Course Project Proposal
ITCS-4010/5010: Cloud Computing for Data Analysis
Due: Wednesday, March 4, 2009

Your course project should be a programming project that involves the implementation of a cloud computing program using Hadoop. Your project proposal should clearly identify and state the topic you will work on for the course project. You will receive feedback on the appropriateness and scope of the project you propose. Course projects must be done individually.

Your project proposal should:

1. Outline the task you plan to work on for the project and identify the MapReduce steps required for your project.

2. Identify the data sets you plan to use for your project.

3. Describe what you expect to deliver as the final project. In describing your project, state what you will definitely accomplish, what you are likely to accomplish, and what you would ideally like to accomplish.

4. In addition to your code, the final project must include a written summary of your work and a web page illustrating your results for a few interesting examples. **Your code must run on the Linux PCs in the Woodward 335 lab.** The final course project is due before the last day of classes.

Evaluation of appropriateness depends on the cloud computing and data analysis (e.g., information retrieval) techniques you will use, and what you are likely to learn from the project. You are encouraged to discuss possible project choices with the instructor in advance, especially if you wish to propose something unusual.

Your project proposal should be a brief document, one or two pages in length, and must be written using a computer. The proposal is due on **March 4, 2009** at the beginning of class.