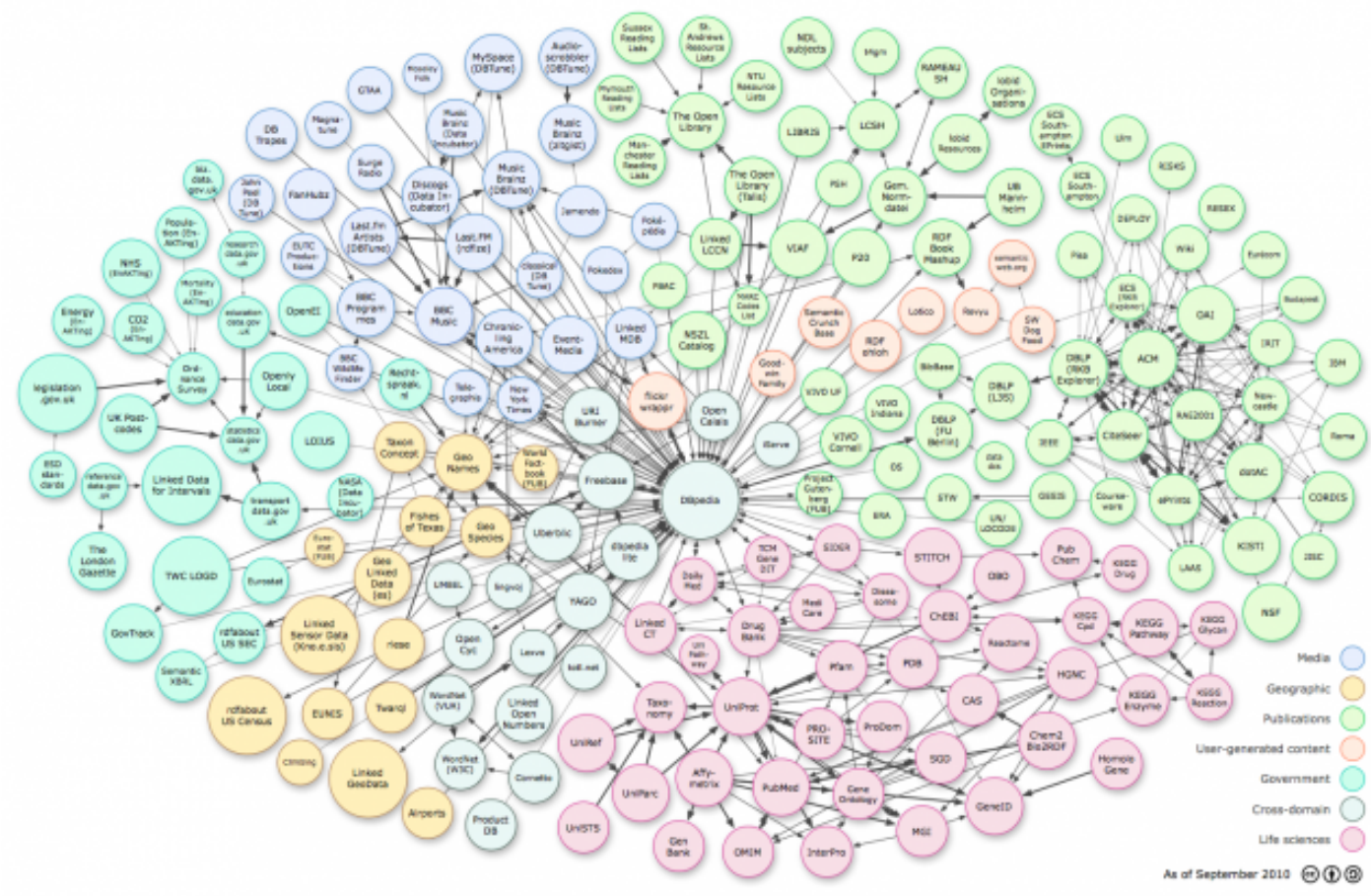
Semantic Computing Course

- Semantic Computing
- Teaching
- Text Mining

The *Semantic Computing* course (SOEN 6211, formerly SOEN 691B) is offered at Concordia University, providing graduate students with a unique opportunity to study research and development of novel semantic software systems. The course is taught by Prof. René Witte and supported by team members from the Semantic Software Lab. Students from other universities in Québec can register for this course through CREPUQ.

1. Contents

This course provide an introduction to selected topics from Semantic Computing, including:



Open Linked Data Cloud

Introduction to Semantic Computing

Introduction to Semantic Computing Paradigms and Applications; Processes, Tools, and Standards for Deriving Intelligence from (User-Generated) Content

Text Mining Systems

Foundations of Language Technology and Language Engineering; Text Mining Systems; Industry Standards; Design and Deployment of Text Mining Solutions; Information Extraction; Case Study: The General Architecture for Text Engineering (GATE)

Tagging and Tag Analysis

Social Tagging vs. Machine-Generated Tags; Folksonomies; Dynamic Navigation and Tag Clouds

Recommender Systems and Collaborative Filtering.

Theory and Implementation of Recommender Engines.

Analyzing the Blogosphere

Blog Tracking; Opinion Mining; Analysis of Microblogs (e.g., Twitter); Standards: RSS, ATOM; Services: Technorati, Bloglines

RDF and Linked Data

RDF; Triplestores; SPARQL; Microformats; GRDDL; Linked Open Data; Case Studies: DBpedia, Freebase, Wikidata, FOAF, GeoNames

Semantic Information Repositories and Information Retrieval

Digital Object Management; Semantic Digital Libraries; Semantic Metadata; Dublin Core; Information Retrieval; Lucene; Solr; FedoraCommons; GATE Mimir

Ontologies and the Semantic Web

Web Ontology Language (OWL); Ontology Editors; Ontology Learning and Ontology Population from Text

Semantic Desktops and Semantic Wikis

Integration of Semantic Technologies into Desktop Environments and Information Systems; Case Studies: Semantic MediaWiki, Semantic Assistants.

Scalable Semantic Systems Engineering

Clustering and Cloud Computing Paradigms; Apache Hadoop and MapReduce; Scalable Web Crawling using Nutch.

The course places a strong focus on the practical aspects of building semantic solutions. Students will have to complete a project on a selected course topic and present their work in class. Additionally, every student (including auditing students) will have to present a research paper during our weekly seminar.

2. Prerequisites

There are no formal prerequisites for this course. However, the course primarily targets students in a research option (PhD, MSc

thesis) that have existing background knowledge in topics such as machine learning, natural language processing, intelligent systems, information systems, or the semantic web, for example through one of these courses offered at Concordia:

- COMP 6321: Machine Learning
- COMP 6531: Foundations of the Semantic Web
- COMP 6751: Natural Language Analysis
- COMP 6781: Statistical Natural Language Processing
- COMP 6721: Introduction to Arti?cial Intelligence
- COMP 6741: Intelligent Systems

Additionally, students taking this course are expected to have good knowledge of an object-oriented programming language (like Java or C++), a scripting language (such as Python or Groovy), as well as general software engineering and information system knowledge. The primary development environment for this course is Linux.

3. Relationship to other courses

A frequently asked question is how this course compares to *COMP 6531: Foundations of the Semantic Web*, taught by Prof. Haarslev. The two courses are largely complementary and can be taken in parallel; however, it is recommended to first complete SOEN 691B, then COMP 6531: The SOEN 691B course focuses on the technical aspects of several semantic computing techniques, in particular text mining and linked data, whereas COMP 6531 focuses on the formal concepts and design of OWL ontologies, including theoretical foundations in description logics and automated reasoning. Together with the NLP courses offered through the CLaC lab and related courses in intelligent systems and machine learning, Concordia students have an opportunity to develop comprehensive skills in *semantic sciences* that are unique across Canada.

4. Schedule

Due to Prof. Witte's sabbatical, the course will not be offered in the 2014/2015 academic year.