INFOST 716 – Thesaurus Construction

DRAFT
SYLLABUS
for
Summer
2017

Instructor: Richard Smiraglia
E-mail: smiragli@uwm.edu
Office Hours: TBA

Location: NWQB 2569
Phone: 414-229-1750
Fax: 414-229-6699

Meeting Times & Location: online

CATALOG DESCRIPTION:

Overview of thesaurus construction, vocabulary control and ontology. Specific techniques for design and construction of thesauri, including domain analysis, vocabulary extraction, concept clustering, ordering, and identifying structural relations, and thesaurus maintenance. 3 credits.

PREREQUISITE:

INFOST 511 Organization of Information, or its equivalent.

GENERAL DESCRIPTION:

Students will become conversant with the principles and standards for thesaurus construction, including but not limited to domain analysis, vocabulary control, and ontology. Students will gain understanding of the broader theoretical context in which thesauri function. We will begin with techniques for analyzing the vocabulary of a domain, then we will study the use of various existing thesauri. Students will build a thesaurus in a domain of their choice. The relationship between domain vocabulary and ontology will be explored.

OBJECTIVES:

Upon completion of this course students will be able to demonstrate the following SOIS Program Outcomes:

• Analyze a domain for ontological content, extract terms and apply vocabulary control (Instructor meeting; Assignments; Discussion).
• Apply international standards (ISO25964 Thesaurus Schemas) to vocabulary of the designated domain (Assignments).
• Develop a functional thesaurus for the designated domain (Assignments).
• Demonstrate familiarity with principles knowledge representation for knowledge organization (Class presentation; Final exam).
• Gain the ability to articulate aspects of ontology, taxonomy, and vocabulary control (Final exam).

ALA CORE COMPETENCIES COVERED:

1A. Foundations of the profession. The ethics, values and foundational principles of information organization.
2A. Information resources. The ontological content of documents.
2C. The management of concepts for vocabulary control.
3A-C. Organization of recorded knowledge. Principles, descriptive skills, and systems of knowledge organization, especially knowledge representation, vocabulary control and thesaurus construction.
6A. Research. Domain analysis as a research method.

DIVERSITY ISSUES:

Diversity is embraced in the ethical treatment of vocabulary for all potential users, and in the use of knowledge organization systems for resource discovery. In particular, treatment of names of concepts in differing domains is addressed.

METHOD:

Lectures will be used to introduce course units. Student work will be presented in seminar. Discussion will be used to extend understanding of the readings. Small groups will be used to learn specific tasks.

REQUIRED READINGS (Both are provided in class):


Other Required Readings:


**ASSIGNMENTS**

The course is developed around the stages of construction of a domain-specific thesaurus. Assignments are:

- Assignment 1. Identify a domain; provide a description of the domain, and a list of essential source documents that can provide domain-specific vocabulary.
- Assignment 2. Extract essential vocabulary from the domain-specific sources identified in assignment 1. Decisions about indexing depth will be critical at this phase.
- Assignment 3. Coordinate the vocabulary. Cluster terms, and create a diagram of hierarchies and facets within the clusters. Identify homonyms and synonyms.
- Assignment 4. Structure the thesaurus and present it in alphabetical order.
- Assignment 5. Re-structure the conceptual nodes as an ontology. Comment.
- Assignment 6. Present a narrative justification of the domain, its vocabulary, your choices of terms, your decisions about clusters and facets.

**COURSE PARTICIPATION**

Course participation will be based on your regular substantive contributions to class discussions and small group tasks.
### COURSE SCHEDULE:

<table>
<thead>
<tr>
<th>On this date</th>
<th>We will take up these topics</th>
<th>You should read these things</th>
<th>And this will be due</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 30, 2017</td>
<td>Module 1: Introduction to Knowledge Representation&lt;br&gt;Introduction to indexing and subject analysis&lt;br&gt;Depth, Analysis, Summarization&lt;br&gt;Module 2: Vocabulary control principles&lt;br&gt;Analysis of major thesauri</td>
<td>Wilson, chap. 5&lt;br&gt;Brown, frames 37-130&lt;br&gt;Mai Beghtol, &quot;Whole, kind, parts&quot;&lt;br&gt;Littletree and Metoyer&lt;br&gt;Getty vocabularies</td>
<td>Assignment 1</td>
</tr>
<tr>
<td>June 6, 2017</td>
<td>Module 3: Thesaurus construction&lt;br&gt;Module 4: Domain analysis</td>
<td>Roe &amp; Thomas, chap. 1-2 (Roe, Aitchison &amp; Dextre Clarke)&lt;br&gt;Aitchison et al., pp. 17-48&lt;br&gt;Lancaster, chap. 4, 6-10&lt;br&gt;Smiraglia and Lee, chap. 1 and 6&lt;br&gt;Aitchison et al., pp. 49-83&lt;br&gt;Soergel, chap. 16</td>
<td>Assignment 2</td>
</tr>
<tr>
<td>June 13, 2017</td>
<td>Module 5: Vocabulary extraction&lt;br&gt;Module 6: Concept clustering</td>
<td>Roe &amp; Thomas, chap. 3-5 (Thomas, Shearer, Nielsen)&lt;br&gt;Soergel, chap. 12&lt;br&gt;Soergel, chap. 14</td>
<td>Assignment 3</td>
</tr>
<tr>
<td>June 20, 2017</td>
<td>Module 7: Identifying thesaural relationships&lt;br&gt;Module 8: Constructing the controlled vocabulary</td>
<td>Dextre-Clarke</td>
<td>Assignment 4</td>
</tr>
<tr>
<td>June 27, 2017</td>
<td>Module 9: From thesaurus to ontology&lt;br&gt;Module 10: Knowledge representation: the broader context</td>
<td>Smiraglia and Lee, chap. 2-5&lt;br&gt;Shiri, chap. 3&lt;br&gt;Willis and Losee&lt;br&gt;Smiraglia and Lee, chap. 7-9&lt;br&gt;Shiri, chap. 7</td>
<td>Assignment 5&lt;br&gt;Final thesauri (Assignment 6)</td>
</tr>
<tr>
<td>July 4, 2017</td>
<td>Module 11: Review and conclusion&lt;br&gt;Presenting your thesauri</td>
<td>Roe &amp; Thomas, chap. 6, 9-12</td>
<td>Final exam</td>
</tr>
</tbody>
</table>

The course is developed around the stages of construction of a domain-specific thesaurus. Assignments are:

**ASSIGNMENT DESCRIPTIONS**
Assignment 1 | Identify a domain; provide a description of the domain, and a list of essential source documents that can provide domain-specific vocabulary.
Assignment 2 | Extract essential vocabulary from the domain-specific sources identified in assignment 1. Decisions about indexing depth will be critical at this phase.
Assignment 3 | Coordinate the vocabulary. Cluster terms, and create a diagram of hierarchies and facets within the clusters. Identify homonyms and synonyms.
Assignment 4 | Structure the thesaurus and present it in alphabetical order.
Assignment 5 | Re-structure the conceptual nodes as an ontology. Comment.
Assignment 6 | Present a narrative justification of the domain, its vocabulary, your choices of terms, your decisions about clusters and facets.

**EVALUATION:**

<table>
<thead>
<tr>
<th></th>
<th>Grade Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discussions and small groups</td>
<td>20%</td>
</tr>
<tr>
<td>Assignments</td>
<td>50%</td>
</tr>
<tr>
<td>Final exam</td>
<td>30%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

**GRADING SCALE:**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>96-100</td>
<td>A</td>
<td>74-76</td>
</tr>
<tr>
<td>91-95</td>
<td>A-</td>
<td>70-73</td>
</tr>
<tr>
<td>87-90</td>
<td>B+</td>
<td>67-69</td>
</tr>
<tr>
<td>84-86</td>
<td>B</td>
<td>64-66</td>
</tr>
<tr>
<td>80-83</td>
<td>B-</td>
<td>60-63</td>
</tr>
<tr>
<td>77-79</td>
<td>C+</td>
<td>Below 60</td>
</tr>
</tbody>
</table>
UWM AND SOIS ACADEMIC POLICIES

The following links contain university policies affecting all SOIS students. Many of the links below may be accessed through a PDF-document maintained by the Secretary of the University:  http://www.uwm.edu/Dept/SecU/SyllabusLinks.pdf. Undergraduates may also find the Panther Planner and Undergraduate Student Handbook useful (http://www4.uwm.edu/osl/students/). For graduate students, there are additional guidelines from the Graduate School (http://www.graduateschool.uwm.edu/students/current/), including those found in the Graduate Student and Faculty Handbook: http://www.graduateschool.uwm.edu/students/policies/expanded/.

Students with disabilities. If you will need accommodations in order to meet any of the requirements of a course, please contact the instructor as soon as possible. Students with disabilities are responsible to communicate directly with the instructor to ensure special accommodation in a timely manner. There is comprehensive coverage of issues related to disabilities at the Student Accessibility Center (http://www4.uwm.edu/sac/), important components of which are expressed here: http://www.uwm.edu/Dept/DSAD/SAC/SACltr.pdf.

Religious observances. Students’ sincerely held religious beliefs must be reasonably accommodated with respect to all examinations and other academic requirements, according to the following policy: http://www4.uwm.edu/secu/docs/other/S1.5.htm.

Please notify your instructor within the first three weeks of the Fall or Spring Term (first week of shorter-term or Summer courses) of any specific days or dates on which you request relief from an examination or academic requirement for religious observances.

Students called to active military duty. UWM has several policies that accommodate students who must temporarily lay aside their educational pursuits when called to active duty in the military (see http://www4.uwm.edu/academics/military.cfm), including provisions for refunds, readmission, grading, and other situations.

Incompletes. A notation of “incomplete” may be given in lieu of a final grade to a student who has carried a subject successfully until the end of a semester but who, because of illness or other unusual and substantial cause beyond the student’s control, has been unable to take or complete the final examination or some limited amount of other term work. An incomplete is not given unless the student proves to the instructor that s/he was prevented from completing course requirements for just cause as indicated above (http://www4.uwm.edu/secu/docs/other/S31.pdf).

Discriminatory conduct (such as sexual harassment). UWM and SOIS are committed to building and maintaining a campus environment that recognizes the inherent worth and dignity of every person, fosters tolerance, sensitivity, understanding, and mutual respect, and encourages the members of its community to strive to reach their full potential. The UWM policy statement (http://www4.uwm.edu/secu/docs/other/S47.pdf) summarizes and defines situations that constitute discriminatory conduct. If you have questions, please contact an appropriate SOIS administrator.

Academic misconduct. Cheating on exams and plagiarism are violations of the academic honor code and carry severe sanctions, ranging from a failing grade for a course or assignment to expulsion from the University. See the following document (http://www4.uwm.edu/osl/dean/conduct.cfm) or contact the SOIS Investigating Officer (currently the Associate Dean) for more information.

Complaints. Students may direct complaints to the SOIS Dean or Associate Dean. If the complaint allegedly violates a specific university policy, it may be directed to the appropriate university office responsible for enforcing the policy.
Grade appeal procedures. A student may appeal a grade on the grounds that it is based on a capricious or arbitrary decision of the course instructor. Such an appeal shall follow SOIS appeals procedures or, in the case of a graduate student, the Graduate School. These procedures are available in writing from the respective department chairperson or the Academic Dean of the College/School.