



# INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

## Whatsapp Analyzer

Rizwana S, Shivraj M, Sidhesh M, Shakthivel A, Arsh P  
SIES Graduate School of Technology, Nerul, Navi Mumbai.

**Abstract:** The proposed system is a whatsapp message analyzer that performs the users and group chat analysis. Analyzing the user behavior with the given messages will be needed in many situations whether it could for extracting the user interest, taking or picking up the point of discussion, understanding user attitude and behavior. By extracting the meaningful text from conversation, a summary of data can be generated to get the statics and summary of the conversation. With the application of machine learning algorithms, semantic analysis is done to get the meaning and summary of the messages. the system provides a valuable tool for users to gain insights into their WhatsApp chat history and understand their social interactions and communication patterns better. The project implementation is user-friendly and efficient, making it accessible to a wide range of users.

**Keyword:** whatsapp analyzer, text analysis, semantic analysis, chat summarization.

### 1 Introduction

Millions of people use the well-known instant messaging service WhatsApp globally. The app is a useful tool for both personal and professional communication because it enables users to share messages, voice notes, images, and videos with one another. With WhatsApp's rising fame, there is a growing demand for research into the conversations that take place on the app. A software programme called WhatsApp Analyzer employs data analysis methods to glean insights from WhatsApp conversations. By extracting the meaningful text from conversation, a summary of data can be generated to get the statics and summary of the conversation. These summary and extractions can be useful for analyzing the behaviors of a user or a group of users in the conversations. These analysis is useful for many applications. The variety of applications include recalling of chat summary, any key points discussed in the chat and focused point of interest etc. which can be useful for social as well as security concerns.

WhatsApp analyzers can use various data analysis techniques such as natural language processing, sentiment analysis, and network analysis to extract insights from WhatsApp conversations. These techniques allow WhatsApp analyzer to identify key topics, track the sentiment of conversations, and visualize the social network of users involved in a conversation.

### 2 Literature Review

There are other similar systems available that can analyze WhatsApp chat data and provide insights. Some examples of such systems [1-6] are discussed as follows.

**Chat Visualizer:** Chat Visualizer is a web-based tool that allows users to upload their WhatsApp chat history and analyze it to gain insights. The tool generates various visualizations, such as word clouds, message frequency charts, and sentiment analysis graphs.

**WAToolkit:** WAToolkit is an open-source project that allows users to analyze their WhatsApp chat data using Python. The project includes various scripts that can extract data from a chat file and generate visualizations.

**Chatistics:** Chatistics is a Python package that provides tools for analyzing WhatsApp chat data. It can extract data from a chat file and generate visualizations such as word clouds, message frequency graphs, and more.

These systems use similar techniques and libraries to analyze WhatsApp chat data, such as Pandas, Matplotlib, and Seaborn. However, the specific insights and visualizations that they provide may differ based on their implementation and the preferences of their users.

### 3 Proposed System

The proposed system consists of various phases of implementation.

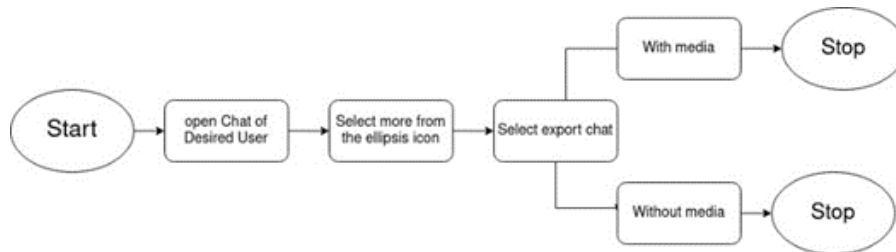


Fig.1 Flowchart of steps to be followed

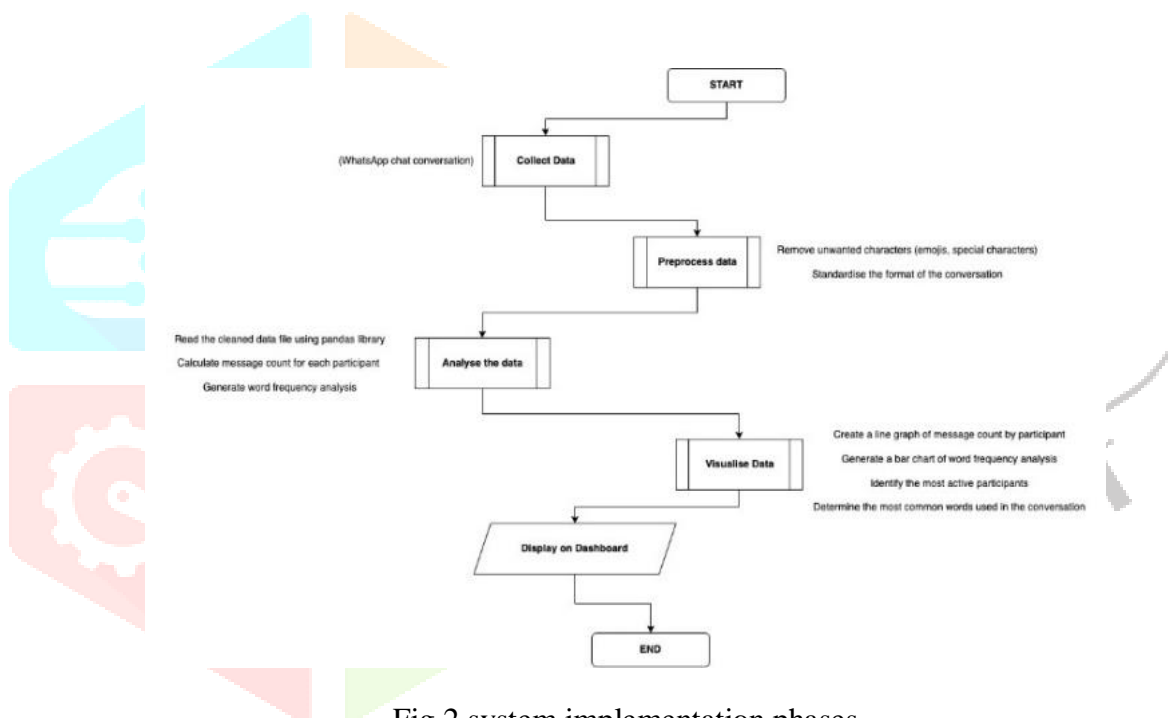


Fig.2 system implementation phases

**Data Collection:** To begin the analysis, the first step is to collect the data. In this project, the data is in the form of a text file that contains the WhatsApp chat conversation. To export the chat conversation from the WhatsApp application, follow these steps:

Open the chat you want to export in WhatsApp.

Tap the three-dot menu button in the upper right corner and select "More."

Select "Export Chat" from the menu.

Choose whether to include media or not, and then select "Email Chat" or "Export Without Media" to save the chat as a text file.

**Data Preprocessing:** After the data has been collected, it needs to be preprocessed to make it usable for analysis. The preprocessing step involves cleaning and formatting the data to remove unwanted characters, such as emojis and special characters, and standardise the format of the chat conversation. This step is necessary to ensure that the data can be analysed accurately and effectively. In the WhatsApp Chat Analysis

project, the preprocessing step is done using Python's regular expressions (regex) module to extract and clean the chat conversation.

**Data Analysis:** Once the data has been preprocessed, the next step is to analyse the data using Python. The sample code provided in the repository uses the pandas library to read the text file into a DataFrame, which is a two-dimensional table-like data structure. The DataFrame can then be used to perform various analyses, such as calculating message counts, generating word clouds, and creating visualisations. Some of the common types of analyses that can be performed on the chat conversation include:

**Message count analysis:** This involves counting the number of messages sent by each participant in the chat conversation. It can help identify the most active participants in the conversation.

**Word frequency analysis:** This involves counting the frequency of words used in the chat conversation. It can help identify the most common words used in the conversation.

**Sentiment analysis:** This involves analysing the tone and sentiment of the chat conversation. It can help identify the overall mood and sentiment of the conversation.

**Data Visualization:** Once the analysis is complete, the results can be visualised using Python's matplotlib library. The sample code in the repository provides examples of how to create various types of visualisations, such as line graphs, bar charts, and scatter plots. The visualisations can be customised to fit the specific needs of the analysis and to help communicate the results of the analysis to others. Data visualisation is an important step in the analysis process because it allows for easy interpretation and communication of the results.

**Interpretation:** The final step in the analysis process is to interpret the results. This involves looking at the data and the visualisations to gain insights into the chat conversation. For example, the analysis could reveal which participants in the chat were the most active, which words were used most frequently, or which times of day were the most active. These insights can then be used to make decisions or draw conclusions about the chat conversation. The interpretation step is important because it allows the analysis to be used for decision-making or to gain a deeper understanding of the chat conversation.

## 4 SYSTEM IMPLIMENTATION

### 1-Data Collection

- a) Open chat in WhatsApp
- b) Tap three-dot menu button and select "More"
- c) Select "Export Chat"
- d) Choose to include media or not
- e) Select "Email Chat" or "Export Without Media" to save as text file

### 2-Data Preprocessing

- a) Use Python's regex module to extract and clean chat conversation
- b) Remove unwanted characters and standardize format

### 3-Data Analysis

- a) Use pandas library to read text file into a DataFrame
- b) Perform message count, word frequency, and sentiment analysis

### 4-Data Visualization

- a) Use matplotlib library to create visualizations, such as line graphs, bar charts, and scatter plots

### 5-Interpretation

- a) Analyze the results of the analysis to gain insights into the chat conversation
- b) Use insights for decision-making or deeper understanding.



WhatsApp chat history and understand their social interactions and communication patterns better. The project implementation is user-friendly and efficient, making it accessible to a wide range of users.

The proposed web development platform has a lot of potential for future growth and expansion. Some possible future directions for the platform include:

Integration with other web development tools and platforms, such as content management systems (CMS) or web hosting services, to provide a more seamless and comprehensive learning experience.

Addition of more advanced and specialized projects, such as e-commerce websites, web applications, and mobile-responsive designs, to help learners develop more advanced skills and stay up-to-date with the latest trends in web development.

Incorporation of social learning features, such as discussion forums or peer review, to facilitate collaboration and community-building among learners.

Development of a certification program or badge system to recognize and reward learners for their achievements and skills.

Expansion of the platform to include other programming languages and frameworks, such as JavaScript, React, or Angular, to cater to a broader audience of learners.

By pursuing these and other future directions, the proposed web development platform has the potential to become a leading resource for learners looking to build their web development skills and stay up-to-date with the latest trends and technologies in this field.

## References

1. Elbagir, S. and Yang, J., 2020. Sentiment analysis on Twitter with Python's natural language toolkit and VADER sentiment analyzer. In IAENG Transactions on Engineering Sciences: Special Issue for the International Association of Engineers Conferences 2019
2. Le, H.T. and Tran, N.T., 2018. A Sentiment Analyzer for Informal Text in Social Media. *Journal of Science and Technology*, 131.
3. Suryawanshi, R., Rajput, A., Kokale, P. and Karve, S.S., 2020. Sentiment analyzer using machine learning. *International Research Journal of Modernization in Engineering Technology and Science*, 2(06), pp.1-12. Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: Grid Information Services for Distributed Resource Sharing. In: 10th IEEE International Symposium on High Performance Distributed Computing, pp. 181–184. IEEE Press, New York (2001)
4. Kothandan, J. and Murugesan, P., 2021. ML based social media data emotion analyzer and sentiment classifier with enriched preprocessor. *Journal of Information Technology Management*, 13(Special Issue: Big Data Analytics and Management in Internet of Things),.
5. Garay, J., Yap, R. and Sabellano, M.J., 2019, February. An analysis on the insights of the anti-vaccine movement from social media posts using k-means clustering algorithm and VADER sentiment analyzer. In IOP Conference Series: Materials Science and Engineering (Vol. 482, No. 1, p. 012043). IOP Publishing.
6. Bandi, A. and Fellah, A., 2019, September. Socio-analyzer: A sentiment analysis using social media data. In Proceedings of 28th International Conference