

SoftDev P04

TPNG: Salt&Pepper

Roster: Tracy Ye, Princeden Hom, Sascha Gordon-Zolov, Ankita Saha

2025-04-23

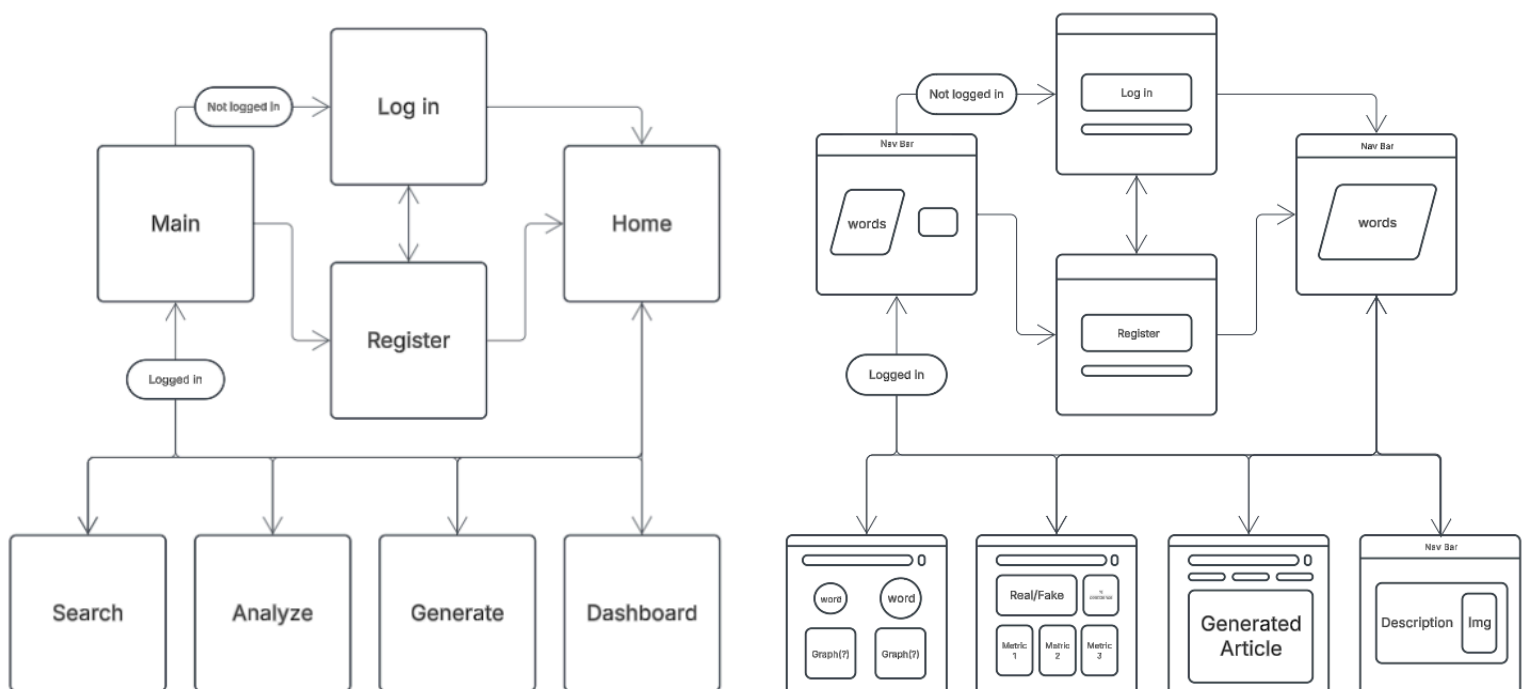
Time spent: 2 hours

Target Ship Date: 2025-04-10

Fake News Analyzer

Our website combs through a dataset of over 38,000 fake and true news articles to determine commonalities that are shared between them that may be used to predict other article's trueness or falseness. The first page users will see is a page of randomized words that vary in size and color. Each word's size will be proportional to the occurrence rate of that word throughout the dataset it was trained on while the color of the bubble will show whether the size is referring to the word's prevalence in fake-news articles or true-news articles. Once the user is logged in, they will have access to more functions, including a search page where they can find a specific word and compare its appearances in fake vs. true articles and an analysis page where the user can input a link to an article on the web and a rating of real or fake will be returned along with a judgement on the objectivity and tone of the article and keywords that helped in making the decision.

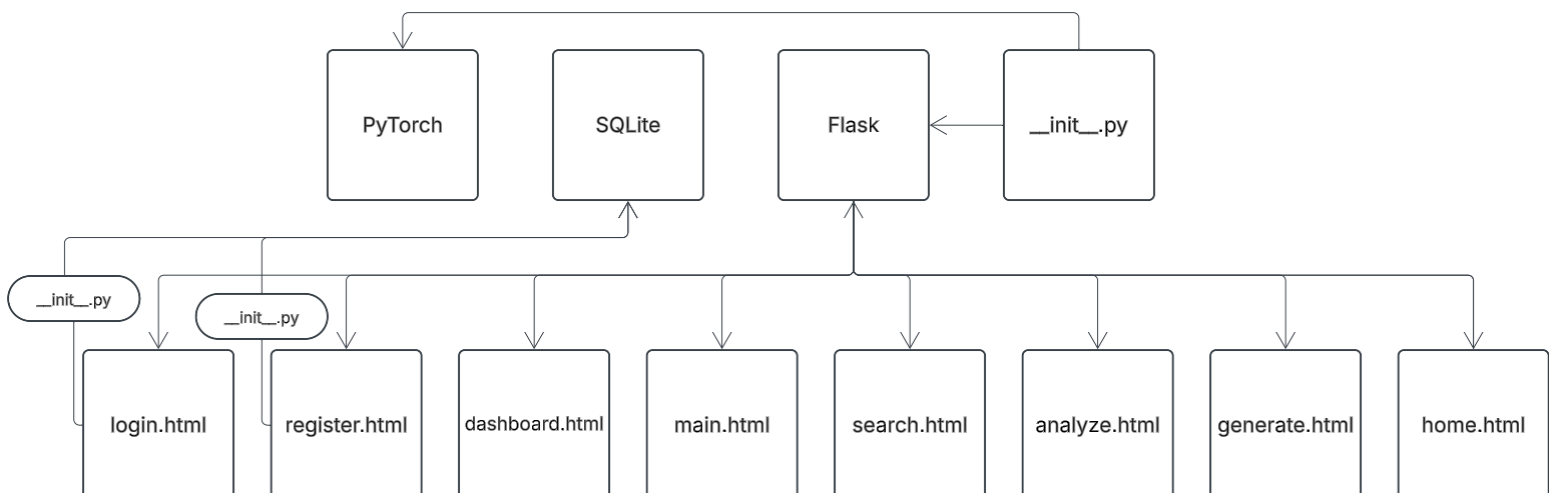
Site Map:



Components:

- `__init__.py`
 - Main python program that will run the flask application
 - Additionally will contain all data processing functions for the app and interact with the database and apis
 - Utilize sqlite to analyze database and calculate word counts
 - Utilize pytorch to create a fake news classifier
- Register.html
 - Allows user to register a new account
- Login.html
 - Allows user to log into a preexisting account
- Home.html
 - Displays D3 diagram, but larger
- Main.html
 - Has redirection to login and register
 - Display D3 diagram with most popular fake and real news words with size being proportional to prevalence
- Dashboard.html
 - Contains information about what the website is
- Generate.html (STRETCH)
 - If we have time we plan to create an algorithm that will be able to generate text that replicates the tone and structure of fake news
- Analyze.html
 - Allows users to upload txt files and will tell the user whether the text is more similar to fake or real news
 - Allows users to input an article and returns whether it is true or fake news
 - Returns the article's objectivity and keywords that were used in the decision making
- Search.html
 - Allows user to input words in a search bar to see the prevalence of that word in the fake and real news articles

Component Map:



Database Organization: SQLite3

user

- Username (text)
- Password (text)

fake_words

- Word (text)
- Count (integer)
- Color (text)

true_words

- Word (text)
- Count (integer)
- Color (text)

Datasets

- [Fake News detection dataset](#)
 - Separated in two different CSV files of fake or real news
 - Contains a title, date, text, and subject feature for each article in a collection
 - It can be used to get random words to see how frequently they appear in fake or real news.

APIs

- TEXT2DATA API (returns the probable “sentiment” behind a statement)
 - <https://text2data.com/Integration>
 - identify the primary emotional opinion within the text—to determine a writer's attitude as positive, negative, or neutral
 - Use to help determine whether an article is real or fake/biased/unbiased, etc.

Data Visualization Library

We will utilize d3 due to its ability to display data in a unique variety of ways beyond just charts, especially since we plan on using custom shapes like bubbles to display data

Front End Framework: Bootstrap:

We decided to use this because it will fit best for our project components, especially those that require specific stylistic changes (color, border, layout, etc.), and is easily understandable.

- Change color of bubble text or words to differentiate fake news vs real news
- Create layout for data visualizations
- Create formatting for articles (grid structure)
- Create a dropdown for different webpages

Task Breakdown:

Tracy - (PM)

- SQLite DB
- API
- Routing
- Droplet

Princeton

- Utilize pytorch to create a fake vs real news classifier
- Research our stretch goal of a fake news generator

Sascha

- Create search function for fake vs real news terms
- Write a function to assign a truth value to a given portion of text
- Research our potential stretch goal of a fake news generator
- CSS

Ankita

- Use code to comb through necessary sets of data and remove null, duplicate, etc. values and neaten to work smoothly for our project
- Use D3 to visualize data for fake news vs. real news and show growing trends, statistics, probabilities
- Use Bootstrap on finished components to make the site aesthetically pleasing