INFO 624 – Final Project

Mat Kelly
Drexel University

Overview

The final project is a major component of this course and accounts for 25% of your final grade (22 points plus 3 points for project abstract, see below). It is designed to give you hands-on experience applying information retrieval concepts to a substantial, open-ended problem. You may complete the project individually or as part of a team, depending on course guidance.

You may choose one of two project types:

Project Type: Search Engine Building

In this option, you will design, build, and evaluate a search engine for a specific domain or dataset. Your work should demonstrate thoughtful decisions about data collection, indexing, query design, ranking, and evaluation. The final submission includes a written report documenting why you built the system, what data it uses, who it serves, and how it works, along with a formal evaluation of its effectiveness. If applicable, you should also submit source code and links to any user interface you developed.

For a search engine building project, please submit your report (`.md` or `.pdf`) along with any source code (if applicable).

Report Requirements

The report should provide documentation about the following:

- **Why**: Explain the purpose of the search engine and what difference you can make.
- **What**: Describe the data and domain in which you build the search engine index. Please provide details about where the data source is and how you collected the data.
- **Who**: Explain whom your search engine may serve and their basic information needs. Ideally, you should be able to describe three use cases (information needs).
- **How**: Describe related decisions you have made and steps you have taken to build the engine.

How Details

1. Describe the data fields related to the project and how they should be analyzed and indexed. Select proper data types, analyzers, filters, etc. for each field, and provide your rationale.
2. Describe your search queries in terms of the use cases (needs):
   - What keywords and query structure should be used?
   - What fields should be searched for potential matches?
   - What fields should be included in the scoring (ranking)? In what manner?
3. Select related similarity, scoring, and boosting methods accordingly. Describe them and provide your reasons.
4. Create your index, mappings, and load data.

How Good: Evaluation

1. Test your query for each use case (information need); you may limit the number of hits for your search.
2. Examine the results and judge each document's relevance in terms of the information need.
3. Provide a formal evaluation in terms of metrics such as precision, recall, F1, and/or nDCG.
4. Discuss the results. You are encouraged to try different settings and compare their results.

Where

Where is your search engine (index)? Provide the names of your indices.

Experiences

Discuss your team experiences working on the search engine. What works? What doesn't? What could have been done better? What have you learned as a team? Include your thoughts on future work.

Code and Interface (if applicable)

If you have built a front end (or anything related) to your search engine index, please provide the link to your (web) interface and/or source code with instructions.

Project Type: Literature Review

In this option, you will write an in-depth literature review on a topic related to information retrieval. The paper should synthesize and critically analyze existing research, connect ideas across sources, and clearly situate your topic within the broader IR landscape. The paper must follow the ACM journal format and include a sufficient number of high-quality references.

Regardless of project type, the goal of the final project is to demonstrate your understanding of information retrieval principles, your ability to apply them thoughtfully, and your capacity to clearly communicate technical ideas, design decisions, and results.

If your project is a literature review, the requirements are as follows:

1. Your paper should be about 8 pages (minimum 6 pages and up to 10 pages), in the ACM journal format.

2. You should review and integrate at least 12 references and provide proper citations.

Evaluation Criteria

Your final literature review paper will be evaluated based on:

- **Paper formatting**: Adherence to formatting requirements (template), paper length, and citation quality.
- **Relevance to information retrieval**: Integration with existing IR research.
- **Quality of presentation**: Writing quality, clarity, literature integration, critical analysis, flow of ideas, and findings.
- **Other aspects specific to the work**

Please convert your file to a PDF and submit the PDF version.

ACM Template:
`https://www.acm.org/binaries/content/assets/publications/taps/acm_submission_template.doc`

Project Abstract

Due: *February 16, 2026, 6 pm ET*
Points: 3 (of the 25 total for the project)

Assignment Description

You are required to submit a project abstract describing your proposed final project. This abstract serves as an early checkpoint to help ensure that your project idea is clear, feasible, and aligned with the goals of the course.

The abstract is not a draft of your final report. Instead, it should concisely explain what you plan to do, why the project is meaningful, and how you intend to approach it.

Length and Format

- 250–400 words
- Submit as plain text in Canvas

Required Content

Your abstract must clearly address the following:

1. **Project Type** Indicate whether you are completing a **Search Engine Building** project or a **Literature Review**.
2. **Project Focus**
   - Search engine project: Describe the problem your search engine will address and its motivation.
   - Literature review: Describe the information retrieval topic or research question you will examine.
3. **Data or Sources**
   - Search engine project: Identify the dataset(s) you plan to use and how you will obtain them.
   - Literature review: Describe the types of academic sources you expect to review.
4. **Users or Research Perspective**
   - Search engine project: Identify the intended users and at least one example information need.
   - Literature review: Describe the analytical perspective you will take (e.g., comparison of methods, trends, or challenges).
5. **Planned Approach**
   Briefly explain how you plan to carry out the project (e.g., indexing and evaluation strategy, or how you will organize and synthesize the literature).
6. **Team Information (if applicable)**
   List all team members and briefly describe how responsibilities will be shared.

Grading Criteria

Your abstract will be evaluated based on:

- Clarity and organization
- Appropriateness of scope
- Alignment with course objectives
- Evidence of thoughtful planning

Feedback will be provided to help you refine your project before final submission.