**CS410 Project Proposal**

1. **What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.**

Greg Harrison (gdh3)

1. **What is your free topic? Please give a detailed description. What is the task? Why is it important or interesting? What is your planned approach? What tools, systems or datasets are involved? What is the expected outcome? How are you going to evaluate your work?**

This project will perform web scraping on allrecipes.com to create a recipe recommendation system that compares the ingredients required for a recipe to a list of ingredients that a user provides, and then recommends recipes in order of popularity. Popularity will be based on the average rating value and the number of ratings. The idea is to create a tool that will recommend recipes for a user based on the ingredients they have so they can make something without having to go shopping.

The outcome will be a recipe data set containing the name of the recipe, the URL for the recipe, the list of ingredients for a recipe, the required quantity of each ingredient, the average rating of the recipe, and the number of ratings. Additionally, I will produce a tool that will compare a list of ingredients from a file with the ingredients in each recipe and produce a list of recommendations.

I will make some effort to avoid recommending duplicate/comparable recipes. This may be done by limiting the number of recipes pulled into the data set, or by checking for this when providing a list of recipes to the user.

1. **Which programming language do you plan to use?**

I will use Python for this project. To pull data from the recipes, I will use the Beautiful Soup Python library.

1. **Please justify that the workload of your topic is at least 20\*N hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.**

Building the web scraper will likely take me 10 – 15 hours. Producing a polished tool to compare the data set of recipes from All Recipes with a list of ingredients and make recipe recommendations will probably take an additional 10 hours.