

Professor Dennis McLeod

Time and Location

Monday and Wednesday

2:00 - 3:50 p.m.

Location: TBD

Brief Course Summary

This course focuses on key issues and the state-of-the-art of information sharing and federation/interoperation of data intensive systems. It stresses the use of data semantics, and explores ontologies and the semantic web. Specific foci of the course include: semantic database models, ontologies; ontology structuring, acquisition, integration, learning, and sharing; ontology applications; the semantic web; semantics-based data mining; relationships with social media; trust and information privacy.

Learning Objectives

- learn the formal underpinnings of data semantics and sharing
- understand semantics for both structured and unstructured data
- learn about and work with ontologies and the semantic web
- understand the concept of database federation and interoperation
- gain experience analyzing technical papers and presenting their essence
- learn to make a strong discussion-oriented presentation
- gain experience working in a group project on a challenging problem in the area

Course Requirements

- Extensive readings
- Attendance, class participation
- Presentation(s) (powerpoint or similar)
- Individual short paper
- Group project
- Periodic (e.g., every two weeks) short in-class "quizzes"; partially "flipped" class

Presentation

Each presentation should cover the following:

- What is/are the research issue(s)/problem(s) addressed in the paper?
- Summarize the technical approach
- Describe how the results in this paper can be used, evaluate the results
- Pose several key issues/questions for class discussion

Important notes on presentations:

- Presentations should consist of at most 25 slides maximum, with a reasonable font size. Please do not put too much text on a slide.
- Please email your complete presentation to the Professor, TA, Course Producer and Grader (all course "staff") by 3:00pm on the Friday prior to your presentation day (for Monday presenters), or by the prior Monday at 3:00pm (for Wednesday presenters). Depending on the quality of your presentation, you may receive comments and suggestions, or be asked to make changes before you give the presentation.
- You must include discussion questions/issues in your presentation.
- After the presentation, please email the final version of your slides to the Professor, TA, Course Producer and Grader.
- Presentation [sample](#)

Grading breakdown of the presentation:

40% for content

30% for delivery

30% for discussion

Group Project

Group projects will involve your participation in a semester-long effort to study, design, and analyze a particular aspect of ontology and semantic web work or application domain, and construct an experimental prototype (possibly an ontology, meta-ontology, etc.). (Details are provided in the course project specification, under Projects on the website for this course).

Project Timeline:

Group Formation: September 14

Custom Project Topic Proposals: September 28

Topic Selection: October 5

Progress Report: October 19-24 (in class)

Final In-Class Presentations: November 28-30 (in class)

Final Report and Material Submission: December 2

Grading breakdown of the group project:

20% for a mid-semester brief progress report

40% for a final in-class presentation

40% for a final report plus material developed in the project

Overall Grading

Grading for the course will be determined approximately as follows:

40% for in-class presentation(s), leading discussion

10% for class participation and attendance (attend at least 80% of the classes to receive full credit)

30% for the group project

20% for the “short” (5-10 page) paper (requires a lot of reading), and every two week unannounced quizzes

Academic Integrity Policy

Academic Integrity

All presentations, papers, projects, and submitted material must be created and written independently, or you will be penalized for cheating. The USC [Student Conduct Code](#) prohibits [plagiarism](#). All USC students are responsible for reading and following the Student Conduct Code, which also appears in the latest edition of "SCampus".

In this course we encourage students to study together. This includes discussion, and of course collaboration on the course group projects. However, all work submitted for the class is to be done by you or as an identified group effort (for group projects).

Students who violate University standards of academic integrity are subject to disciplinary sanctions, including failure in the course and suspension from the University. Since dishonesty in any form harms the individual, other students, and the University, policies on academic integrity will be strictly enforced. We expect you to familiarize yourself with the Academic Integrity guidelines found in the current SCampus. Violations of the Student Conduct Code will be filed with the Office of Student Conduct, and [appropriate sanctions](#) will be given.

Statement for Students with Disabilities

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to my TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30AM-5PM, Monday through Friday. The phone number for DSP is (213) 740-0776."

Related Web Sites

- [Semantic Web](#)
- [OWL Web Ontology Language](#)
- [OWL2 Web Ontology Language](#)
- [Protege Ontology Manager](#)
- [OWL Reasoning Examples](#)
- [Ontology Alignment Initiative](#)
- [Gene Ontology](#)
- [The Wordnet Lexical Ontology](#)
- [USC Semantic Information Research Laboratory](#)
- [Stanford University Database Group](#)
- [University of Wisconsin Database Group](#)
- [University of Maryland Database Group](#)
- [Tips for giving a good scientific talk](#)
- [Efficient Reading of Papers in Science and Technology](#)

Week/Date	Topics
<u>Week 1</u>	
Aug 22 (Session 1)	Course Introduction: Dennis McLeod.
Aug 24 (Session 2)	Introduction: Data Semantics, Interoperation and Federation <ul style="list-style-type: none"> Joachim Hammer, Dennis McLeod. An Approach to Resolving Semantic Heterogeneity in a Federation of Autonomous, Heterogeneous Database Systems. <i>International Journal of Intelligent and Cooperative Information Systems</i>, 2(1): 51-83 (1993). Yolanda Gil. Interactive Knowledge Capture in the New Millennium: How the Semantic Web Changed Everything. <i>Knowledge Engineering Review</i>, 2011. TBD
<u>Week 2</u>	
Aug 29 (Session 1)	Ontologies, Semantic Web <ul style="list-style-type: none"> Stefan Decker, Sergey Melnik, Frank van Harmelen, Dieter Fensel, Michel C. A. Klein, Jeen Broekstra, Michael Erdmann, Ian Horrocks. The Semantic Web: the roles of XML and RDF. <i>IEEE Internet Computing</i> 4(5): 63-74 (2000). Peter F. Patel-Schneider, Dieter Fensel. Layering the Semantic Web: Problems and Directions. <i>Proceedings of the International Semantic Web Conference (ISWC 2002)</i>. Li Ding, Tim Finin. Characterizing the Semantic Web on the Web. <i>Proceedings of the International Semantic Web Conference (ISWC 2006)</i>.
Aug 31 (Session 2)	Ontologies, Semantic Web <ul style="list-style-type: none"> Stefan Decker, Sergey Melnik, Frank van Harmelen, Dieter Fensel, Michel C. A. Klein, Jeen Broekstra, Michael Erdmann, Ian Horrocks. The Semantic Web: the roles of XML and RDF. <i>IEEE Internet Computing</i> 4(5): 63-74 (2000). Peter F. Patel-Schneider, Dieter Fensel. Layering the Semantic Web: Problems and Directions. <i>Proceedings of the International Semantic Web Conference (ISWC 2002)</i>.

	<ul style="list-style-type: none"> Li Ding, Tim Finin. Characterizing the Semantic Web on the Web. <i>Proceedings of the International Semantic Web Conference (ISWC 2006)</i>.
<u>Week 3</u>	
Sep 5	Labor Day
Sep 7 (Session 2)	<p>Information Retrieval and Web Search with Ontologies</p> <ul style="list-style-type: none"> Pablo Castells, Miriam Fernandez, David Vallet. An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval. <i>IEEE Transactions on Knowledge and Data Engineering</i>. 19(2): 261-272 (2007). Seokkyung Chung, Jongeun Jun, Dennis McLeod. WebSim: A Pathway to Unveiling Term Relationships using a Web Search Technology. <i>Proceedings of the VLDB Workshop on Ontologies-based Techniques for Databases and Information Systems (2006)</i>. Vasco Calais Pedro, Eric Nyberg, Jaime Carbonell. Federated Ontology Search. <i>Proceedings of the International Workshop of Semantic Information Integration on Knowledge Discovery (SIK 2006)</i>. Latifur Khan, Dennis McLeod, Eduard Hovy. Retrieval Effectiveness of An Ontology-based Model for Information Selection <i>Proceedings of the International Conference on Very Large Data Bases (VLDB 2004)</i>.
	<p>Short paper Assignment out</p> <ul style="list-style-type: none"> This short paper assignment involves analyzing a research paper on the topic of ontology concept learning for images. (the paper will be posted online at class website). <p>Your task is to read and analyze this paper, writing a ~3-5 page short paper (figures or diagrams are welcome). You should address the issues below. You may address related and other issues as well (you may be creative here). You do not need to cover all of the below, and you can organize your short paper in any manner that works well. - What are the principal contributions of this paper? - How are concepts and relationships acquired/learned? - How does the approach to concept learning in this paper relate to the approaches to concept and relationship acquisition and learning covered in the papers we have read earlier this semester? - In what sense is the work in this paper specific to image semantics? Can it be used for other types of data (e.g., different modalities such as audio, video)? - Do you think the approach described in the paper is scalable to massive data (e.g., billions or even trillions of images)? - In what sense does the work in the paper utilize crowdsourcing?</p>

<u>Week 4</u>	
Sep 12 (Session 1)	Information Retrieval and Web Search with Ontologies <ul style="list-style-type: none"> • Pablo Castells, Miriam Fernandez, David Vallet. An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval. <i>IEEE Transactions on Knowledge and Data Engineering</i>. 19(2): 261-272 (2007). • Seokkyung Chung, Jongeun Jun, Dennis McLeod. WebSim: A Pathway to Unveiling Term Relationships using a Web Search Technology. <i>Proceedings of the VLDB Workshop on Ontologies-based Techniques for Databases and Information Systems</i> (2006). • Vasco Calais Pedro, Eric Nyberg, Jaime Carbonell. Federated Ontology Search. <i>Proceedings of the International Workshop of Semantic Information Integration on Knowledge Discovery (SIIK 2006)</i>. • Latifur Khan, Dennis McLeod, Eduard Hovy. Retrieval Effectiveness of An Ontology-based Model for Information Selection <i>Proceedings of the International Conference on Very Large Data Bases (VLDB 2004)</i>.
Sep 14 (Session 2)	Inter-Relating Ontologies <ul style="list-style-type: none"> • M. Andrea Rodriguez, Max J. Egenhofer. Determining Semantic Similarity among Entity Classes from Different Ontologies. <i>IEEE Transactions on Knowledge and Data Engineering</i>. 15(2): 442-456 (2003). • Muhammad Abulaish Lipika Dey. Interoperability among Distributed Overlapping Ontologies - A Fuzzy Ontology Framework. <i>Proceedings of the International Conference on Web Intelligence (WI 2006)</i>. • Joseph Hassell, Boanerges Aleman-Meza, Ismailcem Budak Arpinar. Ontology-Driven Automatic Entity Disambiguation in Unstructured Text. <i>Proceedings of the International Semantic Web Conference (ISWC 2006)</i>.
	Group Formation – Due

<u>Week 5</u>	
Sep 19 (Session 1)	Inter-Relating Ontologies <ul style="list-style-type: none"> • M. Andrea Rodriguez, Max J. Egenhofer. Determining Semantic Similarity among Entity Classes from Different Ontologies. <i>IEEE Transactions on Knowledge and Data Engineering</i>. 15(2): 442-456 (2003). • Muhammad Abulaish Lipika Dey. Interoperability among Distributed Overlapping Ontologies - A Fuzzy Ontology Framework. <i>Proceedings of the International Conference on Web Intelligence (WI 2006)</i>. • Joseph Hassell, Boanerges Aleman-Meza, Ismailcem Budak Arpinar. Ontology-Driven Automatic Entity Disambiguation in Unstructured Text. <i>Proceedings of the International Semantic Web Conference (ISWC 2006)</i>.
Sep 21 (Session 2)	Resolution of Semantic Heterogeneity <ul style="list-style-type: none"> • Goksel Aslan, Dennis McLeod. Semantic Heterogeneity Resolution in Federated Databases by Metadata Implantation and Stepwise Evolution. <i>VLDB Journal</i>. 8(2): 120-132 (1999). • Sangsoo Sung, Dennis McLeod. Ontology-Driven Semantic Matches between Database Schemas. <i>Proceedings of the International Workshop on Database Interoperability (ICDEW 2006)</i>. • TBD
<u>Week 6</u>	
Sep 26 (Session 1)	Resolution of Semantic Heterogeneity <ul style="list-style-type: none"> • Goksel Aslan, Dennis McLeod. Semantic Heterogeneity Resolution in Federated Databases by Metadata Implantation and Stepwise Evolution. <i>VLDB Journal</i>. 8(2): 120-132 (1999). • Sangsoo Sung, Dennis McLeod. Ontology-Driven Semantic Matches between Database Schemas. <i>Proceedings of the International Workshop on Database Interoperability (ICDEW 2006)</i>. • TBD

Sep 28 (Session 2)	Ontology Creation and Enhancement <ul style="list-style-type: none"> Sudha Ram, Jinsoo Park. Semantic Conflict Resolution Ontology (SCROL): An Ontology for Detecting and Resolving Data and Schema-Level Semantic Conflicts. <i>IEEE Transactions on Knowledge and Data Engineering</i>. 16(2): 189-202 (2004). Jaime Reinoso, Adrian Silvescu, Doina Caragea, Jyotishman Pathak, Vasant Honavar. Information Extraction and Integration from Heterogeneous, Distributed, Autonomous Information Sources : A Federated Ontology-Driven Query-Centric Approach. <i>Proceedings of the International Conference on Information Reuse and Integration (IRI 2003)</i>. Latifur Khan, Dennis McLeod, Eduard Hovy. A Framework for Effective Annotation of Information from Closed Captions Using Ontologies. <i>Journal of Intelligent Information Systems</i>. 25(2): 181-205 (2005).
	Custom Project Topic Proposals – Due
<u>Week 7</u>	
Oct 3 (Session 1)	Ontology Creation and Enhancement <ul style="list-style-type: none"> Sudha Ram, Jinsoo Park. Semantic Conflict Resolution Ontology (SCROL): An Ontology for Detecting and Resolving Data and Schema-Level Semantic Conflicts. <i>IEEE Transactions on Knowledge and Data Engineering</i>. 16(2): 189-202 (2004). Jaime Reinoso, Adrian Silvescu, Doina Caragea, Jyotishman Pathak, Vasant Honavar. Information Extraction and Integration from Heterogeneous, Distributed, Autonomous Information Sources : A Federated Ontology-Driven Query-Centric Approach. <i>Proceedings of the International Conference on Information Reuse and Integration (IRI 2003)</i>. Latifur Khan, Dennis McLeod, Eduard Hovy. A Framework for Effective Annotation of Information from Closed Captions Using Ontologies. <i>Journal of Intelligent Information Systems</i>. 25(2): 181-205 (2005).

Oct 5 (Session 2)	Ontology Creation and Enhancement (cont'd) <ul style="list-style-type: none"> • Fei Wu, Daniel S. Weld. Automatically Refining the Wikipedia Infobox Ontology. <i>Proceeding of the International World Wide Web Conference (WWW 2008)</i> • Lina Zhou. Ontology Learning: State of the Art and Open Issues. <i>Journal of Information Technology Management</i>: 8:241-252 (2007). • Sangsoo Sung, Seokkyung Chung, Dennis McLeod. Efficient Concept Clustering for Ontology Learning using an Event Life Cycle on the Web. <i>Proceedings of the ACM Symposium on Applied Computing (SAC 2008)</i>.
	Topic Selection – Due
<u>Week 8</u>	
Oct 10 (Session 1)	Ontology Creation and Enhancement (cont'd) <ul style="list-style-type: none"> • Fei Wu, Daniel S. Weld. Automatically Refining the Wikipedia Infobox Ontology. <i>Proceeding of the International World Wide Web Conference (WWW 2008)</i> • Lina Zhou. Ontology Learning: State of the Art and Open Issues. <i>Journal of Information Technology Management</i>: 8:241-252 (2007). • Sangsoo Sung, Seokkyung Chung, Dennis McLeod. Efficient Concept Clustering for Ontology Learning using an Event Life Cycle on the Web. <i>Proceedings of the ACM Symposium on Applied Computing (SAC 2008)</i>.
Oct 12 (Session 2)	OWL and RDF(S) <ul style="list-style-type: none"> • Aditya Kalyanpur, Jennifer Golbeck, Jay Banerjee, James Hendler. OWL : Capturing Semantic Information using a Standardized Web Ontology Language. <i>Multilingual Computing & Technology Magazine</i> 15(7) (2004). • Peter F. Patel-Schneider, Ian Horrocks. Position paper: a comparison of two modelling paradigms in the Semantic Web. <i>Proceeding of the International World Wide Web Conference (WWW 2006)</i>. • TBD

<u>Week 9</u>	
Oct 17 (Session 1)	OWL and RDF(S) <ul style="list-style-type: none"> Aditya Kalyanpur, Jennifer Golbeck, Jay Banerjee, James Hendler. <u>OWL: Capturing Semantic Information using a Standardized Web Ontology Language</u>. <i>Multilingual Computing & Technology Magazine</i> 15(7) (2004). Peter F. Patel-Schneider, Ian Horrocks. <u>Position paper: a comparison of two modelling paradigms in the Semantic Web</u>. <i>Proceeding of the International World Wide Web Conference (WWW 2006)</i>. TBD
Oct 19 (Session 2)	Group Project Progress Reports
<u>Week 10</u>	
Oct 24 (Session 1)	Group Project Progress Reports
Oct 26 (Session 2)	Semantics in Data Mining and Knowledge Discovery <ul style="list-style-type: none"> Danushka T. Bollegala, Yutaka Matsuo, Mitsuru Ishizuka. <u>Measuring the Similarity between Implicit Semantic Relations from the Web</u>. <i>Proceedings of the International World Wide Web Conference (WWW 2009)</i>. Benjamin Markines, Ciro Cattuto, Filippo Menczer, Dominik Benz, Andreas Hotho, Gerd Stumme. <u>Evaluating Similarity Measures for Emergent Semantics of Social Tagging</u>. <i>Proceedings of the International World Wide Web Conference (WWW 2009)</i>. Roberto Mirizzi, Azzurra Ragone, Tommaso Di Noia, Eugenio Di Sciascio. <u>Semantic Tags Generation and Retrieval for Online Advertising</u>. <i>Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM 2010)</i>.

<u>Week 11</u>	
Oct 31 (Session 1)	Semantics in Data Mining and Knowledge Discovery <ul style="list-style-type: none"> • Danushka T. Bollegala, Yutaka Matsuo, Mitsuru Ishizuka. Measuring the Similarity between Implicit Semantic Relations from the Web. <i>Proceedings of the International World Wide Web Conference (WWW 2009)</i>. • Benjamin Markines, Ciro Cattuto, Filippo Menczer, Dominik Benz, Andreas Hotho, Gerd Stumme. Evaluating Similarity Measures for Emergent Semantics of Social Tagging. <i>Proceedings of the International World Wide Web Conference (WWW 2009)</i>. • Roberto Mirizzi, Azzurra Ragone, Tommaso Di Noia, Eugenio Di Sciascio. Semantic Tags Generation and Retrieval for Online Advertising. <i>Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM 2010)</i>.
Nov 2 (Session 2)	Semantics in Social Media <ul style="list-style-type: none"> • Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, Krishna P. Gummadi. Measuring User Influence in Twitter: The Million Follower Fallacy. <i>Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM 2010)</i>. • Yolanda Gil, Angela Knight, Kevin Zhang, Larry Zhang, Varun Ratnakar, Ricky Sethi. The Democratization of Semantic Properties: An Analysis of Semantic Wikis. <i>Proceedings of the International Conference on Semantic Computing (ICSC 2013)</i>. • Kristina Lerman, Tad Hogg. Using a Model of Social Dynamics to Predict Popularity of News. <i>Proceedings of the International World Wide Web Conference (WWW 2010)</i>.

<u>Week 12</u>	
Nov 7 (Session 1)	Semantics in Social Media <ul style="list-style-type: none"> • Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, Krishna P. Gummadi. Measuring User Influence in Twitter: The Million Follower Fallacy <i>Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM 2010)</i>. • Yolanda Gil, Angela Knight, Kevin Zhang, Larry Zhang, Varun Ratnakar, Ricky Sethi. The Democratization of Semantic Properties: An Analysis of Semantic Wikis. <i>Proceedings of the International Conference on Semantic Computing (ICSC 2013)</i>. • Kristina Lerman, Tad Hogg. Using a Model of Social Dynamics to Predict Popularity of News. <i>Proceedings of the International World Wide Web Conference (WWW 2010)</i>.
Nov 9 (Session 2)	TBD
	Short Paper Assignment - Due
<u>Week 13</u>	
Nov 14 (Session 1)	TBD
Nov 16 (Session 2)	Trust and Context in Interoperability and Sharing <ul style="list-style-type: none"> • Jennifer Golbeck, Bijan Parsia, James A. Hendler. Trust Networks on the Semantic Web. <i>Proceedings of the International Workshop on Cooperative Intelligent Agents (CIA 2003)</i>. • Zoltan Gyongyi, Hector Garcia-Molina, Jan Pedersen. Combating Web Spam with TrustRank. <i>Proceedings of the International Conference on Very Large Data Bases (VLDB 2004)</i>. • Vesile Evrim, Dennis McLeod. Context-Based Information Analysis for the Web Environment <i>Knowledge and Information Systems (KAIS 2013)</i>.

<u>Week 14</u>	
Nov 21 (Session 1)	Trust and Context in Interoperability and Sharing <ul style="list-style-type: none"> Jennifer Golbeck, Bijan Parsia, James A. Hendler. Trust Networks on the Semantic Web. <i>Proceedings of the International Workshop on Cooperative Intelligent Agents (CIA 2003)</i>. Zoltan Gyongyi, Hector Garcia-Molina, Jan Pedersen. Combating Web Spam with TrustRank. <i>Proceedings of the International Conference on Very Large Data Bases (VLDB 2004)</i>. Vesile Evrim, Dennis McLeod. Context-Based Information Analysis for the Web Environment <i>Knowledge and Information Systems (KAIS 2013)</i>.
Nov 23	Thanksgiving
<u>Week 15</u>	
Nov 28 (Session 1)	Project Presentations/Discussion
Nov 30 (Session 2)	Project Presentations/Discussion
	Final Report - DUE! (Due based on Final exam schedule) <p>Final Report - The project final written report should be done in the form of a technical conference paper. You are welcome to use conference paper templates (LaTeX or Word format) for ACM / LNCS / AAAI / IEEE /etc. if you find them helpful, but this is not a requirement. This report, along with the presentation slides from your in-class presentation, as well as all materials developed for the project (e.g. ontologies, application source code, documentation, etc.)</p> <p>In the report itself – please include the following:</p> <ul style="list-style-type: none"> Problem Statement Scope of the Project Steps Involved Design and Implementation Properties Tools used Sample SPARQL queries