Syllabus
INLS 512, Applications of Natural Language Processing, Spring 2017

Instructor: Stephanie W. Haas   Office: 111 Manning Hall
Phone: 919-962-8360   Email: shaas at email dot unc dot edu

Class Meeting: 307 Manning Hall, Monday and Wednesday 12:20 – 1:35 p.m.
Office Hours: Monday and Wednesday 10:30 – 11:30 a.m. and by appointment

Syllabus Contents
   Course Description
   Rationale and Approach
   Keys to Success
   Learning Objectives
   Course Policies
   Policies on Academic Integrity and Diversity
   Grading Policies
   Communications
   Sakai
See the Course Schedule for additional information on topics, readings, and assignments. See the Assignment Overview for additional information on assignments.

Course Description
INLS 512, Applications of Natural Language Processing
Applications of natural language processing techniques and the representations and processes needed to support them. Topics include interfaces, text retrieval, machine translation, speech processing, and text generation.

This course is open to undergraduate and graduate students who have an interest in the intersection between computers and human language -- how computers can do useful things with language. Students do not need extensive programming skills; a sense of exploration however, is vital.

Rationale and Approach
This course is a survey of applications and their underlying techniques. It is intended for a wide variety of students: experience in linguistics or programming is not required. You and your classmates will have a range of skills and expertise, and the course will provide opportunities to learn from, and help each other.

Your work for this class falls into 4 categories: 1) preparation for class, 2) in-class activities and discussions, 3) homework assignments, and 4) examinations.
   Preparation: The schedule lists the topics and readings for each class meeting. Folders in the Sakai Resources/Class Materials contain materials for each class. Slides for each class will usually be added prior to the class meeting, in case you would like to take notes on them.
   Readings are taken from research literature, textbooks and collections, and manuals and handbooks.
   You are welcome to work together to prepare for class.

   In-class activities: In-class activities are a mix of lecture, discussion, and exercises.

Homework: Homework assignments include journal entries, exercises, preparation of a "Need To Know" presentation and a literature review.

Examinations: There will be a midterm and a final examination.

Keys to Success
- Plan ahead! Success in this course requires project management skills: identify milestones and deadlines, and plan your work accordingly.
- Coordinate the work schedule for this class with the schedules for your other classes, work, and other activities. You are likely to have many deadlines toward the end of the semester, so it's important for you to keep up.
- Give yourself plenty of time to prepare for each class. You may find some of the readings challenging; take the time to get as much as you can from each one. If you are not prepared for class, you will not be able to fully participate in (and benefit from) each class meeting.
- If you have any questions or concerns about the class or your work for the class, please talk to me about them. I cannot help you if I don't know there is an issue, and it is always easier to deal with problems earlier rather than later.

The most important key of all is to take advantage of this course to look at the world in a different way.

- **Enjoy and appreciate human natural language**. Pay attention to language and its context: what people say or write, the vocabulary and technologies they use, what they seek to accomplish with their language. Think about the challenges natural language poses for computational systems.
- **Think about the respective strengths and limitations of people and computers regarding language.** What do people do well? What do current NLP systems do well? How can we design systems that take advantage of their respective strengths, and/or compensate for their respective limitations? What new opportunities do you see for future applications of NLP?
- **Play!** There are many NLP tools and toys available online. Many research projects have websites with examples or demonstrations. Try them with different kinds of language and see what happens.

Learning Objectives
1. Students will learn about characteristics of language and its use, and understand their implications for Natural Language Processing and Natural Language Processing applications.
2. Students will survey a selection of NLP applications, and be able to describe the problems or task addressed, the materials and methods used, how the applications are evaluated, and opportunities for future developments for each application.
3. Students will learn about some basic NLP tools and techniques.
4. Students will identify and investigate a topic of individual interest in some depth by writing a literature review.

Course Policies

Preparation and Attendance.

I will always be prepared for class, and will start class on time. If unforeseeable circumstances prevent this for any reason, I will try to notify you beforehand if at all possible.

I expect the same of you: be prepared for class, be ready to start class on time, and send me an email in advance to explain why you must miss a class.

If you must miss a class session unexpectedly, email or call me as soon as possible (preferably within 24 hours) to explain why.

Class participation contributes to your final grade; you cannot participate if you are not present. Unexcused absences or repeated tardiness will lower your grade.

You are responsible for getting notes for a missed class from a classmate.

Reading assignments, exercises, study questions, and other preparation should be done before the class for which they are assigned so you can ask questions and participate in discussions. Some of the readings are introductory or textbook-like, others are research articles, which may be more complex. Give each reading your honest effort, but don't panic if there's something you don't understand. Make a note of your questions so we can discuss them in class.

If there is something you don't understand, say so! If you don't want to ask during class, come to my office hours or contact me by email.

Electronic Devices in Class

We may use a variety of software tools for in-class exercises and demonstrations. Many of them are web-based, and do not require downloading software, although some may. Please bring your laptop to class on days we do exercises.

During class, laptops, smartphones, and similar devices should be used only for class-related work. Unrelated activities such as texting, browsing, using social media, or playing games divert your attention from the class and are distracting and discourteous to others.

Please mute your phone before class starts.

Textbook

There is no textbook for this class. Readings and other materials come from a variety of journals, conference proceedings, and websites. A list of additional resources, including textbooks and tools, is available in the Sakai/Resources/Official Course Documents folder.

Assignments

Assignments must be submitted on time through the Assignment tool in Sakai.

Although you will be able to submit an assignment after the deadline, late assignments will be penalized 5% for every day it is late.

If you have a real problem submitting an assignment on time, please talk to me *before* the due date. Getting a late start on an assignment does not count as a real problem.

Pay attention to the instructions for each assignment. Be sure you have completed each part of the assignment, and proofread it before you submit it.

Start working on assignments well in advance of the due date. Do not wait until the last minute (or hour or day) to ask questions about the assignment – I may not be available for consultation.

Policies on Academic Integrity and Diversity

Chapel Hill has had a student-administered honor system and judicial system for over 100 years. Because academic honesty and the development and nurturing of trust and trustworthiness are important to all of us as individuals, and are encouraged and promoted by the honor system, this is a most significant University tradition. More information is available at http://studentconduct.unc.edu/honor-system

The UNC Honor Code is in effect for all work in this course. Chapter 5 of the "Instrument of Student Judicial Governance" (https://studentconduct.unc.edu/sites/studentconduct.unc.edu/files/documents/Instrument.pdf) discusses Academic Dishonesty; **you are responsible** for being familiar with, and abiding by all aspects of the "Instrument".

Students often ask what is okay to talk about with other students and what is not. There are some specific guidelines for this course.
- I do encourage you to help each other learn the course material – your fellow students can often be a great resource for learning. For example, you may want to work together when you are reading an assigned reading for class, or discuss your thoughts in response to a question posed in the reading notes.
- I also encourage you to collaborate on in-class discussions and exercises.
- All work you submit should be your own. Provide full citations for any material (quotations, diagram
- You may give and receive assistance regarding the use of hardware and software. For example, you may ask or answer a question such as "how do I [fill in the blank] in NLTK?").
- Individual homework assignments are to be done **individually**. You may consult the course readings, your notes, and even other print or web sources. (Keep in mind, however, that what you find in other sources may not be consistent with what I want you to do.) You may not consult your classmates or other people; all questions should be addressed to me.
- Tasks for collaborative homework assignments should be apportioned equitably. Each member of the group is responsible for contributing to the assignment.
- You must sign (check) the honor statement when you submit each assignment. This confirms that you and the work you submit conform to the Honor Code.

In support of the University's diversity goals and the mission of the School of Information and Library Science, SILS embraces diversity as an ethical and societal value. We broadly define diversity to include race, gender, national origin, ethnicity, religion, social class, age, sexual orientation and physical and learning ability. As an academic community committed to preparing our graduates to be leaders in an increasingly multicultural and global society we strive to:
- Ensure inclusive leadership, policies, and practices;
- Integrate diversity into the curriculum and research;
- Foster a mutually respectful intellectual environment in which diverse opinions are valued;
- Recruit traditionally underrepresented groups of students, faculty and staff; and
- Participate in outreach to underserved groups in the State.

The statement represents a commitment of resources to the development and maintenance of an academic environment that is open, representative, reflective and committed to the concepts of equity and fairness.

Grading Policies

Your final grade will be based on the following:

I Need to Know About	10%
Literature Review Part 1	5%
Literature Review Part 2	10%
Literature Review Part 3	20%
Preparation and Participation	5%
Exercises	10%
Journal Entries	10%
Midterm Examination	15%
Final Examination	15%

See the Assignment Overview for additional information on course assignments; details for each will be distributed as they are assigned.

The following grade scale will be used AS A GUIDELINE (subject to any curve) **for undergraduate** students:

Grade Range	Definition*
A 90-100%	Mastery of course content at the <u>highest level of attainment</u> that can reasonably be expected of students at a given stage of development. The A grade states clearly that the students have shown such <u>outstanding promise</u> in the aspect of the discipline under study that he/she may be strongly encouraged to continue.
B 80-89.9%	<u>Strong performance</u> demonstrating a high level of attainment for a student at a given stage of development. The B grade states that the student has shown solid promise in the aspect of the discipline under study.
C 70-79.9%	A <u>totally acceptable</u> performance demonstrating an adequate level of attainment for a student at a given stage of development. The C grade states that, while not yet showing unusual promise, the student may continue to study in the discipline with reasonable hope of intellectual development.
D 60-69.9%	A <u>marginal performance</u> in the required exercises demonstrating a minimal passing level of attainment. A student has given no evidence of prospective growth in the discipline; an accumulation of D grades should be taken to mean that the student would be well advised not to continue in the academic field.
F 0-59.9%	For whatever reason, an <u>unacceptable performance.</u> The F grade indicates that the student's performance in the required exercises has revealed almost no understanding of the course content. A grade of F should warrant an advisor's questioning whether the student may suitably register for further study in the discipline before remedial work is undertaken.

*Definitions are from: http://registrar.unc.edu/academic-services/grades/explanation-of-grading-system/
(underlining is my emphasis)

The following grade scale will be used AS A GUIDELINE (subject to any curve) for **graduate** students:

Grade Range	Definition*
H 95-99%	High Pass
P 80-94.9%	Pass
L 70-79.9%	Low Pass
F 0-69.9.9%	Fail

*Definitions are from: http://registrar.unc.edu/academic-services/grades/explanation-of-grading-system/

Communications

The best way to get in contact with me (other than talking to me after class or during my office hours) is by email: shaas at email dot unc dot edu. Note that I receive a large amount of email and while I try to reply to student emails within 48 hours, there are times that it may take me 2-3 days to reply. Therefore, it is important that you get started on assignments early, so there is time for me to respond to any questions you may have. I cannot guarantee that I will be able to answer last-minute questions (e.g., within 2 days of the assignment due date).

You may also call me at 919-962-8360.

Email and phone call are both good ways of letting me know if you can't be in class, or if you want to make an appointment with me.

If you want to discuss something we talked about in class, and assignment or some other matter, I prefer we meet in person. Please come to my office hours or make an appointment with me.

I will use the message/email facility within Sakai to send announcements to the class, as well as to individual students.

Sakai

We will use Sakai for almost all course activities. All enrolled students should have access to the UNC Sakai site for this course: http://sakai.unc.edu/

Course Materials

Course materials are stored in folders in the Sakai/Resources tool.

Official Course Documents. This folder includes the syllabus, schedule, an overview of assignments, and other important information.

Class Materials. This folder contains a folder for each class meeting, labeled with the class number and date.

Contents of each class folder may include:
- Overview: reading and other work that you should complete before the class, as well as any in-class activities you should prepare for
- Articles that aren't available on the web
- Slides that will be used during the class (usually added 1 day before the class meeting)

It is your responsibility to check each class folder enough in advance so you can prepare for the class.

Need To Know. This folder will contain the slides and recommended resources you prepare for your Need To Know assignment; these are available to all members of the class.

<u>Submitting assignments</u>
In order for you to receive credit for an assignment, it must be submitted following the submission instructions given in the assignment specifications. Most assignments will be submitted through the Sakai Assignment tool; the exception is the Need To Know assignment.

If for some reason you are unable to submit an assignment to Sakai, as a last resort you may email it to me along with a note about the problem you encountered. Then, as soon as you are able to, it is your responsibility to submit the exact same assignment to Sakai. The email serves as a record that you tried to submit the assignment on time, but to receive credit, your assignment must be uploaded to Sakai.

I will use the Sakai Gradebook to record your course grades.

<u>Dropbox</u>
You each have a drop box in Sakai that is accessible only to you and me. You may store work in progress here. If you have a question about an assignment, and it would be helpful for me to see your work, you may store the draft in your drop box, and refer to it in your emailed question. I will look at only that file to respond to your question. Do NOT submit homework by putting it into your drop box.

INLS 512-001 Applications of Natural Language Processing Spring 2017
Monday & Wednesday, 12:20 – 1:35, Manning 307
Schedule Overview

This version of the *tentative* schedule contains topics, assignments, and due dates, to help you plan your work this semester.

1. Wednesday January 11, 2017: Spoken Dialogue Systems; Speech Interaction (1)
Assignment
- Survey of Skills and Interests (Sakai Test tool) due by Tuesday 1/17/17, 9:00 a.m.
- Read the UNC "Instrument of Student Judicial Governance" (https://studentconduct.unc.edu/sites/studentconduct.unc.edu/files/documents/Instrument.pdf), focusing especially on Ch. 5, which discusses Academic Dishonesty. Then send me an email message that states you have read the policy and will abide by the Honor Code in this course. Your email is due by Tuesday 1/17/17, 9:00 a.m.

Monday January 16, 2017 No Classes

2. Wednesday January 18, 2017: Spoken Dialogue Systems; Speech Interaction (2)
Survey of Skills and Interests due by Tuesday 1/17/17, 9:00 a.m.
Honor Code email due by Tuesday 1/17/17, 9:00 a.m.
Assignment
- Sign up for a Need To Know presentation date (Doodle Poll) by Tuesday 1/24/17, 8:00 a.m.

3. Monday January 23, 2017: Spoken Dialogue Systems; Speech Interaction (3)
Assignment
- Assign Literature Review Part 1, due Monday 2/6/17, 12:00 noon
- Reminder: sign up for a Need To Know presentation date (Doodle Poll) by Tuesday 1/24/17, 8:00 a.m.

4. Wednesday January 25, 2017: Spoken Dialogue Systems; Speech Interaction (4)
Assignment
- Exercise 1: Evaluate speech interaction system. Due Monday 1/30/17, 12:00 noon.

5. Monday January 30, 2017: Spoken Dialogue Systems; Speech Interaction (5)
Exercise 1 due in Sakai 12:00 noon
Assignment
- Journal 1: Thoughts on Spoken Dialogue Systems and Speech Interaction. Due Wednesday 2/1/17 12:00 noon

6. Wednesday February 1, 2017: Biomedical and Health Informatics Applications (1)
Journal 1 due in Sakai 12:00 noon

7. Monday February 6, 2017: Biomedical and Health Informatics Applications (2)
Literature Review P1 due in Sakai 12:00 noon

8. Wednesday February 8, 2017: Biomedical and Health Informatics Applications (3)
Assignment
- Exercise 2: Corpus design. Due Monday 2/13/17 12:00 noon.

9. Monday February 13, 2017: Biomedical and Health Informatics Applications (4)
Exercise 2 due in Sakai 12:00 noon.
Assignment
- Literature Review P2, due Wednesday 3/8/17, 12:00 noon.
- Journal 2: Thoughts on Biomedical and Health Informatics Applications Due Wednesday 2/15/17, 12:00 noon

10. Wednesday February 15, 2017: Natural Language Generation (1)
Journal 2 due in Sakai 12:00 noon

11. Monday February 20, 2017: Natural Language Generation (2)

12. Wednesday February 22, 2017: Natural Language Generation (3)
Assignment
- Exercise 3, Exploring Words and Phrases. Due Monday 2/27/17, 12:00 noon.

13. Monday February 27, 2017: Natural Language Generation (4)
Exercise 3 due in Sakai 12:00 noon.
Assignment
- Journal 3: Thoughts on Natural Language Generation. Due Monday 3/6/17, 12:00 noon

14. Wednesday March 1, 2017: Midterm Examination

15. Monday March 6, 2017: NLP for Digital Humanities (1)
Journal 3 due in Sakai 12:00 noon

16. Wednesday March 8, 2017: NLP for Digital Humanities (2)
Literature Review P2 due in Sakai 12:00 noon

Monday March 13, 2017 & Wednesday March 15, 2017 Spring Break, No Classes

17. Monday March 20, 2017: NLP for Digital Humanities (3)
Assignment
- Literature Review P3, due Wednesday 4/12/17, 12:00 noon

18. Wednesday March 22, 2017: NLP for Digital Humanities (4)
Assignment
- Journal 4: Thoughts on NLP for Digital Humanities. Due Monday 3/27/17, 12:00 noon

19. Monday March 27, 2017: NLP and Social Media (1)
Journal 4 due in Sakai 12:00 noon

20. Wednesday March 29, 2017: No class meeting
Assignment
TBA

21. Monday April 3, 2017: NLP and Social Media (2)

22. Wednesday April 5, 2017: NLP and Social Media (3)

23. Monday April 10, 2017: NLP and Social Media (4)

24. Wednesday April 12, 2017: NLP and Social Media (5)
Literature Review P3 due in Sakai 12:00 noon
Assignment
- Journal 5: Thoughts on NLP and Social Media. Due Monday 4/17/17, 12:00 noon.

25. Monday April 17, 2017: Deception, Argument, and Claims (1)
Journal 5 due in Sakai 12:00 noon
Assignment
- Exercise 4: Need To Know Follow-up. Due Wednesday 4/19/17, 12:00 noon

26. Wednesday April 19, 2017: Deception, Argument, and Claims (2)
Exercise 4 due in Sakai 12:00 noon.
Assignment
- Journal 6: Thoughts on Deception, Argument, and Claims. Due Monday 4/24/17, 12:00 noon
- Exercise 5: Your Literature Review in 3 sentences. Due Monday 4/24/17, 12:00 noon.

27. Monday April 24, 2017: Literature Review Summaries
Journal 6 due in Sakai 12:00 noon
Exercise 5 due in Sakai 12:00 noon

28. Wednesday April 26, 2017: Wrap-up

Tuesday May 2, 2017 12:00 noon Final Exam

Assignments for this course fall into 3 overlapping groups: Exploration, Engagement, and Evaluation.

- Exploration assignments are opportunities for you to explore a topic in NLP that is of particular interest to you. I Need to Know About and the Literature Review focus on Exploration.
- Engagement assignments are designed to engage you with concepts, ideas, questions, and challenges in NLP. Some are independent assignments, others may require discussion and collaboration with your classmates. Preparation, Participation, Exercises, and Journal Entries focus on Engagement.
- Evaluation assignments (aka "examinations") are an additional means for me to assess your learning in this course.

Exploration

I Need to Know About…(10% of course grade)
Assigned Wednesday 1/18/17; presentations scheduled throughout the semester.
Throughout your life, you will need to learn a little about a new idea, technology, tool, or application – enough to know what it is, and whether you should (or want to) spend additional time and effort to learn more. In many workplaces, colleagues share information about something they've just learned, informally or at meetings or brown bag discussions.
The goal of this assignment is to investigate a current idea, technology, tool, application, or problem in Natural Language Processing (NLP), and brief your classmates on what you've learned. Think of this as a collaborative survey of interesting and exciting developments in NLP, that will increase the number of things we can all learn about!

Literature Review (35% of course grade)
Part 1: Assigned Monday, 1/23/17, due Monday 2/6/17, (5% of course grade)
Part 2: Assigned Monday, 2/13/17, due Wednesday 3/8/17, (10% of course grade)
Part 3: Assigned Monday 3/20/17, due Wednesday 4/12/17, (20% of course grade)
The goal of this semester-long assignment is to give you an opportunity to read and synthesize the research literature about a topic of interest to you. Part 1 is a topic paragraph, briefly describing the topic, questions you want to explore, and 1 article you have read about the topic. Part 2 is an extended outline of your paper, including topic statement, major ideas you will cover in each section, and bibliography. Part 3 is the full literature review.

Engagement

Preparation and Participation (5% of course grade)
Active participation requires being prepared for each class meeting. Preparation includes:
- reading/viewing/playing with assigned articles, videos, demos, and other materials.
- understanding core concepts and ideas presented in these materials and/or preparing questions about things you don't understand,

- thinking about study questions, and
- completing exercises and assignments on time.

Participation in class is composed of attendance, paying attention to presentations and others' comments, responses to questions, contributions to discussions, and participation in exercises and activities. Your participation should be courteous and professional at all times.
- I do take attendance – you cannot participate in class if you are not there. Unexcused absences or repeated tardiness will lower your participation grade.
- I expect you to be an active participant in class. Active participation requires being prepared for class (see above).
- One of the strengths of this class is the variety of backgrounds, interests, and expertise that you all bring to it. We cannot learn from your ideas and perspectives if you don't share them with us, in full class discussions and brainstorming sessions, in small group exercises, and in other classroom activities. Three additional points arise from this observation.
  - First, it is likely that each of you is familiar with some topics, and unfamiliar with others. For example, you may be multi-lingual, or have a background in linguistics or cognitive science, or have strong programming skills. Conversely, any of these could be entirely new areas for you. Regardless, you can contribute to our learning as a class. Your question may help all of us understand a concept more clearly; you may be able to provide a clear definition or example of a concept, or you may be able to help a classmate who is puzzled by an exercise.
  - Second, quality of contributions is more important than sheer quantity. It is important that everyone have an opportunity to contribute.
  - Third, there are many ways to participate. Encouraging your classmates to participate, for example by talking to them rather than just to me, asking them a question, leaving a space for other comments, or bringing a comment they made in a small group to the attention of the class, also count as participating in class.

Exercises (10% of course grade)
Several times during the semester, I will assign brief exercises. The purpose of these exercises is to let you work through common NLP problems, and experiment with some (usually) easy-to-use tools or techniques.
- Most of these exercises will require you to submit a solution via Sakai to receive credit for the exercise.
- You will receive credit for your exercise if it demonstrates honest effort to complete it, even it if is not completely correct. You will not receive credit if your exercise demonstrates little to no effort, or if it is not submitted on time.

Journal Entries (10% of course grade)
Your reflective journal is a way of demonstrating your engagement with course concepts, applications, and ideas.

Several times during the semester, I will assign prompts. You will write a paragraph or two about your thoughts and reflections in response to the prompt. Of course, you are welcome to record your reflections about other course topics as well! Submit each journal entry through the Sakai Assignment Tool.

Journal entries will be graded on a 4-point scale.
3 points: excellent entry, demonstrating that you have considered aspects of the topic going beyond what was covered in class or requested in the prompt. For

example, an excellent entry could refer to additional relevant articles or suggest novel examples or applications of the topic.

2 points: satisfactory entry, covering all the aspects requested in the prompt. A satisfactory entry demonstrates that you were engaged in the topic, and have thought about what you learned.

1 points: minimal entry, including some portion of the aspects requested in the prompt. A minimal entry demonstrates that you have some awareness of the topic.

0 points: no entry submitted.

Evaluation

<u>Midterm Examination (15% of course grade)</u>
Wednesday 3/1/17, during class.

<u>Final Examination (15% of course grade)</u>
Tuesday 5/2/17, 12:00 noon

Contribution of Assignments to Course Grade
I Need to Know About 10%
Literature Review Part 1 5%
Literature Review Part 2 10%
Literature Review Part 3 20%
Preparation and Participation 5%
Exercises 10%
Journal Entries 10%
Midterm Examination 15%
Final Examination 15%