**Social Media Data Aggregation System**

**Moire Henry Nyakundi**

**92313**

**Submitted in partial fulfilment of the requirements of the Bachelor of Science in**

**Informatics and Computer Science at Strathmore University**

**School of Computing and Engineering Science**

**Strathmore University**

**Nairobi, Kenya**

**May 2025**

# **Declaration and Approval**

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of our knowledge and belief, the research proposal contains no material previously published or written by another person except where due reference is made in the research proposal itself.

Student Name: Moire Henry Nyakundi

Admission Number: 92313

Student Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Supervisor Name: Stephen Obonyo

Supervisor Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# **Abstract**

Content creators and influencers increasingly manage multiple social media accounts across platforms such as Instagram, TikTok, and X (formerly Twitter), making it difficult to monitor performance metrics efficiently. This project aims to develop a personal social media data aggregation tool that consolidates user analytics from different platforms into a single, unified dashboard. It addresses the gap in existing solutions, which are often business-focused, expensive, or lack meaningful cross-platform insights for individual users.

Existing tools like Hootsuite and Curator.io support content aggregation but are primarily designed for enterprise use, offering limited personalization, high costs, and minimal analytics for personal users. These limitations highlight the need for a lightweight, user-centric tool that emphasizes usability, privacy, and actionable insights.

This project proposes a web-based system built using the Object-Oriented Analysis and Design (OOAD) paradigm and developed through the Agile Scrum methodology. The solution integrates social media APIs via a Node.js backend, stores data using Firebase, and presents visual analytics through a React.js frontend. OAuth 2.0 ensures secure user authentication, while the system's modular design supports extensibility.

**Table of Contents**

[**Declaration and Approval** ii](#_Toc198716435)

[**Abstract** iii](#_Toc198716436)

[**Table of Figures** vi](#_Toc198716437)

[**Abbreviations** vii](#_Toc198716438)

[**Chapter 1: Introduction** 1](#_Toc198716439)

[**1.1 Background** 1](#_Toc198716440)

[**1.2 Problem Statement** 2](#_Toc198716441)

[**1.3 Aim/General Objective** 2](#_Toc198716442)

[**1.4 Specific Objectives** 2](#_Toc198716443)

[**1.5 Justification** 3](#_Toc198716444)

[**1.6 Scope** 4](#_Toc198716445)

[**Chapter 2: Literature Review** 5](#_Toc198716446)

[**2.1 Introduction** 5](#_Toc198716447)

[**2.2 Existing methods and tools for social media data aggregation.** 5](#_Toc198716448)

[**2.3 Challenges of multi-platform data collection, integration, and analysis.** 5](#_Toc198716449)

[**2.4 Review of Existing Systems** 5](#_Toc198716450)

[**2.4.1 Juicer.io** 5](#_Toc198716451)

[**2.4.2 Curator.io** 6](#_Toc198716452)

[**2.4.3 Hootsuite** 7](#_Toc198716453)

[**2.5 Gaps in Related Works** 7](#_Toc198716454)

[**2.6 Technologies to Be Used** 8](#_Toc198716455)

[**2.7 Conceptual Framework** 8](#_Toc198716456)

[**CHAPTER 3: METHODOLOGY** 9](#_Toc198716457)

[**3.1 Introduction** 9](#_Toc198716458)

[**3.2 Methodology** 9](#_Toc198716459)

[**3.2.1 Ideation and Requirements Engineering** 9](#_Toc198716460)

[**3.2.2 Analysis and Design** 9](#_Toc198716461)

[**3.2.3 Development** 9](#_Toc198716462)

[**3.2.4 Testing** 9](#_Toc198716463)

[**3.3 System Analysis** 10](#_Toc198716464)

[**3.3.1 Use Case Diagram** 10](#_Toc198716465)

[**3.3.2 Class Diagram** 10](#_Toc198716466)

[**3.3.3 Sequence Diagram** 10](#_Toc198716467)

[**3.3.4 Entity Relationship Diagram** 10](#_Toc198716468)

[**3.3.5 Wireframes** 10](#_Toc198716469)

[**3.4 System Design** 10](#_Toc198716470)

[**3.4.1 Wireframes** 10](#_Toc198716471)

[**3.4.2 Database Schema** 10](#_Toc198716472)

[**3.4.3 System Architecture** 10](#_Toc198716473)

[**3.5 System Development Tools and Techniques** 10](#_Toc198716474)

[**3.5.1 React.js and Node.js** 11](#_Toc198716475)

[**3.5.2 Firebase** 11](#_Toc198716476)

[**3.6 Deliverables** 11](#_Toc198716477)

[**3.6.1 Authentication Module** 11](#_Toc198716478)

[**3.6.2 User Management Module** 11](#_Toc198716479)

[**3.6.3 API Integration Module** 11](#_Toc198716480)

[**3.6.4 Analytics & Visualisation Module** 11](#_Toc198716481)

[**3.6.5 UI Module** 11](#_Toc198716482)

[**3.6.6 Documentation** 11](#_Toc198716483)

[**References** 12](#_Toc198716484)

# **Table of Figures**

[Figure 1: Juicer.io screenshot 6](#_Toc198716648)

[Figure 2: Curator.io screenshot 6](#_Toc198716649)

[Figure 3: Hootsuite screenshot 7](#_Toc198716650)

[Figure 4: Conceptual Framework 8](#_Toc198716651)

# **Abbreviations**

UI - User Interface

OAuth - Open Authorization

API - Application Programming Interface

# **Chapter 1: Introduction**

## **1.1 Background**

Social media has emerged as a vital platform for communication in the modern age. With 63.9% of the world’s population using social media and the average number of platforms used each month being 6.8 (Chaffey, 2025) there is an increasing reliance on these tools to maintain a social life and for some, generate income. Digital influencers and content creators are some of the most prominent users of social media and depend heavily on performance metrics to track their success and build their audience. However, these metrics are typically platform specific, requiring them to log into multiple services and manually consolidate insights. This wastes valuable time and can lead to human error.

Many people rely on their social media as a source of income. Having more knowledge of various factors, such as growth over time, will help them to gain the maximum value out of these platforms. The problem arises when attempting to aggregate this information for various platforms. The existing tools often only cater to businesses, are limited to a singular platform, or are costly for an individual. The addictive nature of social media feeds and the propensity for aimless scrolling can be time-consuming and counterproductive. (bubblinn, 2021)

The problem is mainly felt by independent content creators, micro-influencers, digital marketing consultants, and social media managers. Without the insights provided by the data, the creators risk overlooking trends or losing growth momentum. This can result in lower revenue, lost interest in the brand, or even burnout. Additionally, without cross-platform comparisons, it is difficult to identify the most valuable platform or successful content style.

While some industry solutions like Hootsuite or Sprout Social attempt to offer insights, many are primarily focused on content scheduling rather than comprehensive analytics. Even those with built-in analytics often lack the deep cross-platform comparative metrics needed by creators looking to optimize and grow. Academic research on influencer marketing and social media analytics has begun to explore the potential of integrated systems (Adishree, 2024) but practical implementations remain scarce, especially solutions that cater to smaller-scale users rather than large agencies. This project proposes to address this gap by building a web-based tool that provides influencers with a centralized dashboard to monitor key performance metrics across multiple social media platforms.

## **1.2 Problem Statement**

Content creators and influencers need access to performance insights across multiple social media platforms, but most currently rely on disparate dashboards and manual tracking. The lack of a centralized solution hampers their ability to make data-driven decisions and grow their audiences effectively. As creators juggle content across Instagram, YouTube, TikTok, X, and other platforms, the absence of an integrated analytics tool creates inefficiencies and knowledge gaps. Existing aggregation tools are often expensive, tailored to fit the needs of businesses and marketing professionals, or focus on content delivery rather than retrieving data.

A lack of data deprives users of information that would be invaluable in managing their online presence. Those that manually aggregate their data must content with many disparate formats from each platform. Creators that use in-app insights waste a lot of time switching between the numerous social media sites.

Time spent switching between different platforms is time not spent elsewhere. The data is useful so a lack of it could cause a loss of revenue. The increasing number of social media platforms has created a tangible challenge for individuals striving to effectively manage and gain a comprehensive understanding of their digital footprint.

## **1.3 Aim/General Objective**

To develop a framework for aggregating social media data across multiple platforms to enhance data-driven decision-making.

## **Specific Objectives**

1. To investigate existing methods and tools for social media data aggregation.
2. To identify the challenges faced in multi-platform data collection, integration, and analysis.
3. To design a cross-platform data aggregation framework.
4. To develop a prototype system that implements the proposed aggregation framework.
5. To test the performance and usability of the developed system across selected social media platforms.

**Research Questions**

1. What are the current methods used for social media data aggregation, and what are their limitations?
2. What challenges are encountered in collecting and integrating social media data from diverse platforms?
3. How can an aggregation framework be designed and implemented?
4. What impact does the proposed aggregation system have on the efficiency of data analysis and decision-making?
5. How well does the system do in testing?

## **1.5 Justification**

Such a platform could enable simplified content management and review, allowing users to more easily review, and organize their own content that they have shared across various platforms over time (Mailchimp, 2025)They don’t need to pay. They don’t need to understand how to standardize the differently formatted data. it promises significant convenience and time efficiency by providing a single point of access to all their social media activities and content, thereby eliminating the need to repeatedly log in and navigate across numerous platforms (Adishree, 2024)The aggregation of data can facilitate gaining personal insights by offering a holistic and longitudinal view of an individual's online presence (Qlik, 2025). Finally, a personal aggregator could empower users with an enhanced sense of data ownership and control over their social media information, which is typically siloed and managed by individual platforms (bubblinn, 2021)Track their professional growth. A free tool. Focused on individual users. Personal social media aggregation emerges as a compelling solution by directly addressing the inherent fragmentation that characterizes the contemporary online landscape. As individuals distribute their presence across a growing array of platforms for various social and professional purposes, the need for a unified perspective on their digital lives becomes increasingly critical. A personal aggregator serves as a bridge across these digital silos, offering a comprehensive overview that not only enhances user understanding but also empowers them with greater control over their online activities and personal data

## **1.6 Scope**

This study focuses on designing and developing a cross-platform social media data aggregation framework. It will include analysing popular social platforms, implementing APIs, and creating a centralized aggregation and visualization tool. It will not delve into deep natural language processing or sentiment analysis models beyond basic filtering and tagging. This system will be developed using React.js for the frontend, Node.js for the backend, and Firebase for the database. OAuth will be used to handle authentication.

**Limitations**

Platform API restrictions, dynamic changes in data access policies, ethical concerns around data privacy, a lot of time must be invested in learning to use the tools to develop the system, managing the different formats that the data will be provided in,

**Delimitation**

Selecting platforms with open APIs, complying strictly with terms of service, and implementing anonymization and data security measures in the prototype system

# **Chapter 2: Literature Review**

## **2.1 Introduction**

This chapter explores the background and environment in which personal social media data aggregation takes place. It analyses the context and current practices in managing social media presence and discusses the technological foundations for this project.

## **2.2 Existing methods and tools for social media data aggregation.**

Social media aggregation, at its core, involves the systematic collection and centralized display of content originating from diverse social media channels within a unified platform. (PR G. , 2025) Currently, managing personal data across platforms is done manually or through platform-specific dashboards. This means users must log in separately to each service to retrieve the required data, often resulting in an incomplete picture of their social media activity. Additionally, each platform structures data differently and often limits access via strict API protocols (SentinelOne, 2025). This project will have to take into account such platform-specific constraints, among others, to be effective. In an ideal world, users would have unrestricted access to all their personal data but current reality has numerous legal, technical, and ethical boundaries that prevent that from being the case. This gap forms the basis of this research: the need for a user-friendly, privacy-conscious aggregator that empowers individuals to view and analyse their own data (Zanini, 2025)

## **2.3 Challenges of multi-platform data collection, integration, and analysis.**

While there exist many data aggregation tools, there are a few persistent deficiencies that this project aims to address. These include fragmented data access, platform-specific API constraints, inconsistent data formats, lack of user-friendly interfaces, privacy risks, tight API access controls, and introducing fees or shutting down third-party API access entirely (Maiz, 2024). This complexity directly affects individuals who wish to gain a holistic view of their social media activity. Content creators must juggle multiple logins, interfaces, and data formats, which wastes plenty of time or lead to missed engagement opportunities, users and brands cannot collaborate efficiently, and users lack tools for self-reflection or archiving.

## **2.4 Review of Existing Systems**

### **2.4.1 Juicer.io**

Juicer.io allows users to aggregate feeds from multiple platforms into a single display. It supports Instagram, LinkedIn, Facebook, Twitter, Tiktok, and Youtube. It offers a free tier but places branding on the aggregated feed and limits the number of platforms connected. It is primarily used for embedding public content on websites rather than providing a private dashboard for personal use.

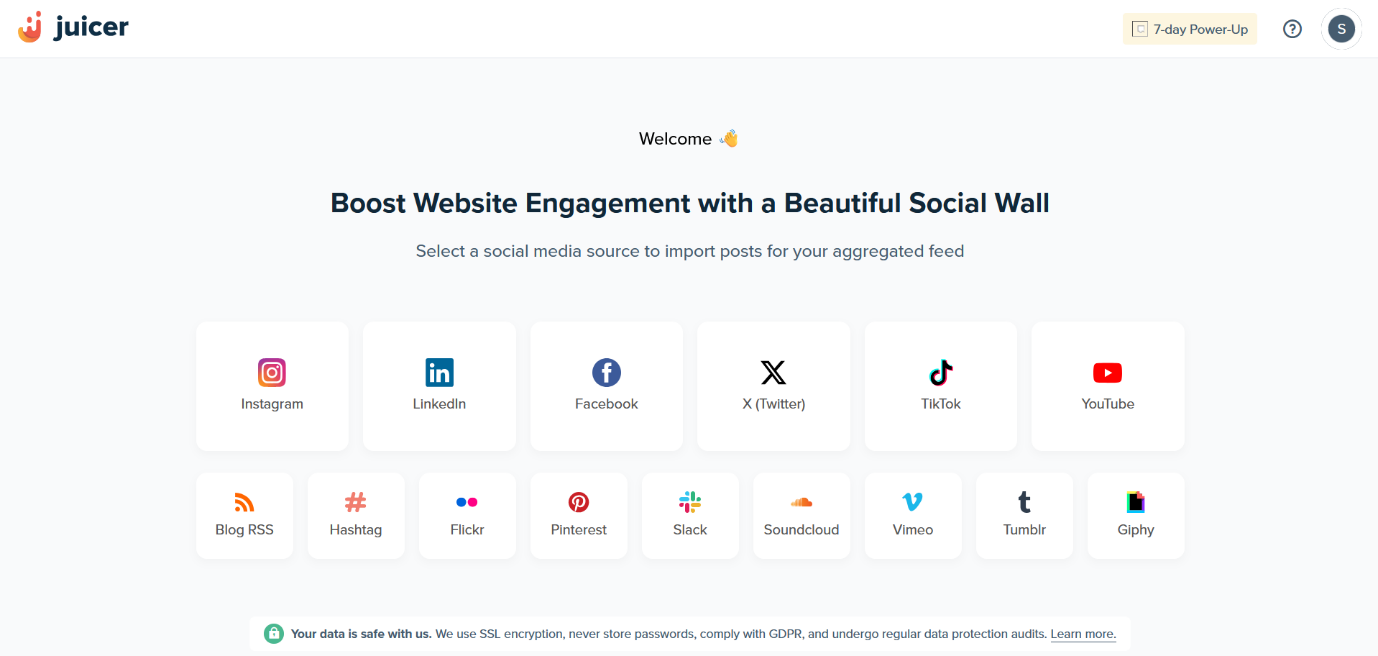


Figure : Juicer.io screenshot

### **2.4.2 Curator.io**

Curator.io aggregates social feeds into embeddable widgets but it is targeted at brands rather than individuals. The free version supports limited updates and embeds, and personalization features for individual content creators are minimal. It does not support analytics or self-reflection tools.

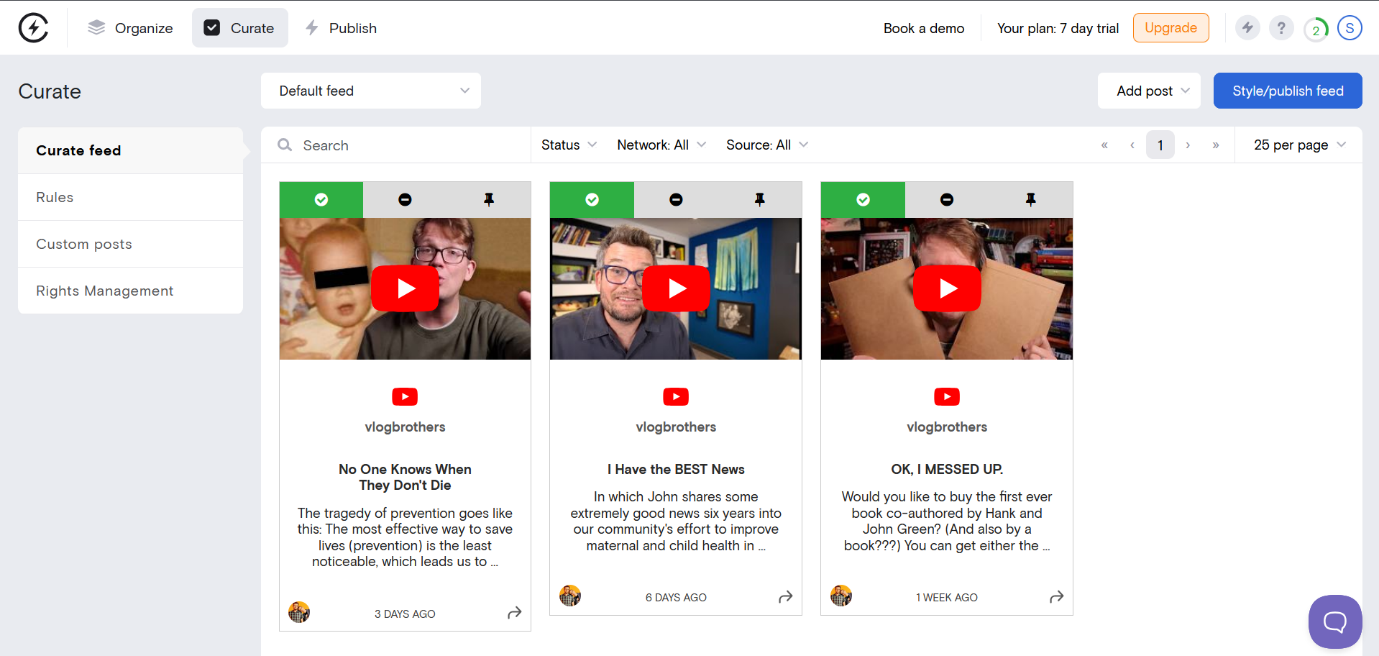


Figure : Curator.io screenshot

### **2.4.3 Hootsuite**

Hootsuite is a comprehensive platform for managing social media marketing, designed primarily for businesses and agencies. While it allows connection to multiple platforms, its complexity, pricing, lack of archival features make it unsuitable for personal use.

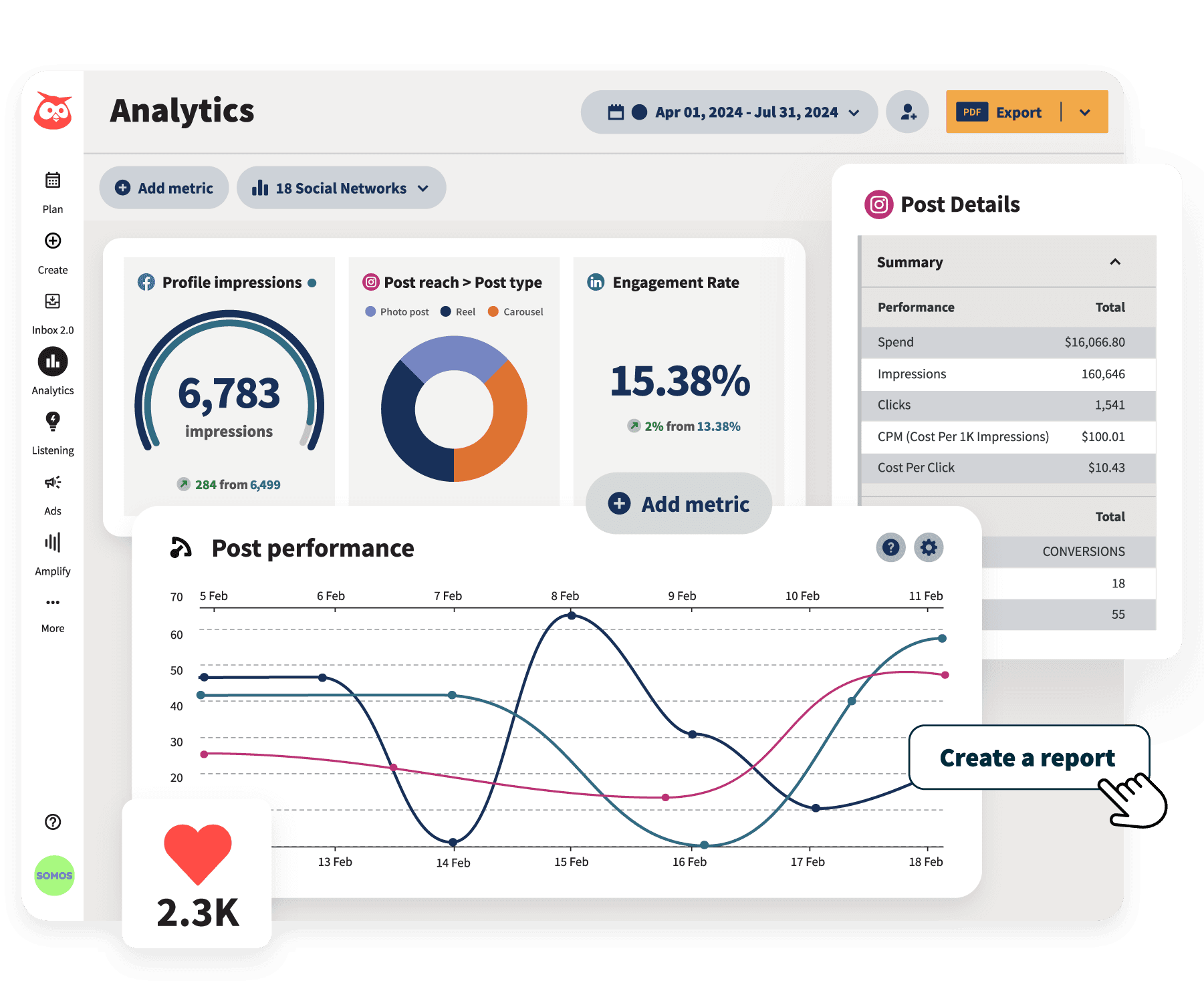


Figure : Hootsuite screenshot

## **2.5 Gaps in Related Works**

The vast majority of existing platforms are primarily tailored to meet the needs of businesses and marketing professionals. (*10 Best Social Media Aggregators to Embed and Display in 2024*, n.d.) In contrast to the abundance of business-focused solutions, the market offers considerably fewer tools and services specifically designed for individual users who wish to aggregate their own social media data from their various personal accounts for their own convenience and use. While some of the aforementioned platforms might offer free or lower-tier subscription plans that could potentially be utilized by individuals, these plans often come with limitations regarding the number of connected social media accounts, or the presence of the platform's branding on the aggregated feeds. (*Free Social Media Aggregator Tools*, n.d.)

## **2.6 Technologies to Be Used**

The project will utilize RESTful APIs provided by social media platforms to fetch user data. OAuth 2.0 will be used for secure authentication and authorization. The backend will be implemented using Node.js, while the frontend will use React.js. MongoDB will serve as the database for storing user data, with Redux used for state management. Chart.js will provide data visualization features. The system will be deployed via Firebase for scalability.

## **2.7 Conceptual Framework**

Users will begin by making an account then linking all the social media accounts. The system will then pull all the relevant data, organise it, and present it in a visual format for the user to see.

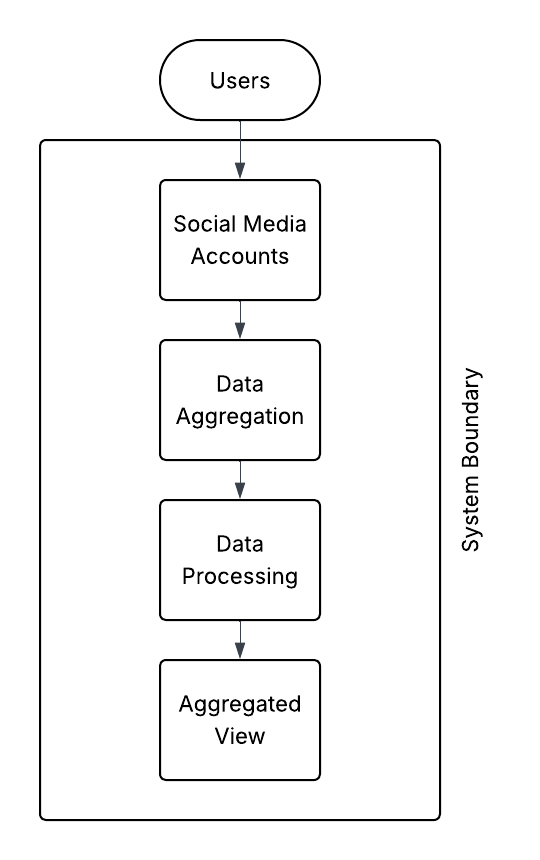


Figure 4: Conceptual Framework

# **CHAPTER 3: METHODOLOGY**

## **3.1 Introduction**

This chapter outlines the methodology adopted for this system. It describes the selected research paradigm, development methodology, specific activities at each stage of the project lifecycle, and the technologies and tools that will be used.

## **3.2 Methodology**

The project adopts the Object-Oriented Analysis and Design (OOAD) paradigm. OOAD was selected because the proposed system involves multiple interconnected and OOAD offers a clear, structured approach to modelling real-world objects and interactions.

The project will follow the Agile software development methodology. This methodology was chosen due to the iterative and adaptive nature of the project, which involves ongoing feedback and adjustments based on testing and API limitations. The rapid prototyping and continuous integration are essential for managing external dependencies like social media APIs.

### **3.2.1 Ideation and Requirements Engineering**

This step involves analysis of user complaints and system limitations in current tools. Existing data from the literature review shall be used for the analysis.

### **3.2.2 Analysis and Design**

This choice is based on the modular nature of the system, where components like user authentication, platform integration, and analytics visualization can be encapsulated as distinct objects.

### **3.2.3 Development**

A web platform is preferred due to its compatibility with API-based systems and broader user reach. The frontend will be developed using React.js, the backend Node.js, Firebase will be used as the database and authentication will be done using OAuth. The coding will be done on Visual Studio code and GitHub shall be used for version control.

### **3.2.4 Testing**

The system will be tested through unit testing to validate individual components, integration testing to ensure modules work together, and system testing to verify the system meets its specifications.

## **3.3 System Analysis**

This section will cover analysis of the system which includes modelling tools like a use case diagram.

### **3.3.1 Use Case Diagram**

This will illustrate how actors/user interact with the system.

### **3.3.2 Class Diagram**

This is used to model objects in the system like users and API connections

### **3.3.3 Sequence Diagram**

This will show the flow of data and interactions during any task the system will carry out.

### **3.3.4 Entity Relationship Diagram**

This will illustrate how data will be stored in the database.

### **3.3.5 Wireframes**

This will be used to model the user interface and navigation structure.

## **3.4 System Design**

This section shows models such as wireframes that satisfy lay out the requirements for the system. These models are wireframes, database schema, and system architecture.

### **3.4.1 Wireframes**

This is a mock-up of the UI the system will use. In this system, wireframes will show how the user will be presented with information.

### **3.4.2 Database Schema**

This is a visualisation of the logic used to store the data and how the information is related to the rest.

### **3.4.3 System Architecture**

This will illustrate how a user interacts with the system, the various components of the system, and how the system will handle data.

## **3.5 System Development Tools and Techniques**

This section describes the tools used in developing the system and the justification for each.

### **3.5.1 React.js and Node.js**

React.js shall be used to develop the frontend while Node.js will be used to develop the backend

### **3.5.2 Firebase**

Firebase will be used as the database to store the data retrieved from the APIs.

## **3.6 Deliverables**

This section describes the deliverables expected from the project.

### **3.6.1 Authentication Module**

This handles user login, registration, and secure OAuth integration for social media platforms

### **3.6.2 User Management Module**

This manages user profiles, roles (admin, user) and linked social media accounts.

### **3.6.3 API Integration Module**

This will interface with external APIs to fetch social media data

### **3.6.4 Analytics & Visualisation Module**

This is responsible for transforming stored data into meaningful charts and summaries.

### **3.6.5 UI Module**

This will handle the frontend interface where users can view and interact with aggregated data.

### **3.6.6 Documentation**

This will be the written report covering the research, development, and design of the system.

# **References**

Adishree, W. (2024, October 16). *What is Social Media Aggregation? A Comprehensive Introduction*. Retrieved from tagembed: https://tagembed.com/blog/what-is-social-media-aggregation/

bubblinn. (2021, February 18). *Looking for a social media aggregator for personal use*. Retrieved from Reddit: https://www.reddit.com/r/socialmedia/comments/lmp1cb/looking\_for\_a\_social\_media\_aggregator\_for/?rdt=55157

Chaffey, D. (2025, February 14). *Global social media statistics research summary*. Retrieved from Smart Insights: https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/

Mailchimp, I. (2025, May 9). *How Social Media Aggregators Streamline Content Creation*. Retrieved from Mailchimp: https://mailchimp.com/resources/content-aggregator/

Maiz, E. (2024, March 18). *Twitter API Pricing and Limits*. Retrieved from Data365: https://data365.co/guides/twitter-api-limitations-and-pricing

PR, G. (2025, April 4). *12 Best Social Media Aggregators to Boost Conversions in 2025*. Retrieved from flockler: https://flockler.com/blog/social-media-aggregators

Qlik. (2025, May 9). *Data Aggregation*. Retrieved from Qlik: https://www.qlik.com/us/data-management/data-aggregation

SentinelOne. (2025, April 23). *What is Data Aggregation? Types, Benefits, & Challenges*. Retrieved from SentinelOne: https://www.sentinelone.com/cybersecurity-101/data-and-ai/what-is-data-aggregation/

Zanini, A. (2025, May 19). *Data Aggregation – Definition, Use Cases, and Challenges*. Retrieved from brightdata: https://brightdata.com/blog/web-data/data-aggregation

*10 Best Social Media Aggregators to Embed and Display in 2024*. (n.d.). Retrieved May 21, 2025, from https://taggbox.com/blog/best-social-media-aggregators/

*Free Social Media Aggregator Tools*. (n.d.). Retrieved May 21, 2025, from https://curator.io/blog/free-social-media-aggregator-tools