course_name	course_subject	course_number
1	focu:/machine_learning	"INTELLIGENT SYSTEMS"	"COMP"	"6741"^^xsd:integer
2	focu:/machine_learning	"MACHINE LEARNING"	"COMP"	"6321"^^xsd:integer
3	focu:/machine_learning	"Intelligent Systems"	"COMP"	"474"^^xsd:integer

5) Which lecture is available for machine learning topic?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX np: <http://www.nanopub.org/nschema#>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
```

```
SELECT ?topic ?lecture ?lecture_name ?lecture_slides
        WHERE
        {
                        ?topic rdf:type focu:Topic .
                        ?topic foaf:name "Machine Learning" .
                        ?topic schema:isPartOf ?lecture .
                        ?lecture rdf:type focu:Lecture .
                        ?lecture foaf:name ?lecture_name .
                        ?lecture focup:includes_slides ?lecture_slides .
        }
```

Output:

	topic	lecture	lecture_name	lecture_slides
1	focu:/machine_learning	focu:COMP_6741Lecture7/	"7_Introduction_to_Machine_Learning"	<file:///D:/CONCORDIA /COMP_6741_Intelligent_Systems /Project_Assignment/comp6741project /database/Concordia_University /COMP_6741/Lecture /7_Introduction_to_Machine_Learning /slides07.pdf>

6) What is the lecture and lab content for Intelligent Systems?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX np: <http://www.nanopub.org/nschema#>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
SELECT ?course_name ?lecture_name ?lab_name ?slides ?worksheets ?lab_slides
        WHERE
        {
                        ?course rdf:type focu:Course .
                        ?course foaf:name ?course_name .
                        ?course foaf:name "INTELLIGENT SYSTEMS" .
                        ?lecture focup:belongs_to ?course .
                        ?lab rdfs:subClassOf ?lecture .
                        ?lecture foaf:name ?lecture_name .
                        ?lab foaf:name ?lab_name .
                        ?lecture focup:includes_slides ?slides .
                        ?lecture focup:includes_worksheet ?worksheets .
```

?lab focup:includes_slides ?lab_slides.

}

ORDER BY ?lecture_name

Output:

Showing 1 to 8 of 8 entries Search: Show 50 entries

	course_name	lecture_name	lab_name	slides	worksheets	lab_slides
1	"INTELLIGENT SYSTEMS"	"1_Introduction"	"1_Python"	<file:///D:/CONCORDIA /COMP_6741_Intelligent_ Systems /Project_Assignment /comp6741project /database /Concordia_University /COMP_6741/Lecture /1_Introduction /slides01.pdf>	<file:///D:/CONCORDIA /COMP_6741_Intelligent_ Systems /Project_Assignment /comp6741project /database /Concordia_University /COMP_6741/Lecture /1_Introduction /worksheet01.pdf>	<file:///D:/CONCORDIA /COMP_6741_Intelligent_S ystems /Project_Assignment /comp6741project /database /Concordia_University /COMP_6741 /Lab/1_Python/Lab1.pdf>
2	"INTELLIGENT SYSTEMS"	"2_Knowledge_Graphs"	"2_RDF"	<file:///D:/CONCORDIA /COMP_6741_Intelligent_ Systems /Project_Assignment /comp6741project /database /Concordia_University /COMP_6741/Lecture /2_Knowledge_Graphs /slides02.pdf>	<file:///D:/CONCORDIA /COMP_6741_Intelligent_ Systems /Project_Assignment /comp6741project /database /Concordia_University /COMP_6741/Lecture /2_Knowledge_Graphs /worksheet02.pdf>	<file:///D:/CONCORDIA /COMP_6741_Intelligent_S ystems /Project_Assignment /comp6741project /database /Concordia_University /COMP_6741/Lab/2_RDF /Lab2.pdf>
				<file:///D:/CONCORDIA /COMP_6741_Intelligent_ Systems /Project_Assignment	<file:///D:/CONCORDIA /COMP_6741_Intelligent_ Systems /Project_Assignment	<file:///D:/CONCORDIA /COMP_6741_Intelligent_S ystems

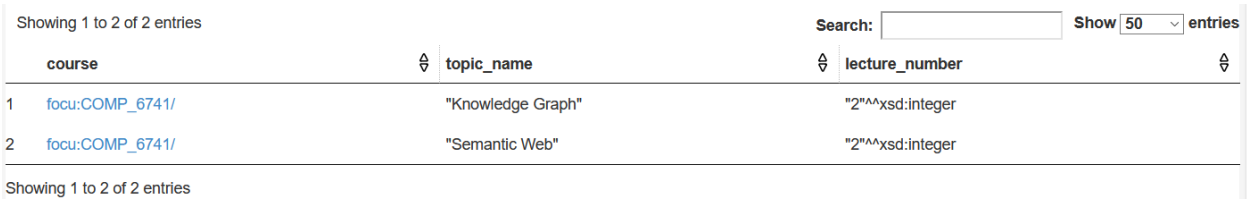
7) What topics are related to Lecture 2 of COMP 6741 course?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
SELECT ?course ?topic_name ?lecture_number
        WHERE
        {
                ?lecture rdf:type focu:Lecture .
                ?lecture npg:number ?lecture_number .
                ?lecture npg:number 2 .
                ?lecture focup:belongs_to ?course .
                ?course focup:belongs_to "COMP" .
```

```
                ?course npg:number 6741 .
                ?topic rdf:type focu:Topic.
                ?topic schema:isPartOf ?lecture.
                ?topic foaf:name ?topic_name .

        }
```

Output:



Showing 1 to 2 of 2 entries — Search: — Show 50 entries

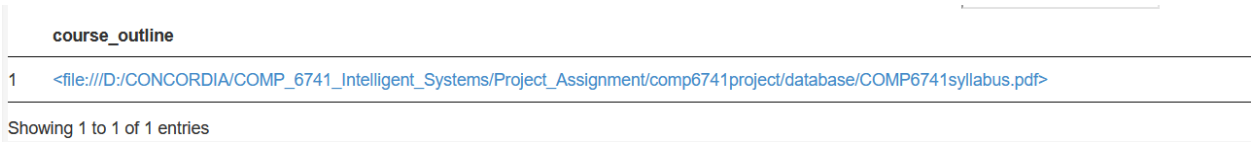
	course	topic_name	lecture_number
1	focu:COMP_6741/	"Knowledge Graph"	"2"^^xsd:integer
2	focu:COMP_6741/	"Semantic Web"	"2"^^xsd:integer

Showing 1 to 2 of 2 entries

8) What is the syllabus for COMP 6741?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
SELECT ?course_outline
        WHERE
        {
                ?course rdf:type focu:Course.
        ?course focup:belongs_to "COMP".
                ?course npg:number 6741 .
                ?course focup:has_outline ?course_outline.
        }
```

Output:



	course_outline
1	<file:///D:/CONCORDIA/COMP_6741_Intelligent_Systems/Project_Assignment/comp6741project/database/COMP6741syllabus.pdf>

Showing 1 to 1 of 1 entries

9) What departments are available under Gina Cody School of Engineering & Computer Science?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
SELECT DISTINCT ?course_subject
        WHERE
        {
                ?course rdf:type focu:Course.
                ?course focup:belongs_to ?course_subject .
                ?course schema:isPartOf "Gina Cody School of Engineering & Computer
Science".


        }
```

Output:

Showing 1 to 15 of 15 entries

	course_subject
1	"MECH"
2	"CIVI"
3	"ELEC"
4	"BLDG"
5	"COMP"

10) What graduate courses are offered by Gina Cody School of Engineering & Computer Science?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX npg: <http://ns.nature.com/terms/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>
PREFIX focup: <http://focu.io/schema/property#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
```

```
SELECT ?course_name
        WHERE
        {
                        ?course rdf:type focu:Course .
                        ?course focup:has_level "GRAD" .
                        ?course schema:isPartOf "Gina Cody School of Engineering & Computer
Science" .
                        ?course foaf:name ?course_name .
        }
```

Output:

Showing 1 to 50 of 331 entries

	course_name
1	"MECH BEHAV/POLYMER COMP MAT"
2	"DISCRETE EVENT SYSTEMS"
3	"STRUCTURAL SYSTEMS/BUILDING"
4	"NATURAL LANGUAGE ANALYSIS"
5	"TOPICS IN INDUSTRIAL ENGR"
6	"GRADUATE SEMINAR IN BCEE"
7	"NUMERICALLY CONTROLLED MACH."

Statistics query:

1. Total number of triples

   SELECT (COUNT(?s) AS ?triples) WHERE { ?s ?p ?o }

	triples
1	"80149"^^xsd:integer

Showing 1 to 1 of 1 entries

2. Total number of Course URIs

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX focu: <http://focu.io/schema#>

SELECT (COUNT(?s) AS ?triples)

WHERE { ?s rdf:type focu:Course . }
```

Showing 1 to 1 of 1 entries

	triples
1	"7244"^^xsd:integer

Showing 1 to 1 of 1 entries