# Project Progress Report: Recipe Finder

Team Tastebuddies, Net IDs: vs27, kg24, sg53

## 1) Which tasks have been completed?

- Installed Pylucene using Docker in the local environment and did a test run in EC2.
- Development of python code completed for the following:
  - Parsing and Cleansing of recipe dataset.
  - Indexing and Storing of recipe dataset.
  - Enhanced performance of indexing using Pandas.
  - While implementing the indexer, we evaluated different indexing options and field types like multivalued fields given in the Lucene library.
  - Retriever to search the recipes based on given 'Ingredient' query term.
  - While implementing the retriever, we evaluated different query parser options and analyzers to search the data using a single term or a phrase.
- Development of API's using Flask.

## 2) Which tasks are pending?

- Development of web pages using Polymer is in progress. (vs27, expected to be completed by 12/02)
- Development of the Content Based Recommender system is in progress. (kg24, expected to be completed by 12/03)
- Deployment of Pylucene in EC2 using Docker. (sg53, expected to be completed by 12/04)
- Deployment of API's and web pages in a web server in EC2. (vs27, expected to be completed by 12/05)
- Documentation and presentation. (expected to be completed by 12/10)
- Planning to do the final submission by 12/10.

## 3) Are you facing any challenges?

- In our project proposal, we mentioned that we would be using Apache Solr to index and store the data. However, we later found that Apache PyLucene is sufficient for our use case and used it for indexing and storing the dataset.
- We have faced a lot of issues in installing Pylucene in the local environment as well as EC2. As a solution to those issues, we found an open source docker based Pylucene repository that helped us in automating the installation of Pylucene.
- We are not facing any challenges as of now.