Course Structure
Web 3 Science: Knowledge & Semantics
Fall 2018

Notice: This course is a graduate-level course, advanced undergraduate students can take the course in case of high motivation and having interest in cutting-edge areas and research

Scheduling: TR: 12:30 pm - 1:45 pm
Venue: MH 207
Office: Room #101B, and # 101C, Music and Theater Building

Assistant Professor: **Dr. Saeedeh Shekarpour**
E-mail: sshekarpour1@udayton.edu
Phone: (937)229 3925
Office Hours: MW: 3:30-4:30

Teacher Assistant: **Sunday Ngwobia**
E-mail: ngwobias1@udayton.edu
Office Hours: MW: 1 pm- 4 pm

Teacher Assistant: **Shubham Kokul**
E-mail: kokuls1@udayton.edu
Office Hours: MWF: 2 pm- 4 pm

Course Description

In this course, the following subjects will be presented:

- ➔ Technologies for developing Knowledge Graphs including:
  - ◆ Semantic Web Technologies: RDF and RDFS
  - ◆ Ontology Engineering
  - ◆ SPARQL Language
- ➔ The major concepts for interlinking knowledge graphs and measuring the quality of data
  - ◆ Linked Data
  - ◆ Open Data
  - ◆ Data Quality
- ➔ Techniques for Knowledge Graph Development and Analytics
  - ◆ Knowledge graph construction
  - ◆ Knowledge graph analytics

Text Resources

- ➔ This topic is a state-of-the-art topic. Thus, I do not rely on a typical textbook, we will explore lots of online resources including W3c recommendations, research papers …
- ➔ During the course, I will upload the link of the resources and our slides.

Grading Schema

Exams + Quizs ……	50%
Project ……	25%
Paper Presentation ……	25%
Assignments ……	10%
Class Activities ……	extra credit
Total ……	> 100%

Important Statements:

- ➔ **Academic Honesty:** You are allowed to collaborate and discuss with other students or search online, but you are not allowed to copy under any circumstances, any case of plagiarism will cause F for all the involving students.
- ➔ Your class activity and presence is a **must**, the more activity the more credit.
- ➔ Projects should be done individually.
- ➔ Paper presentations are a group activities. Each group has 2 members.
- ➔ Please bring your notebook to the Lab sessions.

Course Requirements:

- ➔ Prior and fresh knowledge in programming especially in Java.
- ➔ Please install a Java Editor e.g., Eclipse on your laptop

Course Structure

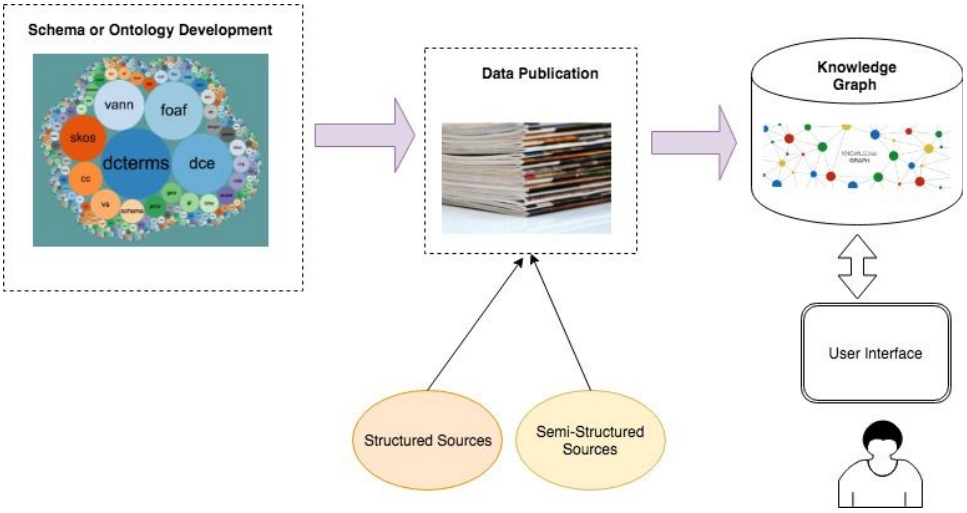
Month	Week	Topic	References
January 2019	Week 1	Introduction & Motivation	
	Week 2	RDF & RDF Schema (RDFS)	
	Week 3	RDF & RDF Schema (RDFS) Lab 1 Quiz 1	Import Jena Library into your Java IDE https://jena.apache.org/ Follow the tutorial to setup with the jar https://www.youtube.com/watch?v=YXjYPz6SHjA
February 2019	Week 4	Ontology Engineering Lab 2	Download Desktop version of Protege or make an account on the online Version Note: Protege only works with Java 8 Please setup environment variable before installing protege https://protege.stanford.edu/ Follow the tutorial to setup protege https://www.youtube.com/watch?v=eZyW7192wSU
	Week 5	**Project Delivery - Part 1**	Do the scheduling in advance
	Week 6		
	Week 7	Mid Term SPARQL Lab 3	
March 2019	Week 8		
	Week 8	**Spring Break**	
	Week 9	SPARQL SPARQL Update Lab 4 Lab 5	You can install an open source triple store: https://en.wikipedia.org/wiki/Comparison_of_triplestores Stardog seems easier: https://www.stardog.com/doc See macOS installation video tutorial link:https://www.youtube.com/watch?v=Kl9KHTyfT6Q
	Week 10		
	Week 12	Linked Data	

April 2019		Open Data Data Quality	
	Week 13	**Project Delivery - Part 2** Knowledge Graph Construction and Analytics	Do the scheduling in advance
	Week 14		
	Week 15	**Paper Presentation** **Q & A**	Do the scheduling in advance
May 2019	**Week 16**		

Scheduling: the scheduling will be done via this excel sheet, please email me in advance to reserve your slot, the strategy for booking is first-come-first-serve.

Course Project

Project Description. This course project is about developing a knowledge graph for a specific domain of your own choice. The project has to be done teamwork (two persons per team). The deliveries for your project are divided into two parts (i) **part-1: ontology design:** in this part you have to model an ontology for your domain and (ii) **part-2: knowledge graph development:** in this part you have to acquire data related to your domain, publish data according to the principles of Linked Open Data and your designed ontology. Each part has to be presented in the class, you have 10 minutes for the first time and 20 minutes for the second part. The structure of your presentations should follow the following structure:



➔ **Project Delivery - Part-1**

1) **Introduction (motivation and description of your domain):** Which domain did you choose to model? What are the motivations behind your choice?
2) **Review of the Existing Ontologies:** you have to make a review of the state-of-the-art ontologies related to your domain. What is missing from the state-of-the-art? What you can design and what you can incorporate?
3) **Modeling Ontology:** Presenting the ontology that you are designed for your domain with respect to the hierarchy of concepts/classes, relations between concepts/classes and datatype properties

➔ **Project Delivery - Part-2**

4) **Data Acquisition:** acquiring data with respect to your ontology

5) **Data Publication:** Converting the format of data to RDF with respect to your ontology
6) **Populating Knowledge Graph:** developing a knowledge graph containing both schema and instance levels and populated by data
7) **Querying Knowledge graph:** formulating several SPARQL queries which transforms natural language queries to SPARQL queries
8) **Implementation details:** which libraries did you use, how is the contribution of the team members in the implementation? Which parts of your ontology come from external open-source projects? Did you upload your code on GitHub?
9) **User Interface (Extra Credit):** developing a UI to navigate, explore and query knowledge graph.
10) **Discussion and conclusion:** what are your achievements? What are the potential improvements for future? What you did not have time to finish, but you think would be a useful addition to your project.
11) **The contribution of team members:** how was the team management strategies and what was challenging?

Project presentation (Phase 1): each team is given 15 minutes to present items (1), (2) and the item (3).

Project presentation (Phase 2): each team is given 25 minutes to present the whole of the project.

Paper Presentation

The aim of this part is to get familiar with the applications of Semantic Web and Linked Data Technologies. Your paper presentation is teamwork, it requires you to choose a topic (please consult with me if you have a challenge in choosing a topic) and then finds an outstanding paper (preferably recent papers after 2015) from a prestigious conference or journal.

The conferences and journals I prefer are listed as follows:

- ➔ The Web Conference (formerly Worldwide Web)
- ➔ International Semantic Web Conference (ISWC)
- ➔ European Semantic Web Conference (ESWC)
- ➔ Web Semantics Journal
- ➔ Semantic Web Journal
- ➔ Others WSDM, CIKM, and KDD

Relevant Topics

1. Entity Linking, Entity Disambiguation
2. Relation Linking
3. Data Quality of Knowledge Graph
4. Knowledge Graph creation, evolution, completion
5. Question Answering and Semantic Search
6. Ontology learning, adaptation, learning
7. Semantic Web/Web of Data and Health
8. Semantic Web and Internet of Things (IoT)
9. Semantic Web and Web 4.0 (Industry 4)
10. Semantic Web/Web of Data and Cognitive systems
11. Semantic Web/Web of Data and Social Media

Group Policies

1. each group should have 2 members
2. Time slot 20 minutes
3. Email me w.r.t. chosen paper or topic, I have to confirm your choice
4. It is not necessary to make slides for your presentation, you can highlight important parts of the paper (i.e., pdf version) and go through that during the presentation.