

## COSI 132A Spring 2017 Information Retrieval Final Project

Your final project is a chance to show that you can use the concepts learned in class to develop an information retrieval application tailored to a specific domain. It also gives you an opportunity to exercise creativity and coordinate software development in a team context.

As in previous programming assignments, **you will submit commented code along with a readme.pdf file** that includes instructions for building and running your application, software dependencies, etc. Include the author of the code within the header of each code module. Include a small sample corpus that can be used to test your code. Additionally, **submit a powerpoint presentation** that will serve as high level documentation of your project. These will be due Wednesday, May 10, the day before your final presentations.

Each team will have two chances to present your work.

**On May 2, team members will present a project status report (using powerpoint slides) describing the project's mission, approach (including who is implementing what), progress and issues.**

**Your powerpoint presentation** should be detailed enough that *it could be understood on its own, without a presenter*. Use complete sentences rather than short bullets whenever necessary for clarity.

The focus of your presentation should be on the creative parts of the application design (the differentiators). Divide your presentation into sections so that each team member has a chance to speak. Each team will have roughly 7 minutes, which should be 5 minutes of planned presentation and 2 minutes of questions/discussion. **Make sure you have a way to connect your laptop to the classroom's hdmi cable.**

**On Thursday, May 11, we will have our final presentations in Volen 119 from 1:30 – 4:30. Each team will have ~18 minutes (15 minutes for slides and demo, 3 minutes for questions/discussion).**

**Your final powerpoint presentation** should serve as documentation for your project and contain the following sections:

1. Project name
2. Team members and their primary contributions (e.g., corpus processing, UI development, elasticsearch schema definition, information extraction, knowledge base development, algorithm implementation)
3. Purpose
4. Method:
  - Overview
  - Corpus acquisition and processing
  - Database design and indexing
  - Runtime functionality
5. Demo illustrating several information seeking scenarios, showing UI, results, and explanation
6. Lessons learned:
  - What worked well
  - What didn't work as well as planned (and why)
  - Next steps, if you were to continue improving this application

Your final oral presentation should use the slides submitted and should cover the sections not covered in the earlier status report. (Team members will be graded individually.) Plan on 18 minutes, including time for a live demo. **Rehearse your presentation to assure it fits into the allotted time.**

**Submissions:** Each team must submit two items to latte; each team member submits a readme.pdf file, as follows.

Each *team* should upload code (once) into latte as a single zip file named "<application\_name>". Use underscore instead of any white space in the title. Depending on the size of your corpus, include a subset of your corpus that can be used to test your application.

Each *team* should upload your team's powerpoint presentation (once) as <application\_name>.ppt

Each *team member* should upload a copy of the team's readme.pdf file.

Include in your readme.pdf file

- Examples of queries/interactions that work over the test subset
- A field entitled "Code submitted by:" which names the team member responsible for submitting the code and powerpoint slides into latte
- A **personalized paragraph entitled "My Contributions"** which describe the parts of the code you were primarily responsible for.