A Web-Based Whatsapp Chat Analyzer Using Python and Data Visualization Techniques

Mr. C. Senthil Kumaran[1], K. Ramachandran[2]

[1]M.C.A., (Ph.D), (HOD & Associate Professor, Department of Master of Computer Applications)
[2]MCA, Christ College of Engineering and Technology
Moolakulam, Oulgaret Municipality, Puducherry – 605010

Abstract—The WhatsApp Chat Analyzer project is made to get useful information from WhatsApp chat files that people export [11]. These days people use apps like WhatsApp a lot to talk to each other for work stuff. The thing is, the information in these chats is not organized and it is hard to look at it. The WhatsApp Chat Analyzer project solves this problem by taking the text from WhatsApp chat files and turning it into information using Python-based data processing techniques [1][2]. Then it shows the results with numbers and pictures, on a website using data visualization tools [4][14].The system lets people put their chat files into it look at the messages and make some charts to understand things like when people talk, what they like to talk about and who says what. This project uses Python, Flask, SQLite and some other tools to make pictures with the data [1][6][18]. It is really easy to use. It helps you understand what is going on in your chats. The system shows you things like timelines what people are doing, which emojis are popular which words people use a lot and who is talking the most, in the chat [7][8].

We live in a time where we do a lot of things online. Messaging platforms like WhatsApp are really important for talking to people [11]. Every day we send a lot of messages on WhatsApp. These messages say a lot, about how we behave how often we talk to each other how we feel and how we interact with each other [12]. Most of us do not have a way to really understand what we are talking about on WhatsApp beyond just scrolling through our chats. The WhatsApp Chat Analyzer project wants to change this. It wants to give us a tool that can look at our chat history and show us what it means in a way that's easy to understand using data science principles [12][13].

This system helps people see how conversations change over time. It shows you who is talking a lot and what people are talking about. The project is simple and easy to use so anyone can use it even if they do not know a lot, about technology. The system is also very accurate which makes it good for everyone to use whether they know about technology or not.

I. INTRODUCTION

This project is about making a website that can look at WhatsApp chat data and understand it easily [6][11]. The main goal of this project is to take the conversations from WhatsApp and turn them into a format that's easy to work with using structured data representations [2]. We want to be able to see what the numbers say about the conversations and show how people are talking to each other.

We also want to make sure that people can log in and sign up safely so they can see their conversation analysis, without anyone else seeing it [6]. The project is designed to be flexible which means we can look at parts of the conversation separately like when people are talking where they are talking from and what words they are using [13]. This way each part of the analysis can work on its own. Ensuring ease of use, minimal user input, and clear visualization of results is also a core objective of this system [14][15].

II. APPLICATIONS

The WhatsApp Chat Analyzer is really useful in a lot of situations. People can use it to look at their group chats see how they talk to each other and find out when they are chatting the most [11].The WhatsApp Chat Analyzer can also be used by students and researchers to study how people behave what they are talking about and how they communicate with each other [7][12].

Companies and teams can use the WhatsApp Chat Analyzer to see how people work together in group chats, on WhatsApp. They can use the WhatsApp Chat Analyzer to find out who is participating and

how they are working together [9][13].The application is really useful, for learning. It can help people understand how to work with data and make websites and manage databases using Python [1][6][18]. This is because it uses Python-based technologies to show how data analysis and web development and database management work. The application is an example of how to use Python for these things.
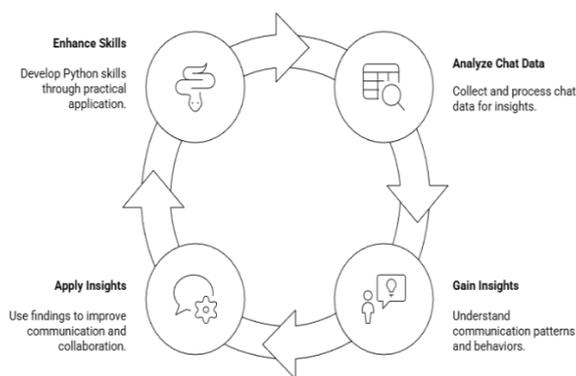
III. ANALYSIS

The WhatsApp Chat Analyzer works like this. First you upload a WhatsApp chat file to a website. This file has to be in text format as provided by WhatsApp export functionality [11]. The system checks if the file is the type. It also makes sure the chat is set up like a WhatsApp conversation. If everything looks good the file is saved on the server for a while so it can be looked at [6]. This way the WhatsApp Chat Analyzer only looks at files that will work so you do not get any mistakes when it is trying to understand the file. This means you can trust the results, from the WhatsApp Chat Analyzer.

When you upload a file the preprocessing part of the program looks at the text from the chat and uses special ways to find the important parts [2][8]. It finds things like when the messages were sent who sent them and what the messages say. The program uses codes to tell the difference between messages from the system messages from users, messages about media and links [11].It takes the information about when the messages were sent and changes it into a format so it is easy to look at the messages in the order they were sent. The program also finds out what year, month, day, hour and day of the week the messages were sent so we can look at the messages in ways based on time [2][12].

When we get the data into a table we can do some analysis to get some information [2]. The system looks at how people send messages how many people are taking part, how many videos and pictures they share which emojis they use and how many links they send [7][11]. We also look at what happens over time so we make lists of what people do every day and every month [12].This helps us see when people are using the chat the most, who is using it a lot and how people are talking to each other in the chat. The system is really good at figuring out things, like when the chat's busiest and what people like to talk about [9][13].

The data we have looked at is then turned into pictures using tools [4][14]. We make things like timelines and charts to show what is going on in a way that's easy to understand [15]. These pictures are made on the spot on the website so users can easily see how people are talking to each other and what is popular, in the WhatsApp group or personal chat [6].Overall, the working process efficiently converts unstructured WhatsApp chat data into structured information, performs statistical and temporal analysis, and presents the results through visual dashboards [12][14][15]. This systematic pipeline ensures accuracy, usability, and meaningful interpretation of chat data for analytical purposes.

IV. SCOPE

The WhatsApp Chat Analyzer is used to look at WhatsApp text files that have been exported [11]. It can handle group chats. The system looks at the text and some basic information about media like messages that have media files. It has some features like showing what happened over time who was active and when what emojis people used and what words were used the most [7][8].

The WhatsApp Chat Analyzer can only look at chat files that're offline right now. It does not work with WhatsApp in time because of privacy issues and limitations with the WhatsApp API [11]. The project is set up in a way that makes it easy to add parts to it later on like new analytical modules and better visualizations [13][14]. This means the project architecture is really flexible and we can make the visualizations of the data better.
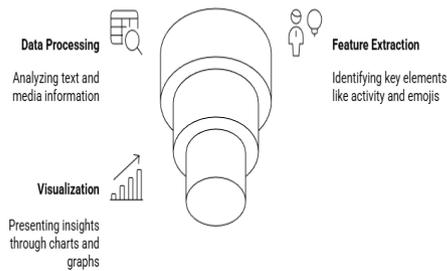
V. EFFECTS

This project shows how we can take the information we get from talking to each other and turn it into something we can really use [12]. People who use this project can see how often they talk to each other and how much they are involved in these conversations. The project teaches us that we need to clean up the information we get and look at it in a way that makes sense [2][13].

It also shows us that when we can see the information in a picture it is easier to understand [14][15]. For people who are learning this project helps them understand how to work with data make the behind the scenes parts of a website work and how to make a website do what we want it to do [1][6]. The project is really about data science and how we can use it to make sense of the communication data.

VI. BENEFITS

The WhatsApp Chat Analyzer is really useful for people who use it and for developers too. It helps save a lot of time because it can automatically look at chat histories [9]. The results are easy to see and understand, which is great [14]. You have to log in to use it so your information is kept separate, from others. It is pretty safe [6].

Developers like it because they get to work with Python, Flask, SQLite and other tools that help make charts and graphs [1][4][6][18]. The WhatsApp Chat Analyzer is also made in a way that makes it easy to fix and add things to it later on [13].

VII. DIFFICULTIES AND CHALLENGES FACED

• Handling multiple WhatsApp chat export formats due to variations in date, time, and message structure across different devices and app versions [11]

• Accurately parsing timestamps containing special Unicode characters such as non-breaking spaces and AM/PM indicators [2]

• Distinguishing between system-generated notifications and actual user messages during preprocessing [11]

• Dealing with empty or messed up chat files is a problem… to make sure we do not miss any information, from the chat files [2]

• When messages go from one month to another or from one year, to the next it is very important to get the date and time right [12]

• Handling large chat files causing performance and memory issues during processing [9][13]

• Preventing errors caused by non-string values when performing text-based analysis operations [8]

• Managing session storage limitations when handling large datasets in a web application [6]

• Generating consistent and readable visualizations for different data sizes and activity levels [14][15]

• Integrating backend data processing smoothly with frontend visualization and navigation [6][17]

VIII. CONCLUSION

The WhatsApp Chat Analyzer project does what it is supposed to do. It takes the raw WhatsApp chat data [11] and turns it into useful information that people can understand using data preprocessing, analysis and visualization techniques [2][4][14]. The WhatsApp Chat Analyzer project makes it easy for people to use a web application to look at how they behave when they chat [6]. The project shows that Python is a tool for solving real world problems with data [1][12]. Even though the project had some problems with figuring out the data and making it look good it still gives people useful results. The WhatsApp Chat Analyzer project is an example of how to use data to learn more, about WhatsApp chat behavior and serves as a foundation for further research in chat analytics [9][13].

IX. FUTURE ENHANCEMENTS

The project can get even better in the future. We can add a feature to see how people are feeling when they talk to each other. This is called sentiment analysis [7][19]. We can also make the project work with languages using NLP techniques [8]. We can give people more options to filter what they see. If we can get the tools we can even make the project work in real time. This means it can analyze conversations as

they happen. We can also make it possible to save reports as PDF or Excel files [6][17].We can also use machine learning to predict what will happen in the future [3][19]. The project can do sentiment analysis and language detection and filtering. The project can also do real-time chat analysis and export reports.

REFERENCES

[1] Python Software Foundation, *Python Language Reference*, Python.org, 2024.

[2] McKinney, W., *Data Structures for Statistical Computing in Python*, Proceedings of the 9th Python in Science Conference, 2010.

[3] Pedregosa, F. et al., *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research, 2011.

[4] Hunter, J. D., *Matplotlib: A 2D Graphics Environment*, Computing in Science & Engineering, IEEE, 2007.

[5] Van Rossum, G., and Drake, F. L., *An Introduction to Python Programming*, Network Theory Ltd., 2014.

[6] Grinberg, M., *Flask Web Development: Developing Web Applications with Python*, O'Reilly Media, 2018.

[7] Jurafsky, D., and Martin, J. H., *Speech and Language Processing*, Pearson Education, 2021.

[8] Bird, S., Klein, E., and Loper, E., *Natural Language Processing with Python*, O'Reilly Media, 2009.

[9] Aggarwal, C. C., *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, 2015.

[10] Tan, P. N., Steinbach, M., and Kumar, V., *Introduction to Data Mining*, Pearson Education, 2016.

[11] WhatsApp Inc., *WhatsApp Features and End-to-End Encryption Overview*, Official Documentation, Meta Platforms, 2024.

[12] Provost, F., and Fawcett, T., *Data Science for Business*, O'Reilly Media, 2013.

[13] Han, J., Pei, J., and Kamber, M., *Data Mining: Concepts and Techniques*, Elsevier, 2011.

[14] Few, S., *Information Dashboard Design*, O'Reilly Media, 2013.

[15] Heer, J., Bostock, M., and Ogievetsky, V., *A Tour Through the Visualization Zoo*, Communications of the ACM, 2010.

[16] W3C, *HTML5 and CSS3 Specifications*, World Wide Web Consortium, 2023.

[17] Mozilla Developer Network, *Web Technologies Documentation*, MDN Web Docs, 2024.

[18] ISO/IEC 9075, *SQL — Database Language*, International Organization for Standardization, 2016.

[19] Goodfellow, I., Bengio, Y., and Courville, A., *Deep Learning*, MIT Press, 2016.

[20] Salton, G., and McGill, M. J., *Introduction to Modern Information Retrieval*, McGraw-Hill, 1983.