

Reddit Sentiment Analysis Web Application

By Mantas Macionis | SETU Carlow

Introduction

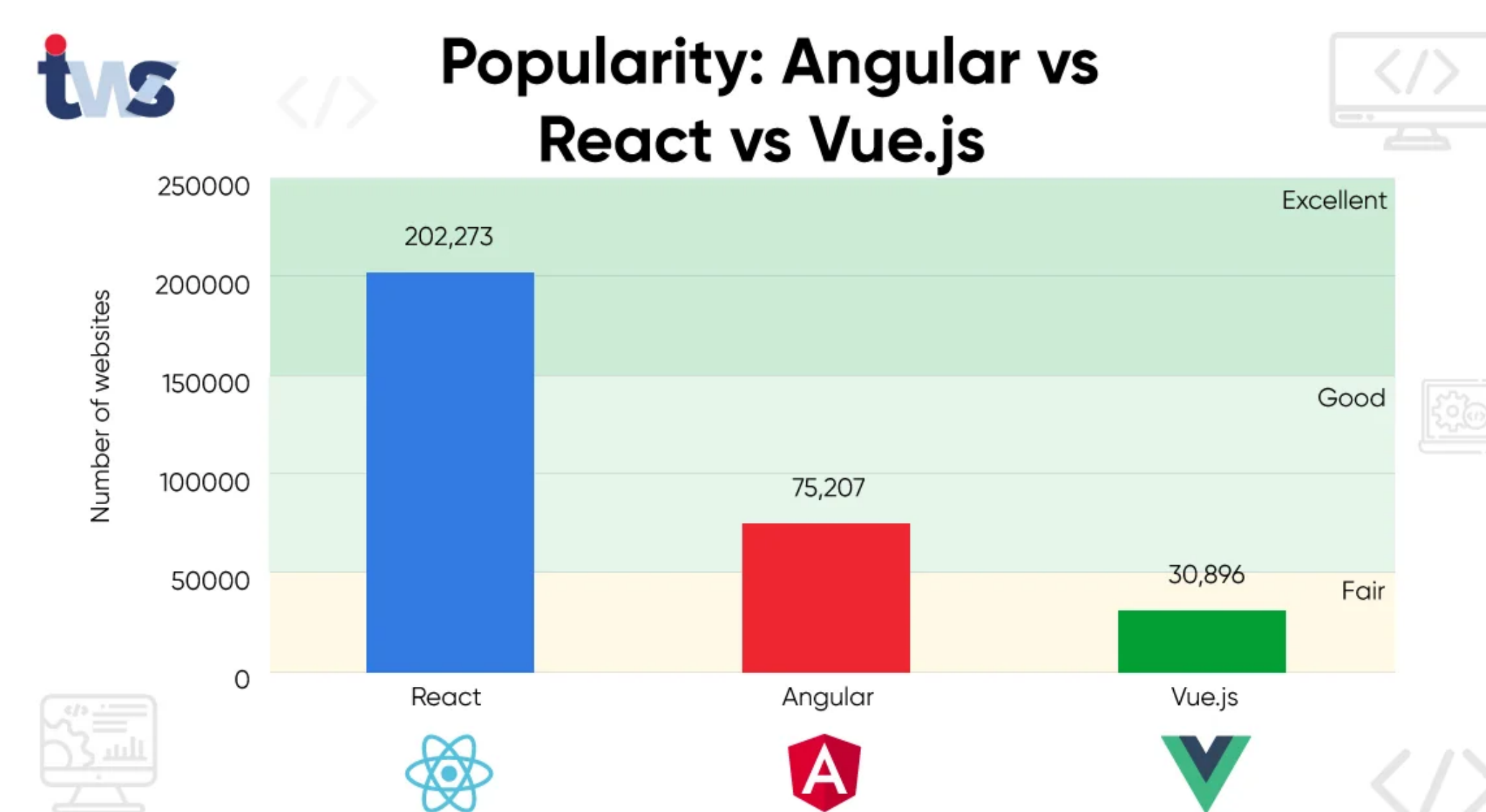
This poster will cover research conducted in relation to the Reddit sentiment analysis web Application.

The purpose of this web application is to give users a simple user interface, where they can specify a board or general search, and a term which they would like to receive a sentiment analysis on.

To create a project of this nature, multiple technological aspects must be considered. These include:

- Frontend.
- Backend.
- Database.
- AI Integration
- Reddit API Integration

[1]



Frontend Technologies

Frontend tech, responsible for the part of the application which users interact with. Including important elements of the user experience such as buttons, text fields and other visual elements.

For this section the main options considered were React.js, Vue.js and angular. Key characteristics include:

- **React.js:** Advantages - Large resources available due to large userbase, Component based Architecture. Disadvantages – Steep initial learning curve, Component based architecture being harder to keep track of with scaling
- **Vue.js:** Advantages – Ease of use, basic syntax, good performance due to virtual DOM(Document object model). Disadvantages – Smallest ecosystem, Limited scalability.
- **Angular:** Advantages – Powerful features, Scalability, Support for web dev practices. Disadvantages – Steep learning curve, Limited flexibility.
- With the following aspects of each main frontend technology considered, Vue.js may be the best option to use for this project, Due to time constraints and low initial scale. Good performance without the steep learning curve in other options is also a major advantage, allowing for a highly responsive site.

[2]

Backend Technology

The backend technology to be chosen for this project depended on aspects such as: The scale of the project, Integration ability with other tech and customization abilities.

The backend tech which encompasses these aspects best is **FLASK**. Flask is a Python web framework which is well suited for building the backend of web apps. Flask can handle tasks such as processing data, interacting with databases and managing server side logic.

Another key aspect of Flask is the use of the python programming language, which will be beneficial in interactions with the reddit API.[3]

Databases

A large set of potential database types could be suitable for this project. Two examples include: PostgreSQL and MongoDB

PostgreSQL is an open source relational database management system, its relational structure makes it suitable for storing data related to user profiles, sentiment analysis results and other relevant information.

MongoDB is an open source NoSQL database management system, it is designed to handle unstructured data in a flexible, scalable and horizontally scalable manner.

AI Integration

The ChatGPT API was chosen for the AI sentiment analysis aspect of this web application.

The ChatGPT API is highly tested and popular, meaning a lot of documentation is available online on how to implement it in projects.

For this project, The GPT 3.5 turbo API has been chosen due to cost efficiency, GPT 3.5 doesn't have the ability to go to a specified site and view the sentiment as live web browsing is not available, instead, the API will be given the same text the traditional sentiment analysis algorithms are, potentially less pre processing will be done, e.g not removing stop words, to test if the GPT API may benefit in accuracy from the extra context which this may bring. [4]

For the traditional machine learning algorithms, the best choices are **Tensorflow** and **scikitlearn**.

TensorFlow is an open source machine learning framework developed by the google brain team, in this project it can be utilised for sentiment analysis model training, Integration with NLP(Natural language processing) Tools and Neural Network based sentiment analysis.

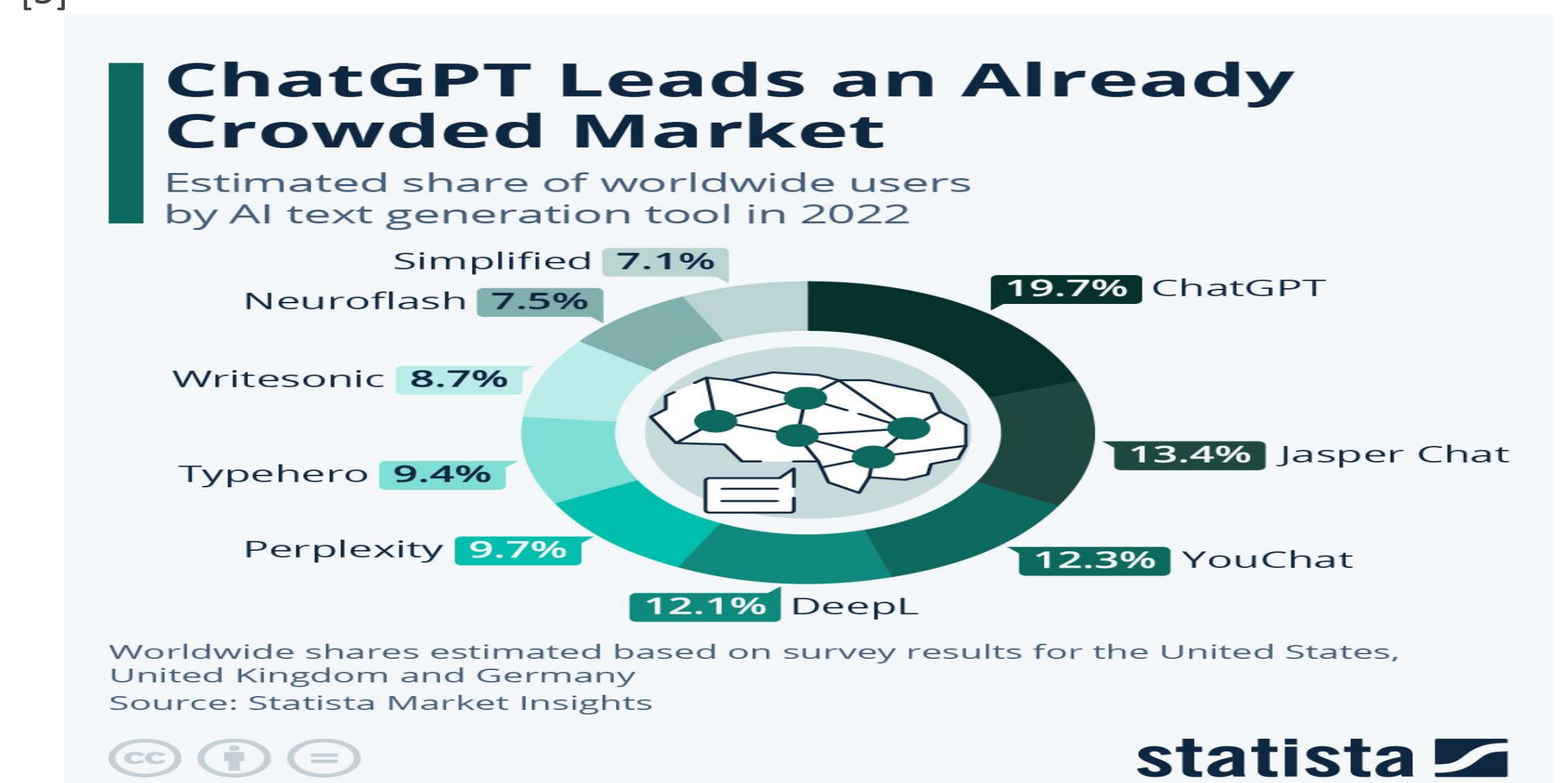
Scikit learn does not involve complex neural networks or deep learning but does allow for traditional machine learning tasks, it is useful for simpler models and algorithms. Scikit learn may also be easier to use than TensorFlow, with a large user community and documentation available online, development speed can be faster.

Reddit API Integration

For integration of the reddit API, the PRAW(Python reddit wrapper) will be used. This will allow for ease of use as Flask is also being used for this projects back end.

PRAW is designed to be user friendly, and simplifies the process of making requests to the reddit API. It allows for convenient data retrieval, Authentication handling and Rate limit management.

PRAW is also open source with a large community and documentation available online, speeding up the development process. [5]

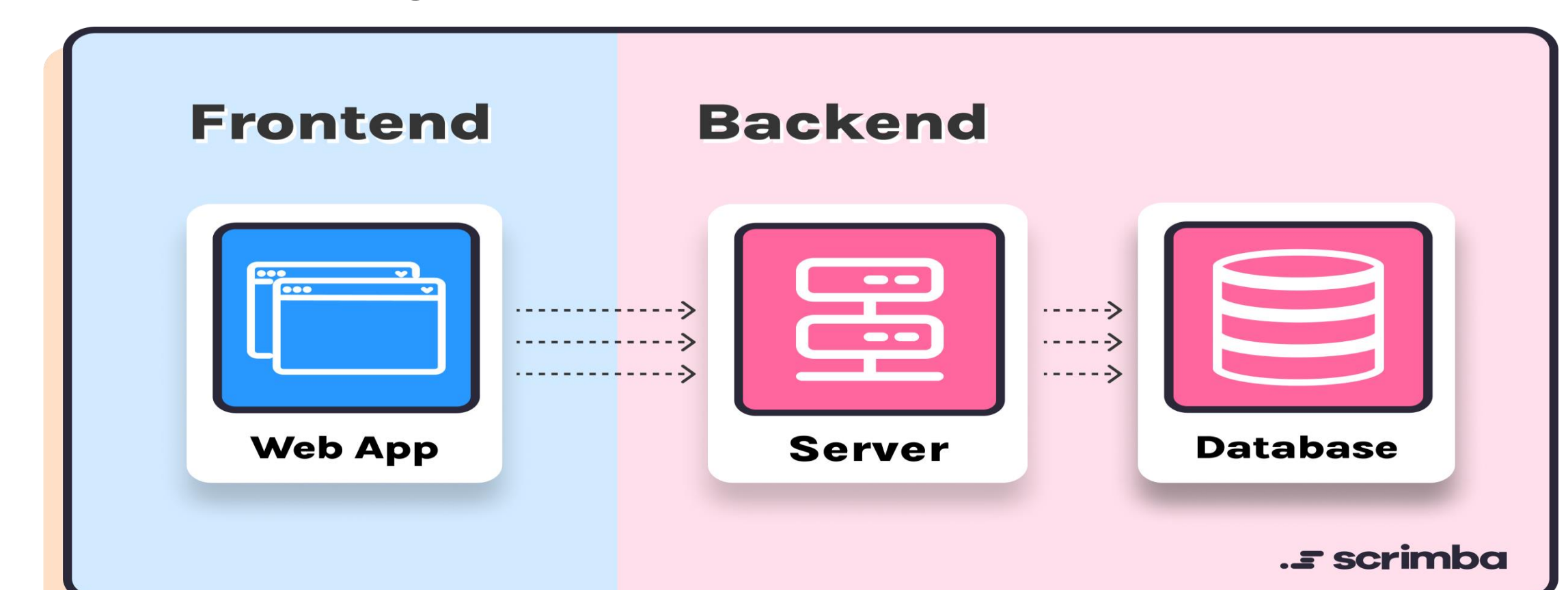


Summary

The technologies used within this project have been chosen for their reliability, compatibility and ease of use, these will be some of the most important factors in the development of this project due to time and monetary constraints.

- Frontend : **Vue, React**
- Backend: **Flask**
- Database: **PostgreSQL**
- AI Integration: **ChatGPT API**
- Reddit API integration: **PRAW**

[6]



References

- [1] https://www.reddit.com/r/angularjs/comments/ik04cp/angular_vs_react_vs_vue_a_complete_comparison/
- [2] <https://www.browserstack.com/guide/angular-vs-react-vs-vue>
- [3] <https://www.geeksforgeeks.org/python-praw-python-reddit-api-wrapper/>
- [4] <https://openai.com/pricing>
- [5] <https://www.statista.com/chart/31199/share-of-worldwide-ai-text-generation-tool-users/>
- [6] <https://scrimba.com/articles/frontend-backend-or-fullstack/>
- [7] Poster template: <https://www.free-power-point-templates.com/presentation-poster-templates/>