

# CS 4395: Web Crawler Report

## 1. How the knowledge base was created

I wanted my web crawler to create a knowledge base around the topic of cakes and baking from my favorite chef, Preppy Kitchen's John Kanell. The first step was to find a suitable starting link that would be able to find URLs from outside the Preppy Kitchen domain while also being able to crawl for relevant recipe blog posts. This ended up being the main page for cake recipes on his blog, which would lead the crawler to Kanell's "About Me" page that has links to his interviews and cake demos on various news sites as well as varying blog posts from his own domain about cake.

I extracted the first 15 links found by the web crawler function, then did the appropriate processing with files to extract the main text from each of the blog posts/news articles. With the main text from each file (corresponding to a unique webpage), I could word tokenize and create a term frequency dictionary for all of the files. I ran the program and printed the top 30 terms, then decided on the terms that seemed the most relevant to me. Those terms were put into a list.

I then iterated through the list of the top 10 most relevant words and checked each of the files to extract sentences that contained those words. With the words as the keys and the sentences as the values, a knowledge base about cakes from Preppy Kitchen from various web pages was created.

## 2. Top ten terms and the corresponding knowledge base

I determined the most relevant ten terms from the term frequency dictionary to be: 'recipe', 'cake', 'flour', 'sugar', 'butter', 'egg', 'minutes', 'bowl', 'baking', and 'batter'. The knowledge base contains sentences from the crawled links that contain each of the words. For example, the word 'minutes' has the following sentences:

```
minutes
', 'bake for 30 minutes, or until a wooden pick inserted into the center comes out with just a few crumbs.
', 'let the cakes cool for 20 minutes in the pans before transferring them to a cooling rack.4.
', 'i recommend chilling the naked cake for 30 minutes before slicing.you can make the cake layers and buttercre
', 'bake for 35 to 40 minutes, then let cool in the pan for 15 minutes before inverting onto a wire rack to cool
', 'transfer the mixture to a sprayed 9-inch springform pan and press the graham cracker mixture into the bottom
', 'beat on medium-low speed until very smooth, about 2 minutes.3.
', 'boil for 5 minutes.
', 'all you need are some simple ingredients you probably already have in your kitchen, and after 45 minutes in
', 'let cool for 10 minutes.
```

The knowledge base for the other terms is below:

```
egg
', 'room temperature butter should make a slight indent when pressed with your finger.eggs – this cake calls for four eggs.

', 'with the mixer running, add the eggs, one at a time, before adding the vanilla.3.

', 'add the vanilla then eggs one at a time, scrape down the bowl between additions, and beat until well combined and fluffy.

', 'if you haven't already, try my homemade vanilla extract recipe.eggs – these eggs help bind the filling together and give

', 'add the eggs, beat just until fully combined, and then pour the filling into the cooled crust.
```

```
batter
', 'so, it's best to use unsalted butter and then add salt to the batter.

', 'make sure you take them out ahead of time as they will incorporate into the batter better when at room temperature.vanilla

', 'divide the batter among the prepared cake pans.

', 'buttermilk – don't skip adding buttermilk to your cake batter, as it gives the cake a tangy flavor and adds fluffiness and

', 'divide the batter among two greased 8-inch cake pans.

', 'always bring the cream cheese to room temperature before using it, as cold cream cheese will lead to a lumpy batter for

', 'wrap the cheesecake in multiple layers of plastic wrap before transferring it to a freezer-safe bag if possible.to prevent

', 'baking powder helps the cobbler topping rise, so if you use expired baking powder, you'll have a dense and flat topping.
```

```
bowl
', 'you can also make this frosting with an electric hand mixer and a big bowl.2.

', 'after the first batch of powdered sugar is added it's time to scrape the bowl down, then sift in more powdered sugar and

', 'in a medium bowl, whisk together the flour, baking powder, and salt.2.

', 'in the bowl of a stand mixer fitted with the paddle attachment, beat the butter on medium speed until creamy.

', 'in the bowl of a stand mixer fitted with the whisk attachment, beat the butter, vanilla, and salt together on medium-high

', 'mix mayonnaise, pepper and cranberry sauce in a small bowl and set aside.

', 'in a medium mixing bowl, whisk together the flour, baking powder, baking soda, and salt.2.

', 'in a large mixing bowl or the bowl of a stand mixer fitted with the paddle attachment.
```

```
flour
', 'in a medium bowl, whisk together the flour, baking powder, and salt.2.

', 'with the mixer on low speed, add a third of the flour mixture, followed by half of the milk.

', 'repeat, alternating between flour and milk until both are fully combined.

', 'in a medium mixing bowl, whisk together the flour, baking powder, baking soda, and salt.2.

', 'with the mixer on low speed, add a third of the flour mixture, followed by half of the milk mixture.

', 'continue, alternating between the flour and milk, mixing just until combined.

', 'cinnamon and nutmeg give the recipe its classic cozy, warm flavor.butter – different butter brands add various amounts of

', 'in a large mixing bowl, whisk together the flour, sugar, baking powder, and salt.

', 'add the cold butter and toss to coat with flour.

', 'work the butter into the flour mixture by pressing it between your fingers until the butter pieces are the size of peas.4
```

```
recipe
[ 'home » recipeswelcome to my recipes category page.

', 'here you can find hundreds of recipes to make your next meal memorable.

', 'from my funfetti cake to a stunning potato tart to the perfect margarita, find the perfect recipe to delight your palate.thanks for stop

', 'this blog is where i share all of my kitchen creations!more about mesign up to receive recipes, recommendations and tips straight to your inbox

[ 'home » recipes » desserts » frostings » buttercream frostingpublished: april 25, 2019 · modified: aug 10, 2020 by john kanellthis post may contain affiliate links

', "it's so creamy and dreamy, you might just want to eat it plain!vanilla buttercream recipes may seem beyond simple but i get a surprising number of requests for the recipe

", 'so let's set the record straight with perhaps the most ubiquitous and variable frosting recipe there is!
```

```
baking
', 'try my vanilla cake recipe, chantilly cake recipe, or chiffon cake recipe.baking powder – always double-check that your baking powder has not expired

', 'in a medium bowl, whisk together the flour, baking powder, and salt.2.

', 'you only need read on →if you want to kick start your fall baking, then make sure these pumpkin whoopie pies are on the top of your list

', 'if you like lemon desserts, try my lemon bars recipe, lemon blueberry cake recipe, and lemon cupcakes recipe as well!leavening agents – baking powder or baking soda

', 'in a medium mixing bowl, whisk together the flour, baking powder, baking soda, and salt.2.

', 'transfer the mixture to a sprayed 9-inch springform pan and press the graham cracker mixture into the bottom and 1 inch up the sides of the pan

', 'nothing fancy!baking powder – make sure the baking powder is fresh.

', 'baking powder helps the cobbler topping rise, so if you use expired baking powder, you'll have a dense and flat topping.milk – milk adds moisture and helps the topping rise

', 'in a large mixing bowl, whisk together the flour, sugar, baking powder, and salt.

', 'pour the apple mixture into a 9×13-inch baking dish.

', 'then, slice the quartered apples length-wise into long line pieces.once the apple cobbler has reached room temperature, you can cover the top with plastic wrap and refrigerate for up to 24 hours
```

```
cake
', 'from my funfetti cake to a stunning potato tart to the perfect margarita, find the perfect recipe to delight your palate.thanks for stopping by!

', 'please read our disclosure policy.this simple and delicious vanilla buttercream frosting is perfect for decorating cakes, cupcakes, and more

', 'let me know how you enjoy this recipe; historically, i have found it to be a crowd pleaser whether decorated or not

', 'this will affect decorations you've piped and may even destabilize layer cakes, causing them to sag or collapse

[ 'home » recipes » desserts » cakes » naked cakepublished: may 21, 2022 · modified: may 21, 2022 by john kanellthis post may contain affiliate links

', 'please read our disclosure policy.a popular choice for a wedding or any celebration, this naked cake is simple and delicious

', 'a moist vanilla cake layered with sweet vanilla buttercream and topped with fresh berries, this chic and elegant cake is perfect for any occasion
```

```
sugar
', 'you can add more salt to taste if needed.confectioners sugar: not all brands are equal and some will dissolve more easily than others

', 'add your powdered sugar to the butter in batches.

', 'make sure to sift your powdered sugar, especially if you're going to be using any piping tips.

', 'after the first batch of powdered sugar is added it's time to scrape the bowl down, then sift in more powdered sugar

', 'make sure you take them out ahead of time as they will incorporate into the batter better when at room temperature

', 'you might also find this sugar labeled as icing sugar or confectioners' sugar.
```

```
butter
['home » recipes » desserts » frostings » buttercream frostingpublished: april 25, 2019 · modified: aug 10, 2020 by john v

', 'please read our disclosure policy.this simple and delicious vanilla buttercream frosting is perfect for frosting cakes

', "it's so creamy and dreamy, you might just want to eat it plain!vanilla buttercream recipes may seem beyond simple but

', 'let me know how you enjoy this recipe; historically, i have found it to be a crowd pleaser whether decorating birthday

', 'if you only have salted butter on hand then go ahead and use it.

', 'cream the softened butter and salt in a stand mixer for about a minute using a paddle attachment.

', 'add your powdered sugar to the butter in batches.

', 'add more cream or milk as desired to thin the buttercream out.
```

### 3. Sample chatbox dialog

A chatbot from this knowledge base could provide instructions for baking that somewhat make sense. It could pull out instructions/advice from the knowledge base and give them to the user based on keywords the user has in their message to the chatbot. If the user does not input a keyword, the bot's response could attempt to guide the user into the direction of talking about baking. The instructions wouldn't give a working recipe of course, but the conversation would make sense and be relevant to the cake/baking topic.

An example dialog could be:

- How do I make a cake?
- Try my vanilla cake recipe, chantilly cake recipe, or chiffon cake recipe.
- I want to put sugar in the cake.
- Make sure to sift your powdered sugar, especially if you're going to be using any piping tips.
- What about flour?
- In a medium mixing bowl, whisk together the flour, baking powder, baking soda, and salt.
- How many minutes does it bake for?
- Bake for 35 to 40 minutes, then let cool in the pan for 15 minutes before inverting onto a wire rack to cool completely.
- This was not a successful recipe.
- Trust me, you won't go back to store-bought once you've tried this recipe!