Foodie: An Enhanced Recipe Search Engine

Grace Ku, Stephanie Cheng, Kenneth Ng

Professor Hamid Tizhoosh

Department of Systems Design Engineering

**Abstract:** As people become increasingly health conscious, the need for a central repository for nutritional information grows more prominent. Our goal was to provide such a repository while seamlessly integrating into users’ day-to-day lives. To achieve this, an enhanced recipe search engine called Foodie is created. The objectives of Foodie include: (1) collecting, filtering, validating and storing recipes and nutritional information by crawling the Internet; (2) providing access to this data through an elegant and intuitive user interface; (3) displaying nutritional information about recipes in an impactful way; and (4) using social networks as means of growing the user base and improving user experience.

# 1. Introduction

The Internet is a powerful tool that houses vast amounts of a variety of information from many sources. However, it is not always easy to find relevant information from this repository. General search engines such as Google attempt to rank this information, so that it surfaces the most relevant information, but it is hard to always be aware of a user's context and higher-level intentions as they are searching. Domain-specific search engine can return more relevant results as stricter criteria can be used to filter out low quality results and allow for sophisticated queries can be executed to obtain specific datasets.

The specific user context this project focuses on is that of health-conscious users. Ideally users should be able to search and sort recipes by ingredient or by nutritional content.

Another existing solution is recipe search engines. After surveying existing implementations of recipe search engines, it was found that none provided nutritional information about recipes. Next, there are online sources for nutritional information such as the US department of Agriculture's (USDA) database of nutritional information [1]. However, it only provides nutritional information for individual ingredients with no context.

# 2. Problem Statement

Food.ie is a web-based search engine that provides reliable information about food, recipes, and nutrition based on information collected from the Internet through reliable sources. This will enable users a single point-of-access to find reliable information about food and nutrition, with the overarching goal of enabling the public to learn more about the foods we are consuming.

# 3. Goals

This project tackles this problem with four main goals in mind:

1. Collecting, filtering, validating and storing recipes and nutritional information by crawling the Internet.
2. Providing access to this data through an elegant user interface.
3. Displaying the data in a meaningful way.
4. Using social networks as a means of growing the user base.

# 4. Solution

The user flow is as follows. The user navigates to the website domain, "foodie.com". The first screen they see is referred to as "the landing page" which includes a search bar:



Figure : Landing page.

At this point the user can search for a recipe or a list of ingredients delimited by commas. An example of a recipe search would be "roast beef". An example of an ingredient search would be "asparagus, quinoa". When the user hits the search button or enter, they are brought to the search results page displayed below:

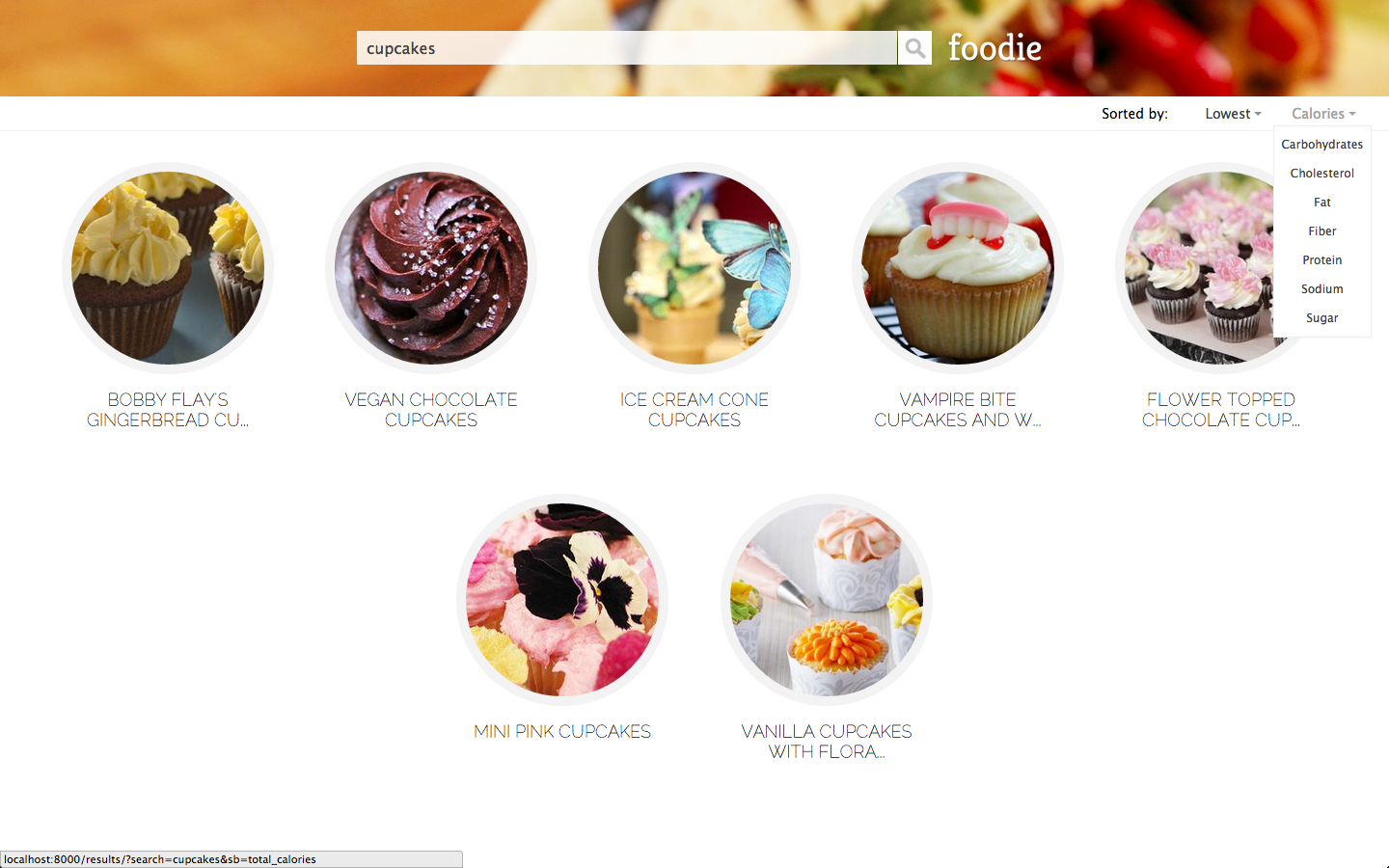


Figure : Search results page.

In the top right hand corner are dropdowns where users can sort the result set by ascending and descending values of nutritional values such as fat, carbohydrates and protein. This sophisticated querying is lacking in existing solutions. Finally, when the user selects a recipe from the list, they are brought to the recipe page below:

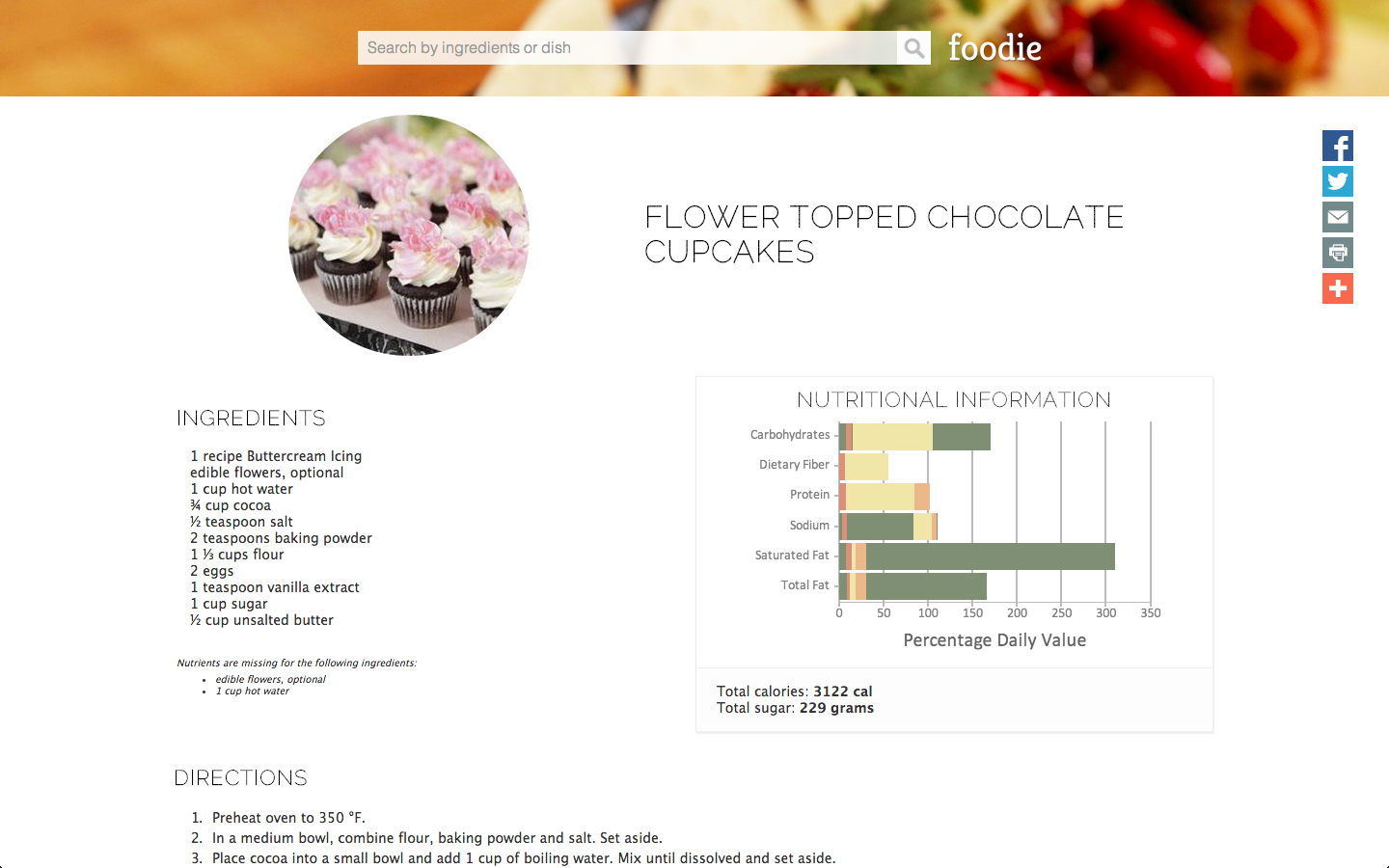


Figure : Recipe page.

On this page are the typical recipe fields you would find on any recipe page. That is, there is the recipe title, a recipe image, a list of ingredients, and the directions. In addition to this is a visualization of the nutritional value of the individual ingredients of a recipe aggregated into one graph. The y-axis are the nutrition labels and the x-axis is the percent daily value. If the user hovers over the graph, more detailed information about each ingredient is provided.

Social media integration provided on the website encourage growth in the user base of Foodie and encourages collaboration and sharing amongst users.

# 6. Technical Challenges

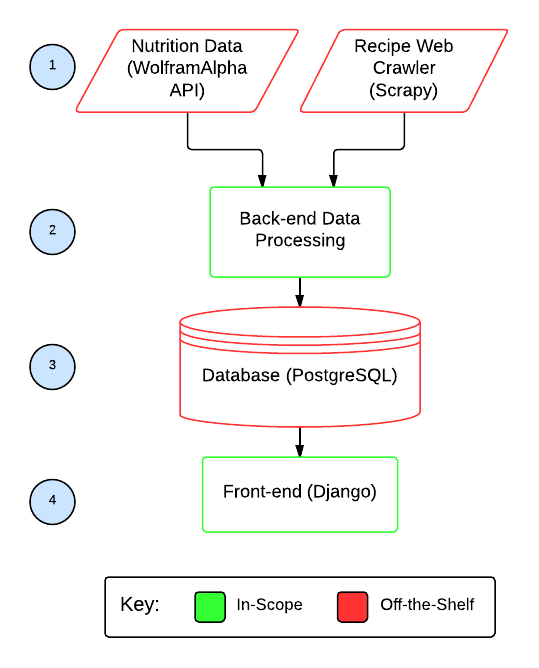


Figure : Technical solution diagram

Above is a technical diagram of the solution. It is decided to use Scrapy for web crawling, the WolframAlpha application programming interface (API) to retrieve nutritional information, the PostgresSQL database to store the information, and the Django web application framework to create the frontend [3][4][5].

Three key technical challenges of this project are: (1) speed, (2) scalability, and (3) data processing data from a variety of sources.

## 6.1 Speed

Because of the solution's reliance on WolframAlpha, a third party, there exists a bottleneck. The results depend on the speed of WolframAlpha (which interprets ingredient input such as "1 cup of sugar" and returns nutritional information). Currently, foodie relies on the natural language processing of WolframAlpha in order to interpret ingredient information. This bottleneck can be overcome by crawling government databases such as the USDA directly. A custom natural language processing engine for interpreting ingredient information would need to be created.

## 6.2 Scalability

Efficiently storing and accessing of high volumes of data can be improved by exploring different types of databases such as graph databases or document-based databases which lend themselves for faster querying and better scaling.

## 6.3 Data Processing

Processing data from a variety of sources proves difficult due to varying formats of ingredients and recipes. This ties back into developing a natural language processing engine that interprets inputs of various formats. For example, the solution should be able to handle scraped ingredients such as "1 cup of sugar" as well as "a cup of sugar". Natural language processing techniques are beyond the scope of this project, but can be employed to tackle this project.

# 7. Impact

The target users for Foodie are health-conscious consumers who want to know more about the nutritional content of their food. Additionally, consumers with dietary restrictions can benefit from the ability to filter out certain ingredients such as dairy products (for those that are lactose intolerant) or meat products (for those that are vegetarian).

The goal of this project is to make a social impact in terms of food awareness. Economically, Foodie also encourages users to cook at home to save money.

# 8. Conclusions

On a high level, Foodie explores two engineering design fields: software engineering and interface design. Software engineering techniques are used to efficiently aggregate existing, freely available information (recipes and nutritional information). Interface design techniques are used provide an interface that enables users to query this data and to obtain useful representations of this data. The end result is an enhanced search engine that allows users to search by ingredient and sort by nutrition, capabilities that are not offered by existing solutions.

# References

[1] USDA National Nutrient Database for Standard Reference, [online], http://ndb.nal.usda.gov/ (Accessed 03 April 2014)

[3] Scrapy, [online], http://scrapy.org/ (Accessed 03 April 2014)

[4] PostgresSQL, [online], http://www.postgresql.org/ (Accessed 03 April 2014)

[5] Django, [online], https://www.djangoproject.com/ (Accessed 03 April 2014)