

`allergens_in_foods_based_on_Chat_GPT.tsv`

This is a tab-separated values file that contains the name of the food derived from Wikipedia and then the 8 allergens determined to be in each food, derived from an Open AI prompt. For each allergen column, a 0 means the allergen is not present, 1 means the allergen is present.

`Challenges_encountered_food_allergens_project.pdf`

A description of a subset of the challenges faced with the food allergens project, including multiple examples.

`Chat_GPT_food_count_in_EHR_and_Chat_GPT_ingredient_list.tsv`

This is a tab-separated values file that contains the name of the food derived from Wikipedia, the count of patients in our electronic health record (EHR) that had the food mentioned, and the list of ingredients for the food derived from an Open AI prompt. The ingredients are all listed in a single column, with each ingredient separate by a bar (|).

`Chat_GPT_Ingredients_and_counts_in_EHR.tsv`

A tab-separated values file that contains a unique list of ingredients derived from the Open AI prompt, followed by a count of how many patients in our EHR had that ingredient mentioned in their notes.

`ingredient_variations_containing_rice_from_CGPT.txt`

This is a list of unique ingredients from the ingredients provided by Open AI (Chat GPT 3.5 turbo), that contain the word 'rice'. This was done as a demonstration to show the variability in how a single ingredient (rice) was mentioned, and how it compared the list of ingredients derived from Wikipedia.

`ingredient_variations_containing_rice_from_wikipedia.txt`

This is a list of unique ingredients from the ingredients extracted from Wikipedia that contain the word 'rice'. This was done as a demonstration to show the variability in how a single ingredient (rice) was mentioned, and how it compared the list of ingredients derived from Chat GPT.

`related_references.txt`

A few references of other works that are related to the current project.

`UpSet_plot_allergens_R_code.txt`

The R code used to make the UpSet plot comparing allergens co-occurring in a food, used in the abstract submission.

`UpSet_plot_allergens_figure.png`

The UpSet plot as a `png` comparing allergens co-occurring in a food, used in the abstract submission.

`UpSet_plot_allergens_figure.pdf`

The UpSet plot as a `pdf` comparing allergens co-occurring in a food, used in the abstract submission.

`Wikipedia_foods_project_methods_details.pdf`

Contains additional details about the methodology used in this project, including the specific Wikipedia data file used, and the prompts used for interacting with Chat GPT 3.5 Turbo.

`wikipedia_foods_title_and_name.txt`

A listing of the foods identified from Wikipedia that had a `main_ingredient` entry. The file contains one food per line, with the name of the food derived from the `title`, a bar (|) separator, and the name of the food derived from the `name` metadata element.

`wikipedia_foods_titles_and_ingredients.txt`

A listing of the foods derived from Wikipedia that had a `main_ingredient` entry, followed by a colon, followed by a list of the ingredients derived from Wikipedia for each food, with each ingredient separated by a bar.

`Wikipedia_unique_list_of_ingredients.txt`

A list of unique ingredients derived from all of the ingredients extracted from Wikipedia. This was done to compare to the list of unique ingredients provided by Open AI Chat GPT 3.5, to show the variability in what was listed.