INFSCI 1560 Information Storage and Retrieval

Department of Computer Science, SCUPI

Spring 2026

Instructor

Ziyi Wang
Email: ziyi.wang@scupi.cn
Office: N527
Office Hours: Friday 10:00-15:00 or by appointment

Lectures

- Section 1: Tuesday 13:50-17:30 @ S103
- Section 2: Thursday 13:50-17:30 @ S103

Course Description

Problems and techniques related to storing and accessing unstructured information with an emphasis on textual information. Overview of several approaches to information access with a primary focus on search-based information access. Covers automated retrieval system design, content analysis, retrieval models, result presentation, and system evaluation. Examines applications of retrieval techniques on the web, in multimedia and multilingual environments, and in-text classification and event tracking.

Prerequisites

- INFSCI 1500 Database Management Concepts and Applications

Course Objectives

Upon finishing this course, the students should be able to

- understand the dimensions of the information retrieval problem
- master the analysis and design of information retrieval systems
- consider the factors which optimize the information retrieval process
- examine current issues in information retrieval

Upon satisfactory completion of this course, students will

- be able to explain core concepts and terms of information retrieval
- be able to explain different retrieval models and basic algorithms
- be able to evaluate existing information retrieval systems and suggest how the systems can be improved
- be able to apply theories to effectively solve information retrieval problems in real world situations

Applicable ABET Outcomes

- Analyze a complex computing problem and apply principles of computing and other relevant disciplines to identify solutions.
- Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.

- Communicate effectively in a variety of professional contexts.
- Recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles.
- Function effectively as a member or leader of a team engaged in activities appropriate to the program's discipline.
- Apply computer science theory and software development fundamentals to produce computing-based solutions.

Textbook

- Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press. 2008. Available at https://nlp.stanford.edu/IR-book/information-retrieval-book.html.

Reference

- Ricardo Baeza-Yates, Berthier Riberiro-Neto. *Modern Information Retrieval, 2nd Edition*. Addison Wesley. 2011.
- Mitra, Bhaskar, and Nick Craswell. *An Introduction to Neural Information Retrieval*. Now Publishers. 2018.
- Stefan Büttcher, Charles L. A. Clarke, Gordon V. Cormack. *Information Retrieval: Implementing and Evaluating Search Engines*. MIT Press. 2010.

Grading

- Attendance 20%
- In-Class Test 20%
- Homework 30%
- Project 30%

Conversion from Numerical Score to Letter Grade

100~90	89~85	84~80	79~76	75~73	72~70	69~66	65~63	62~61	60	<60
A	A–	B+	B	B–	C+	C	C–	D+	D	F

Communication

All lecture notes, learning materials, homework assignments, projects, and announcements will be published on Blackboard (https://pibb.scu.edu.cn/). It is the student's responsibility to regularly check Blackboard in a timely manner. Important announcements and notifications will be sent to QQ group.

Class Policies

Attendance

Class attendance will be checked on a weekly basis and each absence incurs a 2% deduction of your total score. For one-time absence, if you have a strong and valid reason, please provide official proof to the instructor.

In-Class Test

In-class tests are open book, open notes, and open computer. However, the students are not allowed to communicate either in person or through electronic devices. The use of AI is also strictly prohibited. Any misconduct will forfeit the student's right to take the test.

Homework

Homework assignments are submitted every four weeks. Late submission induces a penalty of 25% per week. For example, the first homework assignment is due at the end of week 4, submission at any time in week 5, 6, 7, and 8 induces a penalty of 25%, 50%, 75%, and 100%, respectively. Questions regarding the grading of homework assignment must be raised to the instructor within a one-week time frame.

Project

Students must report their project progress in the presentation sessions. Questions regarding the grading of project must be raised to the instructor within a one-week time frame.

Academic Integrity

The principles of academic integrity require that a student make sure that all work submitted is the student's own and created without the aid of impermissible technologies, materials, or collaborations. Academic integrity policy will be strictly followed.

Classroom Recording

To ensure the free and open discussion of ideas, students may not record classroom lectures, discussion and/or activities without the advance written permission of the instructor, and any such recording properly approved in advance can be used solely for the student's own private use.

Copyrighted Material

All material provided through this web site is subject to copyright. This applies to class notes, slides, assignments, solutions, project descriptions, etc. You are allowed (and expected!) to use all the material provided for personal use. However, you are strictly prohibited from sharing the material with others in general and from posting the material on the Web or other file sharing venues.

Tentative Course Schedule

Week	Topic
1	Introduction
2	Boolean Retrieval
3	Term Vocabulary & Postings Lists
4	Dictionaries & Tolerant Retrieval
5	Index Construction
6	Project Presentation
7	Index Compression
8	Scoring, Term Weighting, and Vector Space Model
9	Computing Scores in a Complete Search System

10	Evaluation in Information Retrieval
11	Reference Feedback & Query Expansion
12	Project Presentation