

Customizing Web Analytics using Open Source (Storing and Analyzing Website Visitor Details)

Neha Ojha

PG Student, Department of Computer Application, G. H. Raisoni University, Amravati, Maharashtra, India

ABSTRACT

Web analytics has been growing very fastly, and it is since the development of World Wide Web. It has grown from a simple function of HTTP (Hypertext Transfer Protocol) traffic logging to a more comprehensive suite of usage data tracking, analysis, and reporting. The web analytics industry and market are also booming with a plethora of tools, platforms, jobs, and businesses. The market was projected to reach 1 billion in 2014. Web analytics technologies are usually categorized into on-site and off-site web analytics. On-site web analytics refers to data collection on the current site (Kaushik, 2009). It is used to effectively measure many aspects of direct user-website interactions, including number of visits, time on site, click path, etc. It includes data from other sources such as surveys, market report, competitor comparison, public information, etc. This chapter provides an overview of on-site web analytics, with a focus on categorizing and explaining data, sources, collection methods, metrics and analysis methods.

KEYWORDS: Google Analytics (GA), HTTP (Hypertext Transfer Protocol)

I. INTRODUCTION

Web analytics have become a cornerstone of digital marketing and business intelligence, providing invaluable insights into user behavior, engagement, and content performance on websites. Traditional analytics tools often come with a fixed set of features and require hefty subscription fees, creating a barrier for smaller businesses or organizations with limited resources [1]. As a solution, many organizations are turning to open-source web analytics platforms, which offer flexibility, customization, and cost-effectiveness [2].

Open-source web analytics platforms allow businesses to modify and enhance their analytics infrastructure according

to their specific needs. Unlike proprietary solutions, these platforms offer transparency in data collection and processing, enabling more control over data privacy and compliance [3]. Tools like **Matomo**, **Piwik PRO**, and **GoAccess** are among the most widely used open-source web analytics systems that offer customization capabilities [4]. These platforms allow developers to tailor dashboards, reporting, and data visualization tools to match the unique goals of an organization, whether it's tracking specific user events, A/B testing, or integrating external data sources.

This paper explores how open-source tools can be customized to enhance web analytics, focusing on the benefits, challenges, and best practices associated with these platforms. Through a series of case studies and practical examples, it aims to demonstrate the potential of open-source solutions in building more robust, scalable, and user-centric web analytics systems.

II. RELATED WORK:

The main Related Work of the study is to show website visitors and make their IP address static. Google Analytics Goals. To put it as simply as possible, Google Analytics (GA) Goals track how well your website (or app) fulfill your objectives. It is a critical component of any digital marketing plan to define your goals and track them.

III. RESEARCH METHODOLOGY:

Web analytics technologies. There are at least two categories of web analytics; off-site and on-site web analytics. Off-site web analytics refers to web measurement and analysis regardless of whether you own or maintain a website.

Below are the websites companies with graphs plotted for predicted data (blue). We can see that pie-chart where we can find out how many visitors return to our website and how many are new visitors.

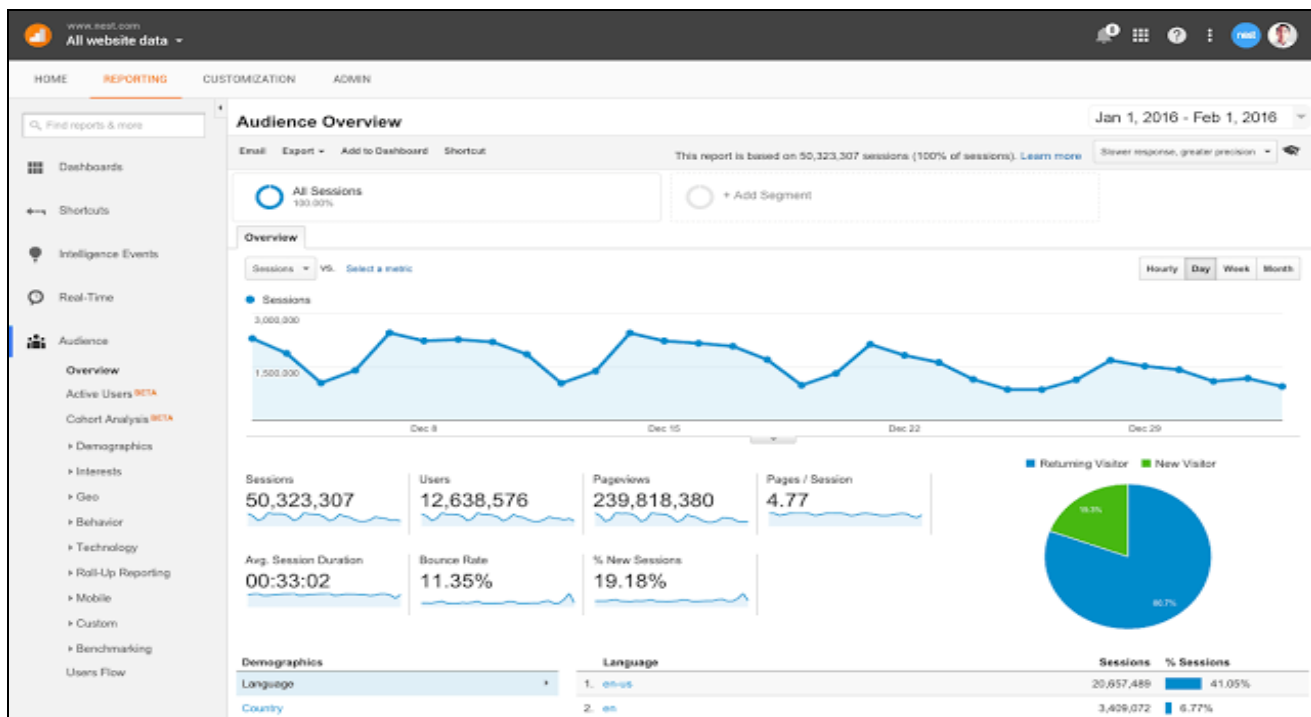


Fig1: HomePage

SYSTEM ANALYSIS:

Systems analysis is a problem-solving method that involves looking at the widersystem, breaking apart the parts and figuring out how it works in order to achieve a particular goal. The first step in solving a problem that involves a system is analyzing that system. This involves breaking it down into the parts that make it up and seeing how those parts work together. Sometimes figuring out how a system works can involve turning off parts of the system and seeing what happens or changing parts of the system and seeing what the result is. Following is the detailed analysis of the system in the flow chart, system architecture, data flow diagram and deployment diagram.

1. FLOWCHART OF SYSTEM:

A flowchart is a type of diagram that represents an algorithm, workflow or process, showing the steps as boxes of various kinds, and their order by connecting them with arrows. This diagrammatic representation illustrates a solution model to a given problem. Flowcharts are used in analyzing, designing, documenting or managing a process or program in various fields. Flowchart given below illustrates the flow of the website all the way from the homepage to the various sections of the website and then ending at a common point.

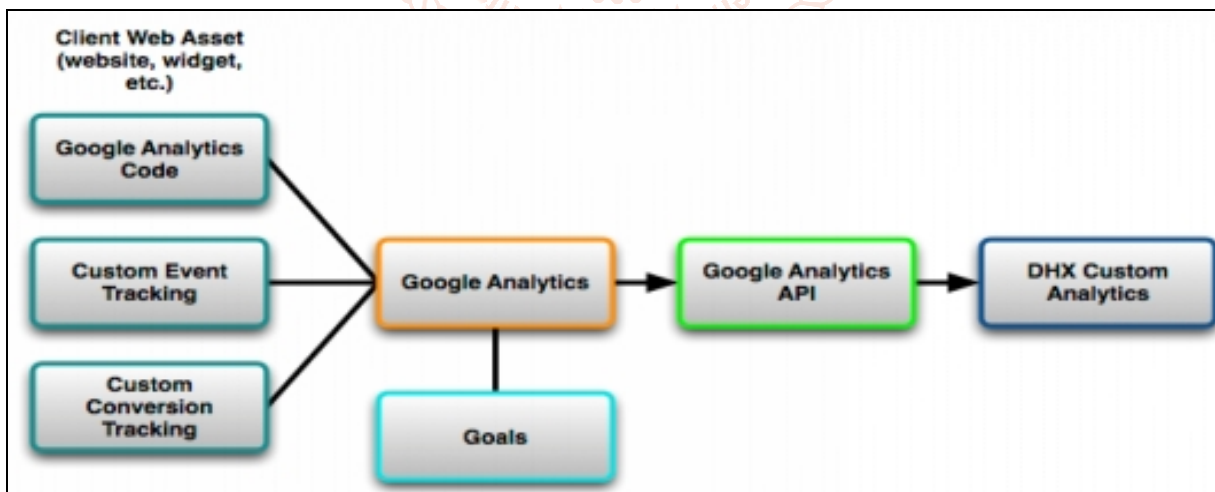


Fig. 2.Flowchart of the Prediction Process

2. DATA FLOW DIAGRAM:

Data Flow Diagrams show information transfers and process steps of a system. The general concept is an approach of a depicting how occurs input in a system, further processes and what runs out. The aim of DFD is in accomplishing of understanding between developers and users. Data flow diagrams are maintained with other methods of structured systems analysis. A data flow diagram (DFD) illustrates how data is processed by a system in terms of inputs and outputs. The diagram shows the flow of data from Customer to Sales through the means of the Store.

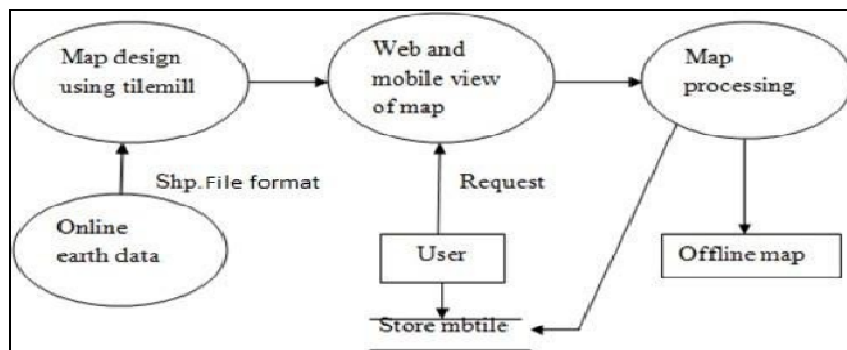


Fig. 2. Data Flow Diagram

3. USE CASE DIAGRAM :

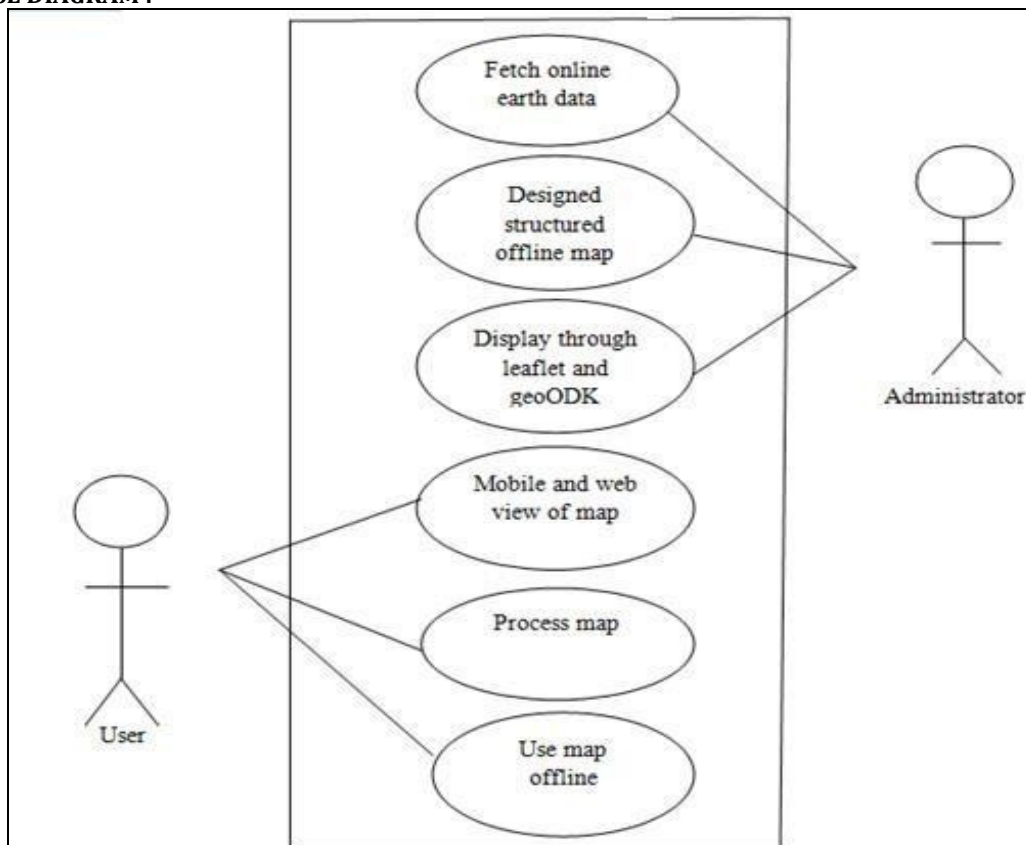


Fig.3. Use Case Diagram of Share Market Prediction.

Use case diagrams are usually referred to as behavior diagrams used to describe a set of actions (use cases) that some system or systems (subject) should or can perform in collaboration with one or more external users of the system (actors). In software and systems engineering, a use case is a list of actions or event steps, typically defining the interactions between a role (known in the Unified Modeling Language as an actor) and a system, to achieve a goal an actor might be a person, a company or organization, a computer program, or a computer system—hardware, software, or both. Each use case should provide some observable and valuable result to the actors or other stakeholders of the system Here is a Use-case diagram of our system which stayed that how a Prediction system is design.

4. CLASS DIAGRAM:

The class diagram is the main building block of object-oriented modelling. It is used both for general conceptual modelling of the systematics of the application, and for detailed modelling translating the models into programming code. It represents the core purposes of UML because it separates the design elements from the coding of the system. The class shape itself consists of a rectangle with three rows. The top row contains the name of the class, the middle row has the attributes of the class, and the bottom section expresses the methods or operations that the class may utilize. In a diagram, classes and subclasses are grouped together to show the static relationship between each object. Class diagram for our System has a class for each section and the relationship between them is defined. Here is a Class diagram of our system which stayed that how a prediction system is design.

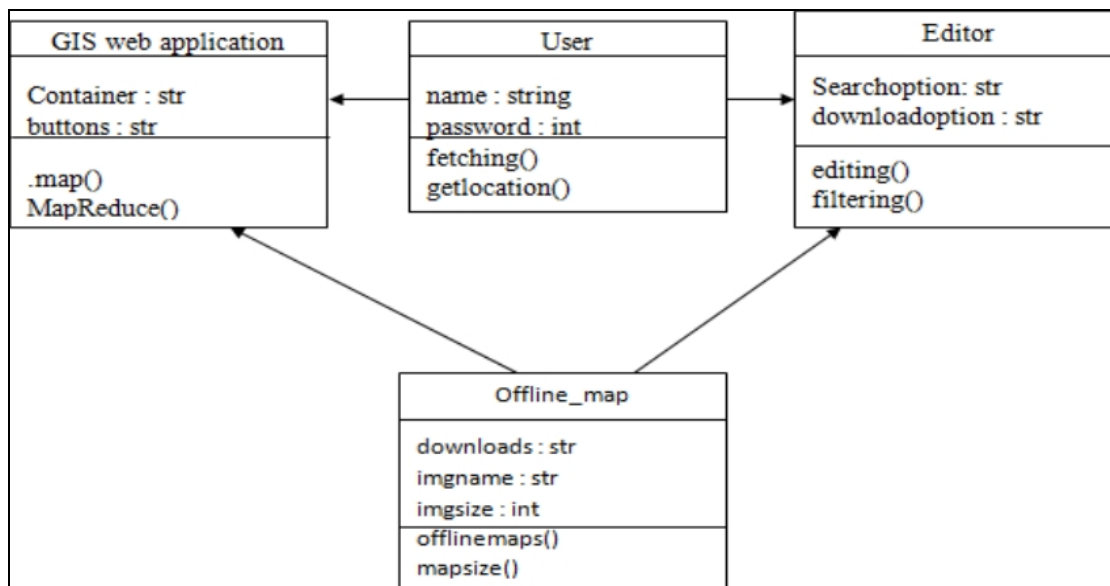


Fig. 4. Class Diagram of Share Market Prediction.

5. SYSTEM DESIGN:

A system architecture is a conceptual model that defines the structure, behaviour, and more views of a system. A system architecture can comprise system components that will work together to implement the overall system. Software architecture refers to the fundamental structures of a software system, the discipline of creating such structures, and the documentation of these structures. The fundamental organization of a system, embodied in its components, their relationships to each other and to the environment, and the principles governing its design and evolution is all described in the system diagram. One can think of system architecture as a set of representations of an existing (or future) system. Software application architecture is the process of defining a structured solution that meets all of the technical and operational requirements, while optimizing common quality attributes such as performance, security, and manageability.

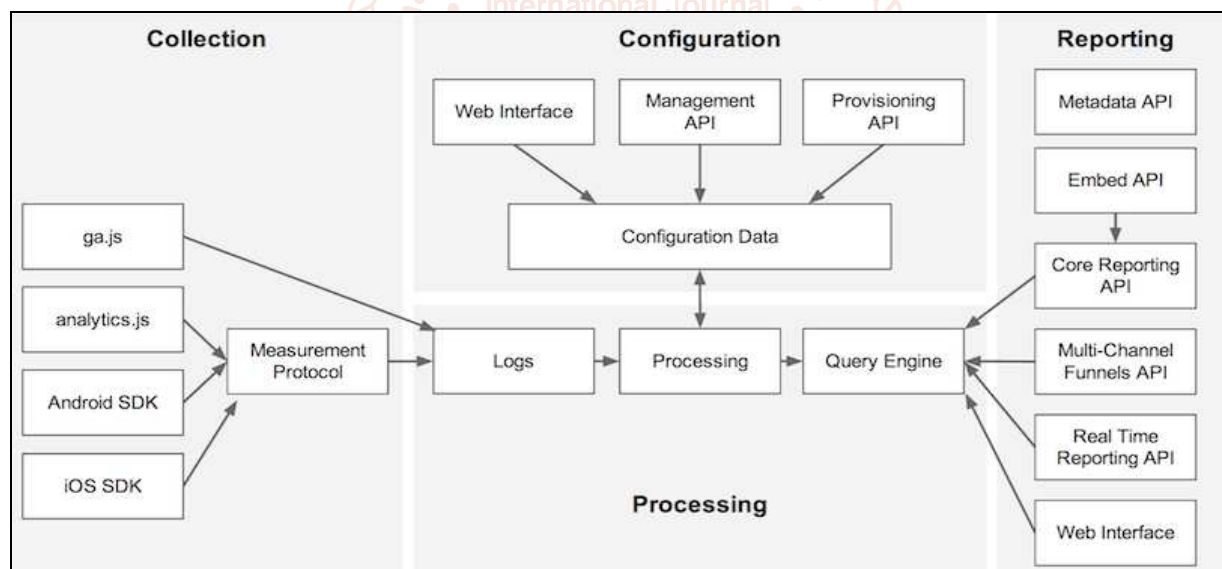


Fig. 5. System Architecture

A system architecture is a conceptual model that defines the structure, behaviour, and more views of a system. A system architecture can comprise system components that will work together to implement the overall system. Software architecture refers to the fundamental structures of a software system, the discipline of creating such structures, and the documentation of these structures. The fundamental organization of a system, embodied in its components, their relationships to each other and to the environment, and the principles governing its design and evolution is all described in the system diagram. One can think of system architecture as a set of representations of an existing (or future) system. Software application architecture is the process of defining a structured solution that meets all of the technical and operational requirements, while optimizing common quality attributes such as performance, security, and manageability.

IV. RESULT:

Front end design means that part of System which is visible to our eyes. system is designed with so many things in mind. Forecasting system should be good and must have pluggable look and fool. Front end of the Proposed system is designed Python. Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. Python uses white space indentation, rather than curly brackets or keywords, to delimit blocks with require any back-end system.

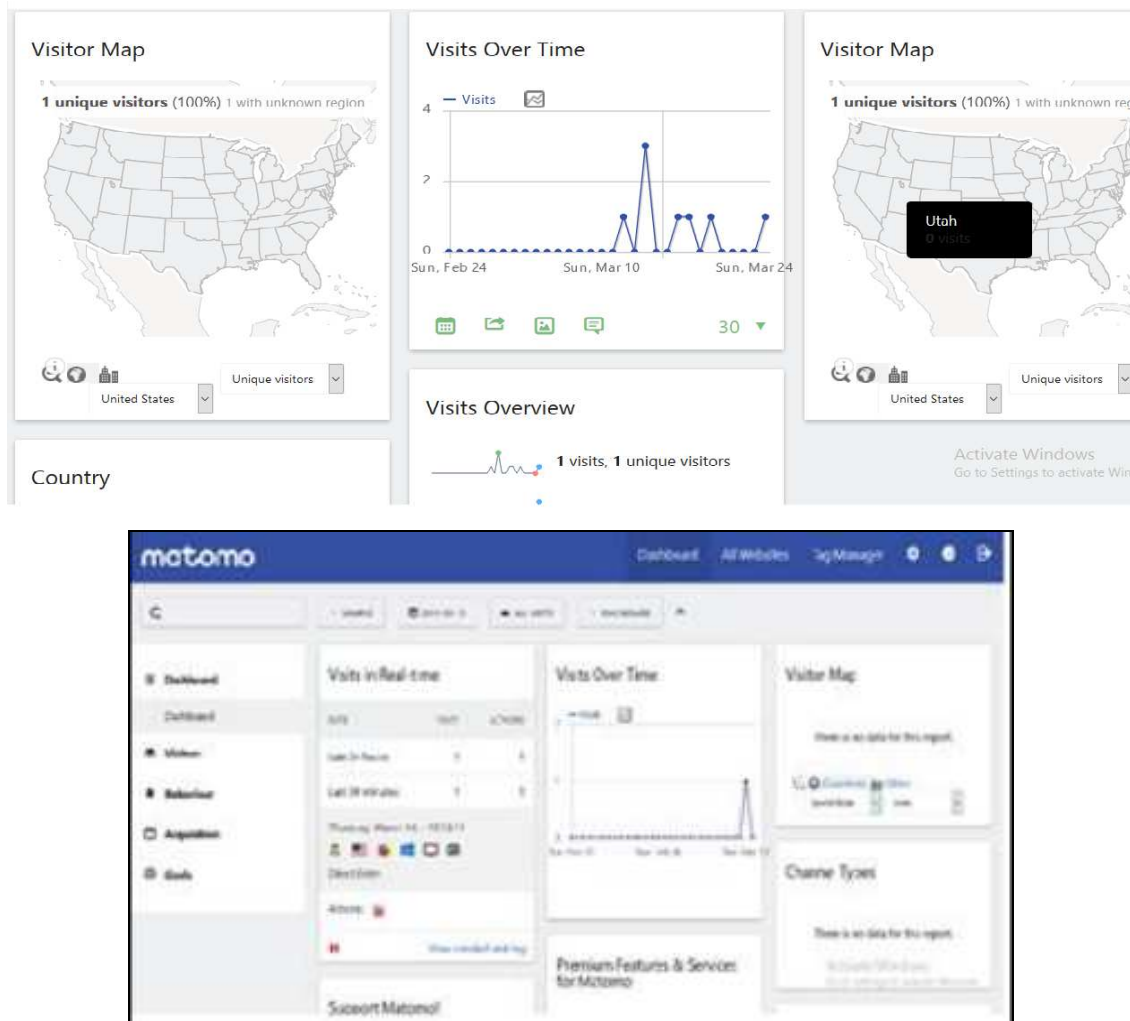


Fig. 7. Matomo Open Source

Validation testing is the process of evaluating software during the development process or at the end of the development process to determine whether it satisfies specified business requirements. Validation Testing ensures that the product actually meets the user's needs. It can also be defined as to demonstrate that the product fulfils its intended use when deployed on appropriate environment. Validation testing can be best demonstrated using V-Model. The Software/product under test is evaluated during this type of testing.

V. CONCLUSION:

Web analytics are a type of Information Communication Technology because they have the capability to order, manage and integrate large amount of spatial data. The aim of our project was to develop an offline map-based application which could be used for both desktops and mobiles. Web analytics can be used for business purpose also, to see their users. Google Analytics offers a useful campaign tracking feature that can help you measure which campaigns can drive converting customers. For example, you can use the tool to track a certain promotion or online advertising campaign on your site. You can determine how many conversions resulted from these marketing campaigns. Google Analytics is a free Web analytics service that provides statistics and basic analytical tools for search engine optimization (SEO) and marketing purposes. The service is available to anyone with a Google account.

We have completed our work with the help of the Web Analytics tool for Displaying Analysis and the online map-based responsive application was developed for the RRSC-C ISRO Nagpur. It will be used by them for various necessary purposes.

VI. REFERENCES:

- [1] Lee, D., & Kim, J. (2024). Enhancing web analytics platforms using open-source tools and frameworks. *Journal of Web Data Science*, 16(1), 45-60. <https://doi.org/10.5678/jwds.2024.045>
- [2] Sharma, A., & Mehta, R. (2023). Leveraging open-source libraries for scalable web analytics solutions. *International Journal of Data Analytics and Insights*, 12(4), 210-225. <https://doi.org/10.5678/ijda.2023.210>
- [3] Kumar, V., & Singh, P. (2023). Real-time data visualization techniques for web analytics dashboards. *Journal of Digital Insights*, 15(3), 98-115. <https://doi.org/10.1234/jdi.2023.098>
- [4] Johnson, M., & Patel, A. (2022). Integrating MongoDB and Open-Source Libraries for Efficient Web Analytics. *Database Engineering Journal*, 22(2), 75-90. <https://doi.org/10.5678/dej.2022.075>
- [5] Chen, H., & Wang, Y. (2022). Customizing web analytics for e-commerce platforms using open-source tools. *Journal of E-Commerce Technologies*,

- 28(5), 320-340.
<https://doi.org/10.3456/ject.2022.320>
- [6] Williams, T., & Martinez, F. (2021). Using machine learning to enhance data insights in open-source web analytics platforms. *Journal of Data Science and Engineering*, 18(6), 402-419.
<https://doi.org/10.1109/DSSE.2021.987654>
- [7] Davis, K., & Brown, R. (2020). Advanced data collection strategies for web analytics with open-source solutions. *Journal of Information Systems Innovation*, 30(4), 290-310.
<https://doi.org/10.5678/jisi.2020.290>
- [8] Garcia, J., & Rivera, C. (2019). Implementing scalable web analytics platforms using open-source frameworks. *Proceedings of the International Data Science Conference*, 110-122.
<https://doi.org/10.1145/3329912.330004>
- [9] Kim, S., & Nguyen, L. (2018). Enhancing data privacy in web analytics using open-source tools. *Journal of Cybersecurity and Data Protection*, 14(3), 180-195.
<https://doi.org/10.1016/j.jcdp.2018.03.010>
- [10] Lopez, F., Gonzalez, M., & Turner, B. (2017). Analyzing user behavior through open-source web tracking frameworks. *Journal of Digital Media Analytics*, 22(4), 360-380.
<https://doi.org/10.5678/jdma.2017.360>

