1. Course Information
This course provides an introduction to some of the best practices underlying Information Retrieval (IR) systems. It aims to help the students understand some of the theories and their application for classifying relevant and irrelevant information, storage and retrieval of documents, and the principles related to keeping humans in the loop. Focus is on text documents and their retrieval. Topics covered include document representation, document ranking, relevance ranking and feedback, indexing and searching, user interfaces, and the evaluation of information retrieval system effectiveness.

Students who successfully complete the course will have gained:
- Knowledge of the variety of available IR systems, and the methods embodied in these systems;
- Understanding of fundamental concepts related relevance and ranking that are key to success of information retrieval systems;
- Skills in the critical analysis and evaluation of the performance of IR systems.

Prerequisites and Technology Background: This course involves math. Therefore, students are expected to be comfortable in probably and statistics. Familiarity with computers and some programming experience are highly desirable, but not necessary.

Course format: This course will include a combination of lectures, classroom exercises, as well as seminar-style readings, presentations, and discussions.


Course Website and Blackboard: Blackboard will be used to provide essential course materials, the most current syllabus, and assignment documents.

2. Course Requirements
Readings: Students are expected to read the assigned materials before each class. For specific readings, please see the semester schedule at the end of this syllabus.

Exam: One midterm exam will be given in class. There is no final exam.

Assignments: Homework assignments are given in the form of problem sets. Each problem set will include essay-type questions, questions designed to show understanding of specific
concepts that may involve calculations, and hands-on exercises involving existing IR engines. Students should complete each assignment in teams and hand-in the work on time.

**Final Project**: The final course project will be an extensive literature review paper on an area of IR. Oral presentation will be given to the class.

**Style Manuals and Guidelines**: Reports should be word-processed and should be double-spaced. Students are required to cite sources, if any are used in their written reports, according to either the American Psychological Association (APA) or Turabian style manual. Choose only one style manual and use it throughout the report.


3. **Student Performance Evaluation**
The assignments, exam, and the project contribute to the final grade as follows:

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Percentage of Total Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Participation</td>
<td>20%</td>
</tr>
<tr>
<td>Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Exam</td>
<td>30%</td>
</tr>
<tr>
<td>Final Project and Presentation</td>
<td>30%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>

Grades are determined on a 100-point scale as follows:

**Letter Grade**

<table>
<thead>
<tr>
<th>A</th>
<th>A-</th>
<th>B+</th>
<th>B</th>
<th>B-</th>
<th>C+</th>
<th>C</th>
<th>C-</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeric Range</td>
<td>95-100</td>
<td>90-94</td>
<td>85-89</td>
<td>80-84</td>
<td>75-79</td>
<td>70-74</td>
<td>65-69</td>
<td>60-64</td>
<td>50-59</td>
</tr>
</tbody>
</table>

4. **Course Policies**

**Class Attendance and Participation**: Students are expected to be prompt and prepared for class as well as participate in the classroom and online discussions. Students are asked to notify the instructor in advance if they cannot attend class, must arrive late or leave early, expect to submit work late, or intend to withdraw from the course.

**Late Submissions**: Late submissions will not be accepted without the express permission of the instructor and students who submit their work late will lose half a letter grade per late day at the discretion of the instructor. A late final project or final exam will be penalized a full letter grade for each day late.

**Incompletes**: A tentative grade of “I” for Incomplete is given only when the student has nearly completed the course but due to circumstances beyond the student’s control cannot complete the course on schedule. The student is responsible for contacting the instructor and requesting an “I” grade in advance of the semester end. The conditions of the “I” grade, including the timeline for the completion of the work will be specified by the instructor. The “I” grade is automatically changed to “E” unless the work is completed as agreed between the student and the instructor.
**Academic Dishonesty:** The instructor has a zero tolerance policy for academic dishonesty, plagiarism, and cheating. Any such activity will be reported to the appropriate office.

**Students with Disabilities:** Reasonable accommodations will be provided for students with documented physical, sensory, systemic, cognitive, learning and psychiatric disabilities. If you believe you have a disability requiring accommodation in this class, please notify the instructor.

5. **Tentative Course Calendar**

*All reading assignments must be completed prior to the following week’s lectures. The lecture and reading lists are subject to change.*

1. Week of August 28 – Topics: Introductions, housekeeping, math review.
2. Week of September 4 – Topics: Architecture of a Search Engine. Reading: CMS Chapter 1&2
3. Week of September 11 – Topics: Crawls and Feeds. Reading: CMS Chapter 3
4. Week of September 18 – Topics: Processing Text. Reading: CMS Chapter 4
5. Week of September 25 – Topics: Processing Text. Reading: CMS Chapter 4
7. Week of October 16 – Topics: Queries and Interfaces. Reading: CMS Chapter 6
8. Week of October 23 – Midterm review. Reading: CMS Chapter 1-6
9. Week of October 30 – Midterm
10. Week of November 6 – Topics: Retrieval Models. Reading: CMS Chapter 7
12. Week of November 20 – Topics: Social Search. Reading: CMS Chapter 10
13. Week of November 27 – Presentation of final research projects
14. Week of December 4 – Presentation of final research projects