IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# User Engagement Analytics

**Shivam Pandey, Sahitya Uniyal, Shashank Sahu, Yash Gahlot, Ankita Singh**

Student, Student, Student, Student, Assistant Professor
Computer Science Engineering,
SRM Institute of Science and Technology, Ghaziabad, India

***Abstract:*** A web app that takes data from users regarding the time they spent on various activities throughout the day. This data is collected on a regular basis and is used to analyze user behavior and predict various trends. Behavior analysis can be used by the user to change unwanted habits to more productive habits. Data can also be used by ML prediction models to predict various trends and to analyze market situations for certain products. Additionally, it can also be made available to the public with the help of an API service.

***Index Terms*** **- Application Programing Interface, User Engagement, Machine Learning**

## I. Introduction

User Engagement is the all-inclusive term for the various interactions the users have with the activities in their daily life. User engagement is also a prime way to analyze how users interact with entities around them, be it positive or negative. It helps in analyzing whether one's actions or activities throughout the day are aligned in the direction of the desired result or is it leading to something not desired. With time the users are spending more of their time on online activities than outdoor activities. According to a survey people like to spend more of their time on mobile phones rather than doing outdoor activities and they also report being online 'almost constantly'. Fast changes like these always spark fears about possible negative impact on our health.

In this project, we propose a web app that collects data from users every day to create a trend graph/analysis after the end of every week. The goal of this project is to help users shape their interactions with daily activities. The app will use data collected through the survey to analyze the user's interaction pattern with different activities. Initially, the survey will take key user information like name, age, nationality, etc., for building the user's profile and classifying the data more efficiently. Collected data will be segregated into several different categories, the primary one being Mobile, Indoor, and Outdoor. Additionally, we will be developing a database that will be used for achieving the above-mentioned goal. To achieve this goal, we will be analyzing the data and providing the results from several viewpoints. Users can access the generalized data and results from the analysis of the database. Or access personal data through their profile, this data will show changes in the user's interaction over time. Users can then decide whether the changes made were positive or negative, and what steps should be taken moving forward. The database also has scope across several different fields like entertainment, machine learning, targeted ads, app/ web development, etc., just to name a few. Through this project, we aim to provide people with an opportunity to improve or change themselves.

This project will contribute towards mental health, physical fitness and the rise in the addiction of technology. It will make people more aware of these things.

## II. Literature Survey

In this literature review, we rigorously screened and analyzed previously published papers on the topic. The literature review gave us many insights into user engagement, user behavior and digital wellbeing. The table below contains some of the articles and documents published about user engagement.

Table 1: Tabular summary of Literature Review

| Author (Year) | Title of Paper | Methodology | Dataset Used | Limitations |
|---|---|---|---|---|
| "Heather L. O'Brien, Elaine G. Toms"<br><br>(2008) | "What is User Engagement? A Conceptual Framework for Defining User Engagement with Technology" [1] | "Participants, Interview Protocol, Procedure, Data Analysis" | "Strauss Corbin, McCarthy and Wright's, Strauss' grounded theory, Merriam-Webster Online" | The research findings are limited to a specific sample and time, so they are not conclusive, but they align with previous research and may have potential for generalization |

| | | | | |
|---|---|---|---|---|
| "Iryna Susha, Åke Grönlund, Marijn Janssen" (2015) | "Organizational measures to stimulate user engagement with open data" [2] | "It reviewed literature on the problem of poor data usage and conducted four case studies in different organizations and countries" | "Evans and Campos, Gurstein, Davies and Bawa, Zuiderwijk et al, Zuiderwijk and Janssen" | This research relies on generic theoretical recommendations that may change over time. There is also a lack of evidence of actual output. |
| "Alberto Monge Roffarello, Luigi De Russis" (2019) | "The race towards digital wellbeing: Issues and opportunities" [3] | "ParseHub2(web scraping tool), within-subject experiments" | "Reviews done in different years: 3% in 2015, 8% in 2016, 25% in 2017, 64% in 2018" | Potential bias in user reviews, a smaller number of participants, and a lack of consideration for overuse of other technology devices. |
| "Zhuojun Gu, Ravi Bapna, Jason Chan, Alok Gupta" (2022) | "Measuring the Impact of Crowdsourcing Features on Mobile App User Engagement and Retention: A Randomized Field Experiment" [4] | "Thomas JS (2001)'s methodology for linking customer acquisition to customer retention, Field experiment on a "Catchphrase" game app" | "Field experiment on different users, manipulating tests for users, different records." | Narrow focuses on fundamental aspects of crowdsourcing and the potential for bias due to examining only a casual context. The limited time frame raises questions about the long-term impact. |
| "Abdulsalam Mustafa Salihu, Nor'ashikin Ali, Jaspaljeet Singh Dhillon, Gamal Alkawsi, Yahia Baashar" (2022) | "User Engagement and Abandonment of mHealth: A Cross-Sectional Survey" [5] | "Data were coded based on themes and analyzed accordingly, categorizing the data accordingly" | "Surveys done over the period of September 2020 to January 2021, 17-item survey" | Recall bias in the checkpoint assessments, a lack of identification of the theories applied in the mHealth apps, and the absence of demographic information on respondents' nationality, race, or profession. |
| "Mohamed Basel Almourad, Amen Alrobai, Tiffany Skinner, Mohammed Hussain, Raian Ali" (2021) | "Digital wellbeing tools through users' lens" [6] | "SPACE and GDW, were selected and extensively investigated to collect evidence of their capabilities, design and potential use" | "Google Play and Apple Store sites, the application website, the total extract covers 350 reviews" | Lack of research on users' perceptions of digital well-being applications, including their self-regulation features. |

The literature review we conducted for our project proved to be valuable in several ways. We gained insight into how to collect highly organized and effective data through the research papers and articles examined in the review. The literature review highlighted several limitations. These limitations include the quality and quantity of data used, and difficulty in interpreting complex results. The things learned clarified user engagement and how it is used in digital well-being and how companies make use of it to retain users.

## III. Proposed Methodology

In today's world, handling large amounts of data presents a formidable challenge that requires careful planning, organization, and execution. To overcome this obstacle, it is crucial to focus our efforts on establishing clear goals and objectives for our analytical models while considering possible limitations or restrictions. Following this, we must collect data from different sources while undertaking various pre-processing steps such as data cleaning, integration techniques followed by transformative measures aimed at ensuring our dataset is clean, rationalized and properly structured for analysis. Analytical methods including clustering, classification and visualization algorithms are then applied efficiently with customized strategies built around requirements set forth by clients. To maximize performance whilst optimizing storage space needs, normalization coupled with indexing have been employed as time tested techniques proven to be extremely effective towards these ends.

Indexing methods specifically have demonstrated improved information retrieval with quick searches ensured through creating search trees or hash tables that systematically organize large datasets helping to optimize analysis time.

In order to present personalized recommendations that cater to specific user groups or individual users themselves advanced techniques like data visualization along with clustering strategies have been put into place in order identify unique emerging patterns within each dataset whilst also providing customized recommendations based on specific goals. Utilizing methods like performance measurement, user feedback gathering, and continuous improvement efforts are essential for maximizing outcomes while minimizing potential inefficiencies.

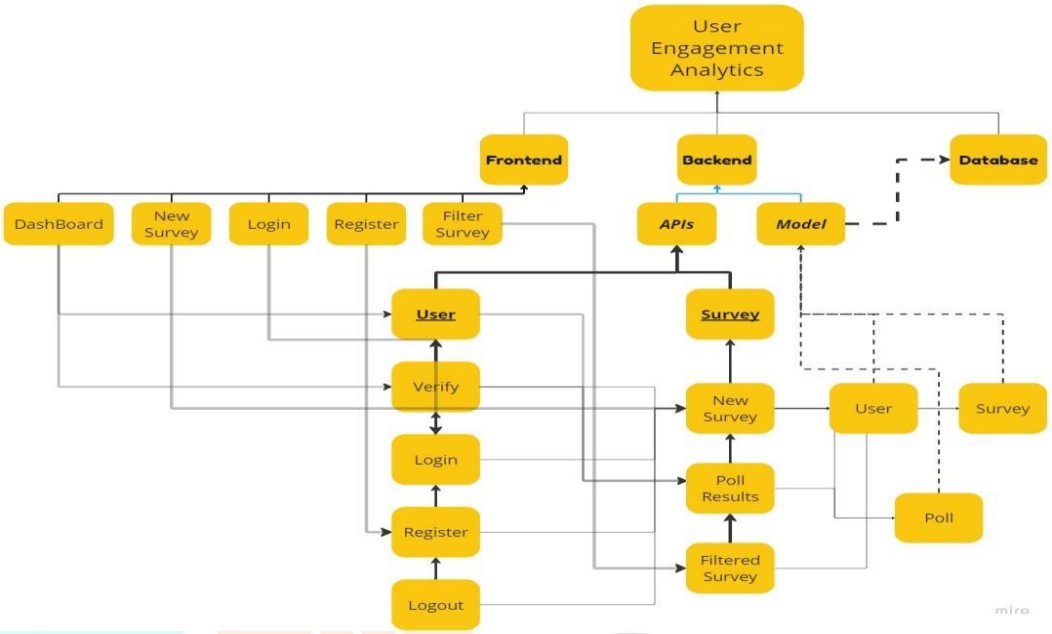


figure 1: proposed methodology

## IV. Existing Problems and Proposed Solution

One of the most significant challenges is the data required to do an analysis. In addition to the issue related to data quantity there are also issues regarding data quality, variation in data user community and user privacy. With large amounts of data, indexing or clustering on various attributes is required, which is performance-intensive when changes or edits are made. The right authentication and authorization methods will be utilized to boost security and after a predetermined period, the user will be logged out to increase the security of the session. And to preserve data integrity, no user may alter the data of another user. For the issue of data clustering in case of adding more categories/attributes we will be using normalization on the data/table.

## V. System Architecture

The system architecture includes MongoDB, which is a source-available document-oriented database program. It is classified as a NoSQL database and uses JSON-like documents with optional schemas. MongoDB supports ad-hoc queries, indexing for efficient data retrieval, and aggregation using the aggregation pipeline, map-reduce function, and single-purpose aggregation methods. It also allows server-side JavaScript execution for queries and aggregation functions.

In addition to MongoDB, the architecture includes Express.js, a back-end web application framework for building RESTful APIs with Node.js. Express.js is known for its simplicity, performance, and flexibility. It is used for building dynamic websites, developing RESTful APIs, and handling server-side tasks such as processing data and managing database connections. Express.js has a lightweight architecture that can handle large amounts of traffic and can be easily integrated with other technologies and frameworks.

## VI. System Specification

table 2: system specification

| Description | Type/Version |
|---|---|
| Processor (CPU) | Intel(R) Core (TM) i7-10750H |
| GPU | Intel(R) UHD Graphics |
| Disk | 512 GB SSD |
| Memory | 16 GB DDR4 |
| External Components | Keyboard and Mouse |
| Operating System | Windows 11 |
| MongoDB | 6.0 / 2022 |

## VII. Result And Analysis

Aim of the project was to create a webapp to collect the data from the users on their daily activities to analyze and provide a common graph and a personalized graph on various activities. Below are the results of analysis along with some of the key functions of the webapp.

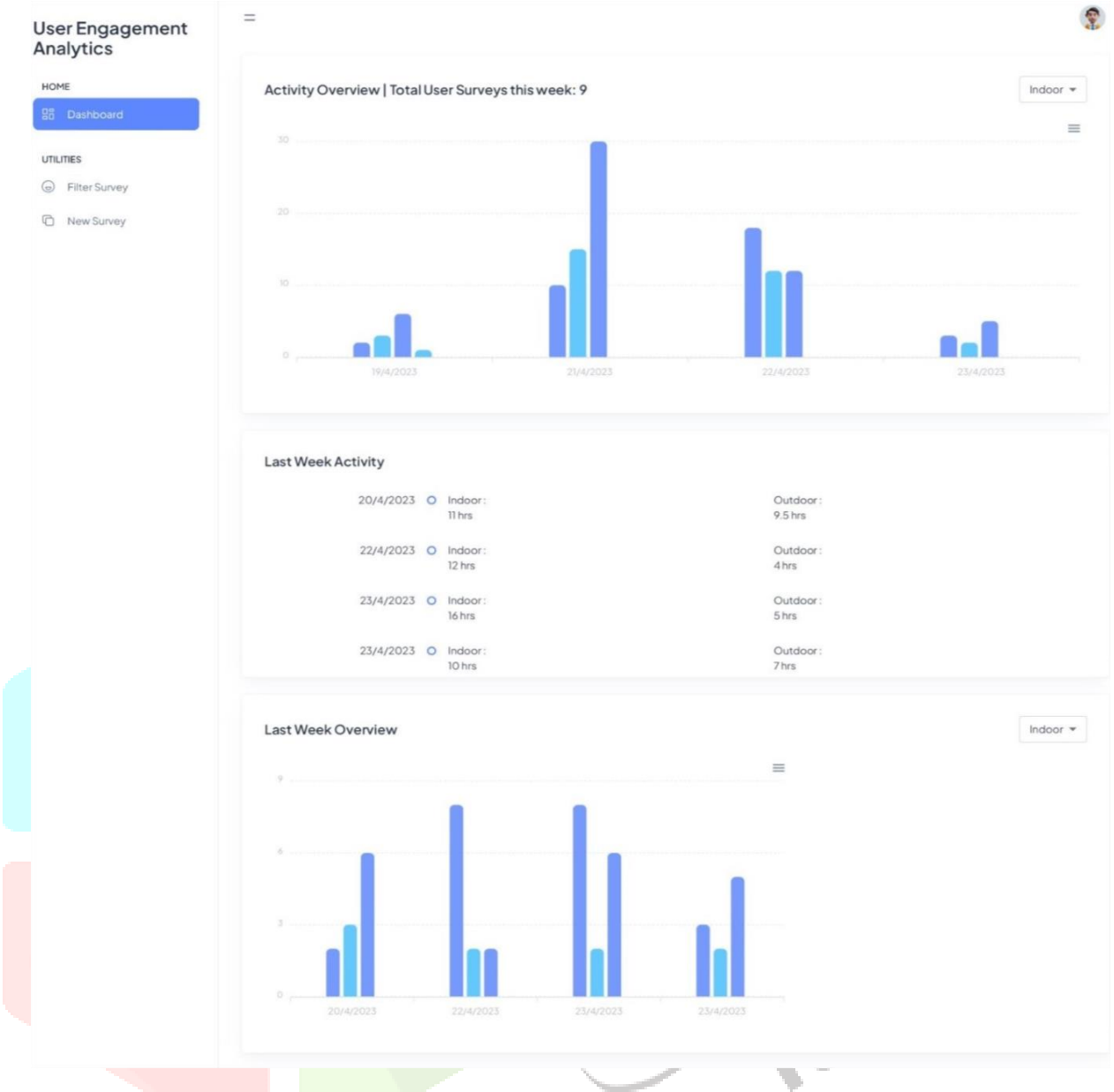


figure 2: Dashboard

Image 2 shows the dashboard of the profile section. It shows the personalized graph analyzed based on user's activities. It also has options to track activities of previous week. Personalized graphs help users by helping them to focus on habits they wish to change and track their progress when doing so.

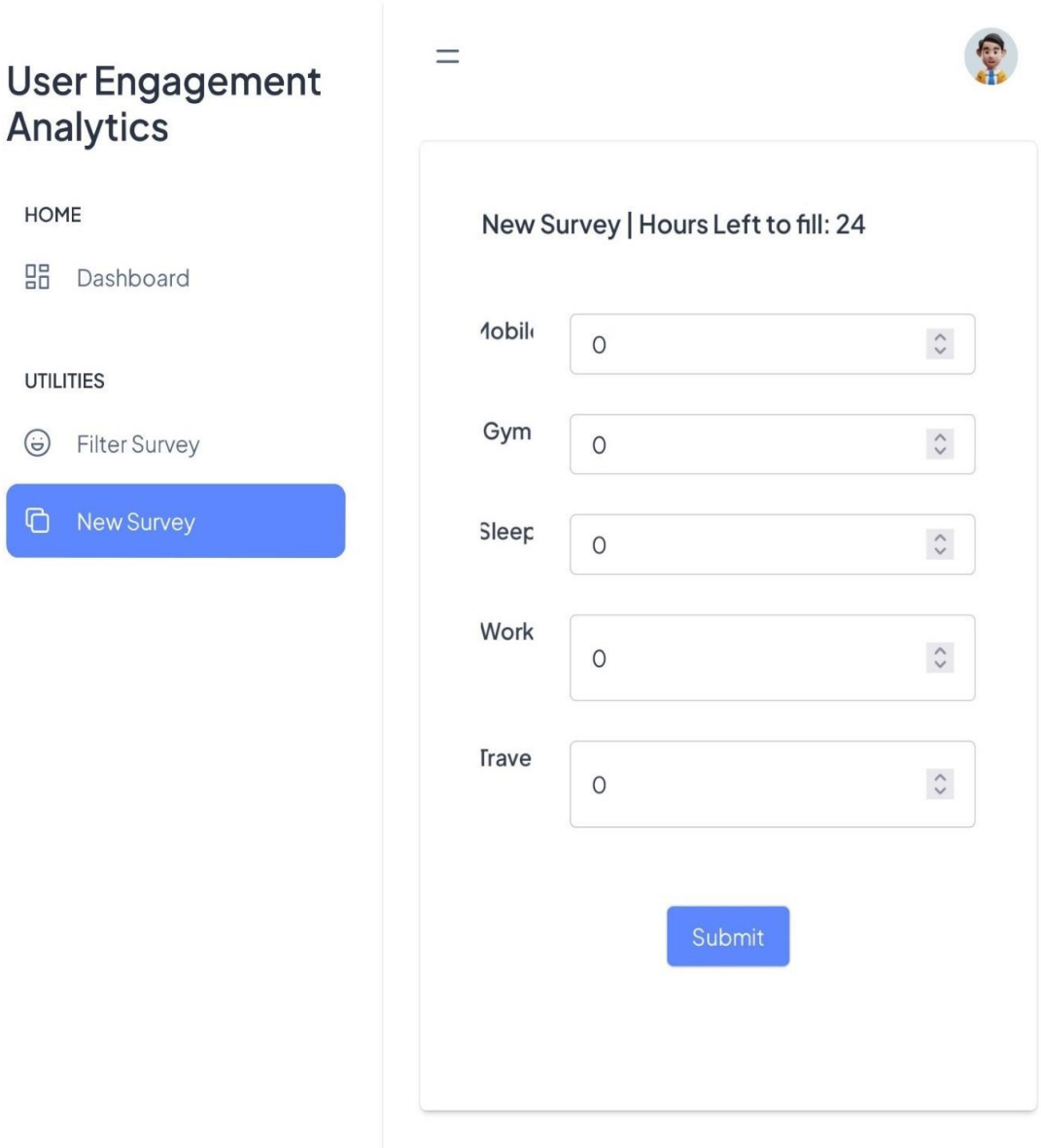


figure 3: new survey form

Image 3 shows a form that takes number of hours on different activities. The form takes time for popular activities, new activities can be added by admin. This is then analyzed and converted into visualized format, in this case graph.

Image 4 below shows filter survey function through which users can find their graphs according to their preference by providing some criteria as per user's requirement.

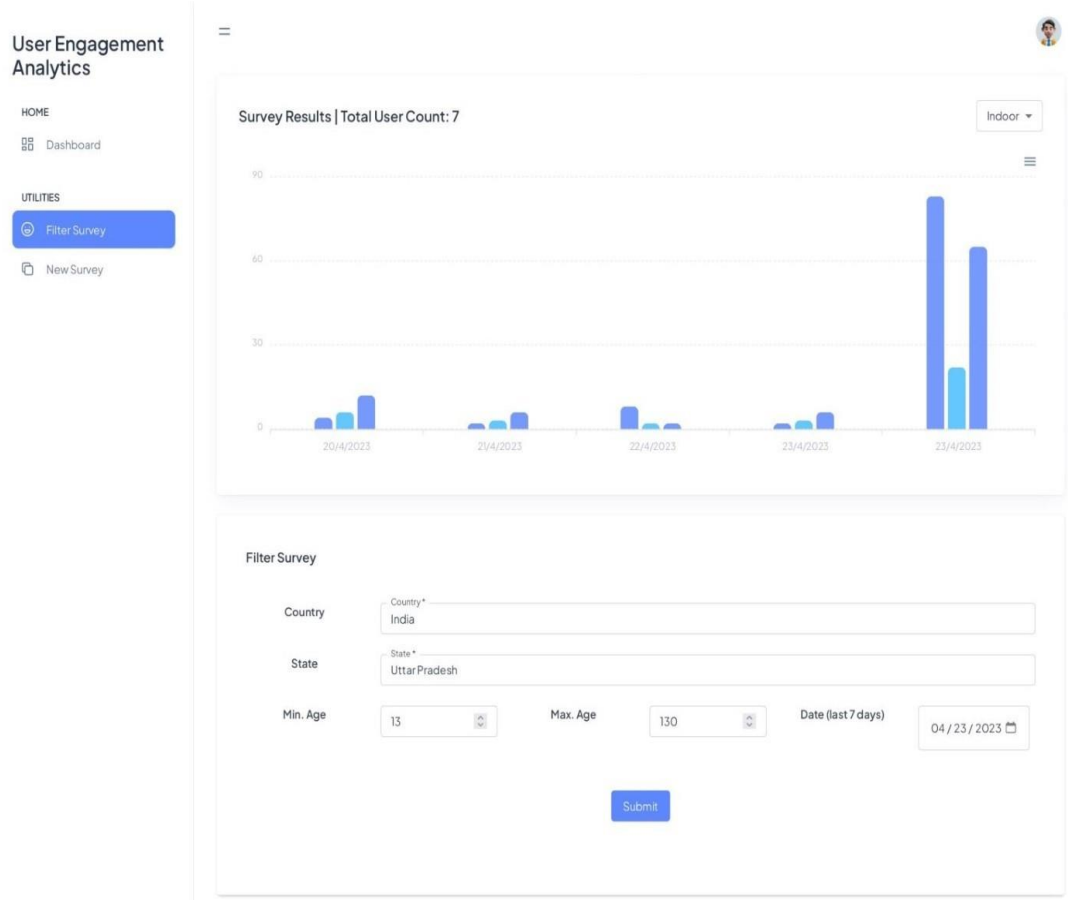


figure 4: filter survey form

## VIII. Conclusion And Limitations

The project has successfully collected data from users and generated usage graphs that optimize their daily routines and identify trends in user behavior for companies. This has allowed users to analyze their activities and make informed decisions to improve their lives, while also helping companies better understand their customers and develop products and services to meet their needs. However, the reliance on manual user input for data collection can be time-consuming and error-prone, and with the increasing volume of user data, effective analysis requires more advanced techniques. Additionally, as the database expands, robust security measures are necessary to prevent data breaches. While collaboration and integration with other organizations may present opportunities, technical and logistical challenges arise in establishing APIs and partnerships for data sharing. Improvements in data analysis speed, accuracy, storage, and retrieval methods become necessary as the user base and data volume grow, requiring continuous enhancements and resource allocation to ensure efficient and timely access to data for users and organizations.

## References

**[1]** H. L. O'Brien and E. G Toms, (2008). "What is user engagement? a conceptual frame- work for defining user engagement with technology." Journal of the American society for Information Science and Technology, 59(6), 938–955.

**[2]** Susha I., Grönlund Å., and Janssen M. (2015). "Organizational measures to stimulate user engagement with open data." Transforming Government: People, Process and Policy, 9(2), 181–206.

**[3]** A. Monge Roffarello and L. De Russis, (2019). "The race towards digital wellbeing: Issues and opportunities." Proceedings of the 2019 CHI conference on human factors in computing systems. 1–14.

**[4]** Z. Gu, R. Bapna, J. Chan, and A. Gupta, (2022). "Measuring the impact of crowdsourcing features on mobile app user engagement and retention: A randomized field experiment." Management Science, 68(2), 1297–1329.

**[5]** A. S. Mustafa, N. Ali, J. S. Dhillon, G. Alkawsi, and Y. Baashar, (2022). "User engagement and abandonment of mHealth: a cross-sectional survey." Healthcare, Vol. 10, MDPI. 221.

**[6]** M. B. Almourad, A. Alrobai, T. Skinner, M. Hussain and R. Ali, (2021). "Digital wellbeing tools through users' lens." Technology in Society, 67, 101778.

**[7]** H Buabbas, Ali Jasem, A.-M. M. A. A.-T. B. A. B. M. A. (2020). "Responding to the threat of chronic diseases in india." BMC Pediatrics, 20.