Teaching

Data Semantics

Spring 2020

Course Information

Course Date and Time: Wednesday 12:00-3:00PM
Location: UTA 1.210A
Instructor: Ying Ding
Office Hour: Wednesday 10-11:45AM, or by appointment (Email: ying.ding@ischool.utexas.edu)

Course Description

The current Web has experienced tremendous changes to connect data, people and knowledge. There are a couple of exciting efforts trying to bring the Web to its full potential. The Semantic Web is one of them which is heavily embedded in the Artificial Intelligence area with the long-term goal to enhance the human and machine interaction by representing data semantics, integrating data silos, and enabling intelligent search and discovery.

This course aims to provide the basic overview of the Semantic Web in particular, and data semantics in general, and how they can be applied to enhance data integration and knowledge inference. Ontology is the backbone of the Semantic Web. It models the semantics of data and represents them in markup languages proposed by the World Wide Web Consortium (W3C). W3C plays a significant role in directing major efforts at specifying, developing, and deploying standards for sharing information. Semantically enriched data paves the crucial way to facilitate the Web functionality and interoperability.

This course aims to provide the basic overview of what the Semantic Web is and how it can be applied. It contains three parts: Semantic Web language, RDF graph database (i.e., RDF triple store), and its applications. The fundamental part of the course is the Semantic Web languages. It starts from XML and goes further to RDF and OWL. The RDF graph database part introduces different APIs of Jena and its reasoners. The application part showcases current trends on semantic applications.

Prerequisites

Basic knowledge of HTML and XML is desired.

Course Objectives

This course aims to develop a critical appreciation of semantic technologies as they are currently being developed. At the end of this course, students should be able to

- sketch the overall architecture of the Semantic Web.
- identify the component technologies of the Semantic Web and explain their roles.
- illustrate the design principles of the Semantic Web by applying the technologies.
- understand certain limitations of the Semantic Web technologies, and be aware of the kinds of services it can and cannot deliver.

The course aims are achieved through:

- Lectures covers basic knowledge of the Semantic Web
- Projects applying semantic technologies to concrete problems of information delivery and use
- Assignments of practicing and utilizing key semantic technologies

Course Materials

This course is developped mainly based on materials from w3schools and W3C

Recommended books for this course:

- Allemang, D., & Hendler, J. (2011). Semantic Web for the working ontologist. 2nd Edition, Morgan&Kaufmann Publisher. [ISBN:978-0-12-385965-5]
- Heath, T., & Bizer, C. (2011). Linked Data: Evolving the Web into a Global Data Space. Morgan&Claypool Publisher. (here is the free version)
- Daconts, M.C., Orbst, L.J., & Smith, K.T.(2003). The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management. New York: Wiley. [ISBN: 0-471-43257-1]
- Antoniou, G., & van Harmelen, F. (2004). A Semantic Web Primer. Cambridge, MA: MIT Press. [ISBN: 0-262-01210-3]
- Powers, S. (2003). Practical RDF. Sebastopol, CA: O'Reilly.[ISBN: 0-596-00263-7]
  (One copy is reserved at IU Library Media & Reserve Service Department (http://www.libraries.iub.edu /index.php?pageId=307) under Prof. Ding's personal reservation)

Artificial Intelligence

- Russell, S., & Norvig, P. (2003). Artificial Intelligence: A Modern Approach. 2nd Edition. Pearson Education.

Software

- Oxygen XML editor
- Protege Ontology editor: available at protege website
- Jena-2.5.5, ARQ-2.2, and Tutorial.zip
- Instructions on installing Java and Jena on Windows or Mac

Others

- Semantic University
- Semantic Web Synthesis
- A online course on the usage of Linked Data

Calendar

Week	Lecture	Lab	Assignment
Week1	Introduction	XML (for more details, see XML Workshop), work on XML exercises on the last two slides of XML.ppt, submit answer to your canvas folder	
Week2	DTD, XML Schema	work on XML Schema Exercise.doc, submit answer to your canvas folder	Announce Assignment 1, Group Project (form a group:<=3)
Week3	RDF	Turtle, RDF examples, work on your RDF example, work on RDF Exercise.doc, submit answer to your canvas folder	
Week4	RDF Schema	Linked Open Data, submit your RDF example to your canvas folder, work on RDFS Exercise.doc, submit answer to your canvas folder	

Week5	OWL, Protege Lab	Work on OWL exercise.doc, submit answer to your canvas folder, go through the Protege Lab by yourself	Group Project (send title and team info to TA)
Week6	OWL, Protege Lab	go through the Protege Lab by yourself, work on Assignment 1	Announce Assignment 2
Week7	Jena RDF API	run RDF API tutorials by yourself, if you prefer to use Eclipse. Same tutorial on Apache Jena 3 (how to set up Jena 3 on Eclipse; Jena 3 RDF API), work on Assignment 2	Assignment 1 due, Project work
Week8	Jena Ontology API	Same tutorial on Apache Jena 3 Jena 3 Ontology API; run ontology tutorials by yourself	Assignment 2 due, Announce Assginment 3
Week9	Jena Reasoner	run reasoner tutorials by yourself, work on Assignment 3	Project Work
Week10	Sparql	work on Sparql Exercise.doc, submit answer to your canvas folder	Assignment 3 due, Announce Assignment 4
Week11	Sparql by Example	run these sparql examples by yourself, work on Assignement 4	Project Work
Week12	Jena Sparql API	Jena Sparql API for Mac, run Sparql exmaples by yourself, work on Assignment 4	Project Work
Week13	Jena Examples	Expert Search in PostgreSQL, run these two examples by yourself, workon Assignement 5	Assignment 4 due, Announce Assignment 6
Week14	Data Search	run data search tutorial by yourself, Data Search Tutorial, Dean Allemang tutorial	Assignment 5 due
Week15	, Project Presentation	Project Presentation	Assignment 6 due
Week16			Project Report due (before 5PM)

Homework

Project: You can select Engineering Project or Survey Project based on your own interest

Engineering Project: Building up a website using learned Semantic Web technologies or a semantic search system based on Jena

- 2-3 students form one group, or you can also work alone
- Select one interesting domain or problem, building up one ontology, using RDF to represent your data
- Setting up a running website which is able to display and query data
- Or creating a semantic search system by using Jena reasoning supports

Survey Project: Surveying the State of the Arts of the Semantic Web, Web2.0, or related areas

- Select one interesting area, identifying existing researches, comparing existing approaches, Summarizing the current status
- potential topics

Expected Result:

- Presentation or demo (individual): present your contribution in this project, if you have finished early and want to present it ealier than the deadline, you can send this request to the instructor.
- Project report (group): write the report together and state clearly the contribution from each team member. It should be 4-6 pages in ACM format, see one example.
- Engineering Projecct Example for a semantic website, a semantic search system, and more examples about previous projects
- Survey Project Example

Assignment: There will be six individual assignments on XML, RDF, and Jena. Please work on these assignments individually. Submit your assignments to your canvas folder

- Assignment 1 (10%): RDF XML OWL exercise, and your protege lab pizza.owl
- Assignment 2 (10%): Jena RDF API (java code on reading one RDF file into Jena and printing it out, submit codes and screenshots, for examples, please see Jena RDF API tutorials)
- Assignment 3 (10%): Jena Reasoner (java code on one example of using jena reasoner, submit codes and screenshots,for examples, please see Jena reasoner tutorials)
- Assignment 4 (10%): Jena Sparql queries (10 sparql queries on one RDF file, submit the RDF file, queries and screenshots, 2 of the sparql queries must use Jena API (like example of Bloggers.java))
- Assignment 5 (10%): Create your own example of Linked Open Data (>2 datasets, adding connections to other datasets (e.g., Dbpedia, Freebase) using owl:sameAs or rdfs:seeAlso). Use Heath and Bizer's book on Linked Data as the guidelines, see one example (Notes: to view this example, please save the owl file and view it using notepad or any text editor).
- Assignment 6 (20%): One Jena example (Java code on using reasoner and sparql queries, submit codes and screenshots, for examples, please see Jena example tutorials)

Resources

- Repository for Linked Open Data: http://datahub.io
- MIT Simile Project: http://simile-widgets.org/
- Nano Publication: http://nanopub.org/wordpress/
- Data-Driven Document: https://d3js.org/
- How to link resources to create LOD: sameas.org, identifiers.org, bridgeDB
- Semantic Web Mailinglist Archive: http://lists.w3.org/Archives/Public/semantic-web/, http://lists.w3.org/Archives/Public/public-lod/
- Semantic Web Synthesis: http://www.morganclaypool.com/toc/wbe.1/1/1
- Semantic Web at W3C: http://www.w3.org/standards/semanticweb/
- Semantic Web Conference: http://swsa.semanticweb.org/content/international-semantic-web-conference-iswc
- OpenCitations project: http://opencitations.net

Grading

The graded work in this course comprises

- Project (30%): Presentation or demo (20 points) + Project Report (10 points)
- Assignments (70%): six individual assignments. Note: All the exercises in the labs are not counted for your final grade.
- Class participation (required)
- Feedback and suggestions to improve this course (always welcome): extra credit

Your written, web-based, and oral work will be evaluated according to four criteria; it must:

- Be clearly written, marked up, and/or presented, and checked for spelling and grammar;
- Demonstrate a degree of insight into the concepts, issues, and trends in both the areas you investigate in the assignments and in the course content;
- Demonstrate a degree of originality in your reviews, analyses and projects; and
- Display familiarity with the appropriate literature.

To receive a passing grade in this course, you must turn in all of the assignments and the term project and complete all the presentations. You cannot pass this course without doing all of the assigned work (which includes the final presentation), however, turning in all of the work is not a guarantee that you will pass the course.

Borderline grades will be decided (up or down) on the basis of class contributions and participation throughout the semester.

A	4.0	Outstanding achievement. Student performance demonstrates full command of the course materials and evinces a high level of originality and/or creativity that far surpasses course expectations.
A-	3.7	Excellent achievement. Student performance demonstrates thorough knowledge of the course materials and exceeds course expectations by completing all requirements in a superior manner.
B+	3.3	Very good work. Student performance demonstrates above-average comprehension of the course materials and exceeds course expectations on all tasks as defined in the course syllabus.
B	3.0	Student performance meets designated course expectations and demonstrates understanding of the course materials at an acceptable level.
B-	2.7	Marginal work. Student performance demonstrates incomplete understanding of course materials.
C+	2.3	Unsatisfactory work. Student performance demonstrates incomplete and inadequate understanding of course materials.
C	2.0	
C-	1.7	Unacceptable work. Coursework performed at this level will not count toward the MLS or MIS degree. For the course to count toward the degree, the student must repeat the course with a passing grade.
D+	1.3	
D	1.0	
D-	0.7	
F	0.0	Failing. Student may continue in program only with permission of the Dean.