ROCHESTER INSTITUTE OF TECHNOLOGY
ENGL 351 LANGUAGE TECHNOLOGY
Time: TuTh 11:00-12:15  Room: LBR-3244

Instructor: Emily Prud’hommeaux  E-mail: emilypx@rit.edu
Office: LBR-1317  Office hours: TBD & by email appt.

Course Description: We will explore the relationship between language and technology from the invention of writing systems to current natural language and speech technologies. Topics include script decipherment, machine translation, automatic speech recognition and generation, dialog systems, computational natural language understanding and inference, as well as language technologies that support users with language disabilities. We will also trace how science and technology are shaping language, discuss relevant intelligent computing concepts, and examine the ethical implications of advances in language processing by computers. Students will have the opportunity to experience computational text analysis with the Python programming language, speech analysis with Praat and other speech technology packages. This is an interdisciplinary course and technical background, while helpful, is not required. This course is also part of the Human Language Technology & Computational Linguistics immersion, the Language Science minor/immersion, and the Digital Literatures & Comparative Media minor.

Course topics:
- Writing systems, technology’s impact on language, decipherment
- Python programming for text analysis, speech analysis with Praat
- History of natural language processing and speech technologies
- Conceptual overview of core language technologies of today (spell checkers, speech synthesis, language learning, automatic speech recognition, machine translation, searching, dialog systems, document classification)
- Case studies with natural language corpora (data)
- Relevant concepts in intelligent computing, and ethical implications of language technologies

Course objectives: Upon successful completion of this course, you will be able to...
- Relate language to technology in socio-historical contexts
- Explain concepts related to the study of writing systems, and reflect upon how technology shapes language
- Conceptually explain how core language technologies work, and analyze their limitations and implications
- Recognize the value of data-driven natural language problem solving and insights in intelligent computing contexts, etc.
- Use Python for basic text analysis, and Praat for fundamental speech analysis
- Read short research articles on language technology topics for gist and conceptual understanding
- Discuss scientific principles (formulating a research question, addressing the question using experimental methods, evaluating the results in light of the research question), verification and vetting by scientific community
Course Policy: RIT requires students to attend classes. Please come to class and arrive on time. This is for your own benefit. You are responsible for knowing about material covered, including lectures, peer presentations, exercises, and group work. If you miss a session, prepare the readings and assignments for that class, borrow notes from a classmate, and talk to me if you have any lingering questions.

Classroom and expectations: This course takes place in a computing lab. We will use the special features of this classroom and engage with hands-on work. Please be professional and do not abuse the computer access in the lab. (Using social media, chatting, and surfing unrelated to class activities are examples of unprofessional behavior.) Learning in this course is cumulative, where each new topic may presuppose that you have acquired an understanding of concepts covered previously. Please take a critical point of view to the studied topics and confidently question course materials.

Equal access: RIT is committed to providing reasonable accommodations to students with disabilities. If you would like to request accommodations such as special seating or testing modifications due to a disability, please contact the Disability Services Office. It is located in the Student Alumni Union, Room 1150; the Web site is www.rit.edu/dso. After you receive accommodation approval, it is imperative that you contact me during office hours so that we can work out whatever arrangement is necessary.

Readings and media: The required course texts are:

The optional Python textbook is freely available on the web: Downey, A. Think Python (http://www.greenteapress.com/thinkpython/). Additional readings and media are available in myCourses, online, or via the library course reserves.

Course website: Class notes, the survey link, peer review forms, report criteria, etc., are in myCourses. You will need a working RIT email account so that you receive any announcements channeled though myCourses, as it uses RIT email addresses.

Final grading components:
- Preparation, participation, and attendance (class survey, peer reviews, lab/class work, preparing readings, discussion posts): 15%
- Fortnightly in-class "quick writes": 10%
- Problem sets (three): 20%
- Midterm and final exam: 30% (15% midterm + 15% final)
- Term Project: report and final presentation: 25% (15% report + 10% presentation)
Preparation, participation, and attendance: You should participate actively and make contributions in class. I expect that you are willing to learn, inquisitive, respectful, and that you engage in interactive activities in class (pair/group work, brain-storming, discussions, presentations, peer reviewing, etc.) and online, when assigned (viewing media, posting reflections on the discussion threads).

Quick writes: In-class, fortnightly "quick writes" give you feedback on a regular basis on your progress in the class and your understanding of the course contents.

Problem sets: These are exercises that allow interaction with language technology applications, analytical tasks, and involve work with Python, Praat, and other software packages. Submit your problem set report in class and your code in the Dropbox before class in a zipped file named PS[#]_{YOURLASTNAME(S)}.zip or a tarball named PS[#]_{YOURLASTNAME(S)}.tgz. Your report should include the command(s) and instructions required for running your code.

Midterm and final exam: The midterm takes place in week 8 and covers contents up to week 7. The final covers cumulative contents and takes place in finals week.

Term project (individual or max. 3 collaborators): For the term project, you will select one of two options: (1) a single-authored formal paper report focused on thematic literature review or an applied appropriate pilot usability study; or (2) a group-authored formal paper report paired with an implementation involving code development and usually data.

The term project is sequenced the following way:

1. Abstract proposal for instructor feedback: March 10
2. Article reflection (part of problem set 2): April 7
3. Draft of term project report for peer review feedback: April 28
4. Term project presentation: May 5 or 7
5. Final term project report: May 12

Abstract proposal: A 0.5- to 1-page project abstract proposal is submitted for instructor approval and feedback. Submit in class and in the assigned Dropbox.

Article reflection: You or your group will submit a one-page summary and critique on a self-selected article from the ACL (http://aclweb.org/anthology-new) or the ACM Digital Library (http://dl.acm.org/). The article should cover a language technology topic linked to your project. Submit with Problem Set 2.

Draft of term project reports: Drafts will peer reviewed. Submit the draft of your report in class and in the assigned Dropbox. If applicable, append code, data, or applied usability instruments.

Term project presentations: In week 14, you or your group will present your final project to the class. Presentations are allotted 10-15 min., depending on if it’s a group or an individual presentation, followed by a short Q&A. Submit the presentation materials in the assigned Dropbox on the day of your presentation (one per group is fine). Presentations are peer reviewed with a web form.
Final term project report: For the project report, you should use the standard two-column ACL template. Please use the style files for NAACL 2015; see http://naacl.org/naacl-pubs/. The report is 4 pages in the assigned format, plus one or more reference pages.

Source requirements for the report: If you chose to complete a thematic literature review, I expect it to cover at least eight academic, peer-reviewed references. In the case of a usability study or a group project, I expect your literature review section to cover at least four academic, peer-reviewed references. Approved sources include journal articles, conference proceedings papers, theses, dissertations, and books or book chapters from academic presses. Please consult with Lara Nicosia, the COLA Librarian, if you are unsure if a source is peer-reviewed or not.

Late work policy: For each day an assignment is late, I will deduct 10% of the total possible points for that assignment from your grade. For examples, if you get everything right but turn in your assignment 2 days late, you’ll get 80% of the possible points. Assignments will not be accepted more than one week after the original due date.

A note on assignments: Working in a responsible and ethically sound way with peers is an important skill in the intellectual process. There are both group and individual assignments. Throughout the course, follow the specific conditions in the table below, in regards to academic honesty. For extra writing consultation, turn to the Writing Commons: http://www.rit.edu/academicaffairs/writing/contacthours.

<table>
<thead>
<tr>
<th>Preparing readings</th>
<th>Student collaboration is expected and encouraged.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem sets: solving (&amp; write-up if team PS)</td>
<td></td>
</tr>
<tr>
<td>Reflections: preparing</td>
<td></td>
</tr>
<tr>
<td>Project work, report, and presentation (if option 2)</td>
<td></td>
</tr>
<tr>
<td>Quick writes</td>
<td>Individual exercises and collaboration of any kind is unacceptable.</td>
</tr>
<tr>
<td>(Problem sets: write-up if individual PS)</td>
<td></td>
</tr>
<tr>
<td>Reflections: write-up</td>
<td></td>
</tr>
<tr>
<td>Project work, report, and presentation (if option 1)</td>
<td></td>
</tr>
<tr>
<td>Midterm and final exams</td>
<td></td>
</tr>
</tbody>
</table>

Final grading: Your final letter grade will be assigned based on the following scale:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>93-100%</td>
</tr>
<tr>
<td>A-</td>
<td>90-92.99%</td>
</tr>
<tr>
<td>B+</td>
<td>87-89.99%</td>
</tr>
<tr>
<td>B</td>
<td>83-86.99%</td>
</tr>
<tr>
<td>B-</td>
<td>80-82.99%</td>
</tr>
<tr>
<td>C+</td>
<td>77-79.99%</td>
</tr>
<tr>
<td>C</td>
<td>73-76.99%</td>
</tr>
<tr>
<td>C-</td>
<td>70-72.99%</td>
</tr>
<tr>
<td>D</td>
<td>60-69.99%</td>
</tr>
<tr>
<td>F</td>
<td>&lt;60%</td>
</tr>
</tbody>
</table>
**Code of Academic Integrity:** As an institution of higher learning, RIT expects students to behave honestly and ethically at all times, especially when submitting work for evaluation in conjunction with any course or degree requirement. The Department of English encourages all students to become familiar with the [RIT Honor Code](http://infoguides.rit.edu/content.php?pid=170660) and with RIT's Academic Integrity Policy. As an additional resource on specifically the appropriate practices for citing work considered for your assignments in this course, see [http://infoguides.rit.edu/content.php?pid=170660](http://infoguides.rit.edu/content.php?pid=170660).

**WEEKLY OUTLINE**

You are expected to complete readings and assignments BEFORE the assigned class session. This syllabus is subject to change. Changes will be announced via email and reflected in the schedule at [http://tinyurl.com/qc6mh9q](http://tinyurl.com/qc6mh9q)

**Week 1**

1/27  Course introduction
1/29  Writing as language technology
      Reading: Sproat Ch. 1, 2

**Week 2**

2/3   Python: Variables, expressions, statements
      Optional Reading: Downey Ch. 1 and 2
2/5   How writing represents language, Decipherment
      Reading: Sproat Ch. 3, 4.0-4.2.3

**Week 3**

2/10  Rosetta Stone documentary viewing
      Reading: Sproat 4.2.4-4.2.5
      Reading: *Pathological Science* (in myCourses)
      **1-2 paragraph reflection on above article due**
2/12  Encoding language on computers
      Reading: DB&M 1.1-1.4
      Quick Write 1

**Week 4**

2/17  Writers’ aids
      Reading: DB&M 2.1-2.3
      **1-2 paragraph reflection on Rosetta Stone documentary due**
2/19  Writers’ aids
      Python: conditionals, control flow
      Reading: DB&M 2.4-2.5
      Optional reading: Downey 5.1-5.7
Week 5
2/24  Language tutoring
       Reading: DB&M 3.1-3.4.1
       Problem Set 1 released
2/26  Language tutoring
       Python: strings and string functions
       Reading: DB&M 3.4.2-3.6
       Optional Reading: Downey Ch. 8
       Quick Write 2

Week 6
3/3    Searching
       Reading: DB&M 4.1-4.3
       Problem Set 1 due
3/5    Searching
       Python: lists and dictionaries
       Reading: DB&M 4.4-4.5
       Optional Reading: Downey Ch. 10, 11

Week 7
3/10   Regular expressions
       Term project abstract proposal due
3/12   Fun with words
       Quick Write 3

Week 8
3/17   Midterm review
3/19   Midterm

Week 9
3/31   Classifying documents
       Reading: DB&M 5.1-5.4
       Problem Set 2 released
4/2    Classifying documents
       Reading: DB&M 5.5-5.6
       Quick Write 4
Week 10

4/7  History of mechanized speech and language tech  
    Reading: Sproat 6.0-6.4  
    Problem set 2 due (includes article reflection for term project)
4/9  Machine translation  
    Reading: DB&M 7.1-7.9, Sproat chapter 8

Week 11

4/14  Speech technology  
      Reading: Sproat 7.0-7.1  
      Problem Set 3 released
4/16  Speech technology  
      Reading: Sproat 7.2-7.3  
      Quick Write 5

Week 12

4/21  Speech technology
4/23  Dialog systems  
      Reading: DB&M 6.1-6.5  
      Problem Set 3 due

Week 13

4/28  Dialog systems  
      Reading: DB&M 6.6-6.10  
      Draft of paper due for peer feedback
4/30  Assistive technology  
      Quick Write 6  
      Peer feedback due

Week 14

5/5  Term project presentations
5/7  Term project presentations

Week 15

5/12  Final Review  
      Term project report due 11:59pm

Final Exam: Wednesday, May 20, 10:15, in our regular classroom (LBR-3244)