****

**DESIGN AND IMPLEMENTATION OF A WEB-BASED RADIO AUDIENCE MEASUREMENT SYSTEM**

**BY**

**AYEGBA JESSE-JOSEPH, ANIBE**

**(16CK020793)**

**SUBMITTED TO**

**THE DEPARTMENT OF ELECTRICAL AND INFORMATION ENGINEERING,**

**COLLEGE OF ENGINEERING,**

**COVENANT UNIVERSITY**

**IN PARTIAL FULFILLMENT FOR THE AWARD OF THE DEGREE OF BACHELOR OF ENGINEERING (B.ENG.), ELECTRICAL AND ELECTRONICS ENGINEERING, COVENANT UNIVERSITY, OTA, OGUN STATE, NIGERIA**

**JANUARY 2021**

1. DECLARATION

I hereby declare that the work detailed in this report was completed by me under the supervision of Engr. Omoruyi Osemwegie in the Department of Electrical and Information Engineering, Covenant University. Also, I affirm that as far as I could possibly know, no piece of the report has been submitted here or somewhere else in an earlier application for the honor of a degree. All sources of information utilized thus have been properly recognized.

…………………………………….

AYEGBA JESSE-JOSEPH ANIBE

16CK020793

1. CERTIFICATION

This is to certify that the project titled "Design and Implementation of a Web-based Radio Audience Measurement System” by AYEGBA JESSE-JOSEPH, ANIBE, meets the requirements and regulations governing the award of the Bachelor of Engineering, B.Eng. (Electrical and Electronics Engineering) degree of Covenant University and is approved for its contribution to knowledge and literary presentation.

|  |  |
| --- | --- |
| **Student:** | Name: AYEGBA JESSE-JOSEPH ANIBE  Sign: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| **Supervisor:** | Name: Engr. Omoruyi Osemwegie  Sign: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| **HOD:** | Name: Prof. Emmanuel Adetiba  Sign: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| **Internal Examiner:** | Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Sign: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| **External Examiner:** | Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Sign: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |

1. DEDICATION

This project report is dedicated to God Almighty my Creator, who has been my rock, guide, source of inspiration, wisdom, knowledge and understanding. He has been my strength and in his arms I have been sheltered. I also dedicate this work to my lovely siblings who have encouraged me every step of the way and whose encouragement has made sure that I give it all it takes to complete that which I have started. To my parents Mr. and Mrs. Ayegba who have been affected every step of the way by this quest, the words that will express how much appreciation I have for you have not been invented yet. So, in all sincerity I say thank you and God bless you.

1. ACKNOWLEDGEMENT

My sincere appreciation goes to God Almighty for enabling me to complete this project. I would also like to express my deepest gratitude to the members of my family for their unending support.

I would like to express my appreciation to my project supervisor, Engr. Omoruyi Osemwegie, for providing me with the necessary guidance needed to complete this project.

# INTRODUCTION

## BACKGROUND OF THE STUDY

The traditional definition of audience measurement is the estimation of the number of people watching a certain TV show or channel, or the number of listeners who are tuned to a particular radio program or channel. Audience measurement takes into account, the behaviour of the audience as well as its demographics [1]. Direct and indirect measurement methods are used, and the results of a carefully selected sample are usually extrapolated to give estimates for the entire population. Audience research is an important aspect of television and radio production broadcasting as well as newer forms of media material. Internet, IPTV, mobile phones, and personal computers are all examples of delivery methods. Audience measurement can be used for everything from self-promotion to fine-tuning a service [2].

Content consumption measurement is one of the solutions that audience measurement technologies aim to bring about. One of the most common approaches used by service providers or broadcasters to get important data for improving service offerings or setting advertising rates is to evaluate content consumption. Its uses are much broader than that. Without accurate audience data, many firms may be reluctant to join in the new delivery platforms.

### BRIEF HISTORY OF AUDIENCE MEASUREMENT

First launched in the late 1940's soon after the start of commercial broadcasting, the audience measurements allowed radio broadcasting business to flourish through networks which offered advertisers, who paid for the estimated number of ears listening on commercials, a way to quantify the financial value of radio audiences. The first measuring techniques had several limitations because the acquisition of reliable, large-scale data was costly. Despite the limitations, standards for measurement remained largely unchanged for years until the explosion of digitally accessible data resulted in such devices as cable boxes, video on demand boxes and cell phone as well as web apps, internet browser clicks, web queries, and social media activities. Radio listeners now leave digital footprints that may be used to follow practically every part of their everyday lives, allowing for large-scale data aggregation for individual users and groups, as well as tracking of more individuals on more dimensions for more programs. Data is now more substantial, real-time, and less expensive to get, allowing for precise and fine grained radio audience monitoring [1].

### AUDIENCE MEASUREMENT IN DIGITAL SIGNAGE

The digital signage service displays advertising and valuable information on terminals with electronic displays, and it is also possible to collect data using a variety of sensors, such as cameras. In comparison to standard DID (Digital Information Device) services, which only supply one-way content, digital signage services can offer more advanced features like user interactivity and audience measurement. It is feasible to give appropriate material to users and boost advertisement effects by measuring audience behavior. Digital signage services are becoming more popular these days for a variety of reasons allowing for increased contact and intelligence services especially now that digital signage devices are being put in a variety of locations including public spaces [3]. Bus stops, hallways, and shopping malls are examples of public spaces. Static signs are losing their impact, but digital signage, on the other hand, are designed to capture consumers' attention and convey messages that are tailored to improve their experience. It is useful for obtaining audience and environmental data to aid the kind of content to present to then. This project presents a web-based approach for gathering these data, using machine learning to accurately predict and identify different sound data.

Diagram

Description automatically generated

The general digital signage architecture with audience measurement [3].

## PROBLEM STATEMENT

With the rapid technological advancements that the world is experiencing, a shift from the traditional way of doing things to a more efficient way is required. Currently, radio audience measurement is done in a way that requires full human input by recalling their listening patterns which is prone to some inaccuracies. It has been noted that humans are inherently fallible, and errors are highly inevitable. Therefore, there are high possibilities of errors gotten from data taken from human inputs to lead to errors in the calculation of radio audience listenership.

## AIMS AND OBJECTIVES

This project aims to design and build a web-based audience measurement system that will be used in the estimation of the number of people who are tuned to, and actively listening to radio stations. The implementation of the application was done as a modern system that would reduce all forms of human error and, also make the data gotten from the system reliable.

The objectives of this project are to:

1. To design a mobile and web application using figma.
2. To build out the mobile application using JavaScript with a library called React native.
3. To build out the web application application using css and JavaScript with a library called React JS.
4. To use a cloud storage platform like firebase to store the audio recordings of participants.
5. To use machine learning to accurately predict the nature of the recordings that participants have uploaded.
6. To determine the effectiveness of the system by comparing the results to a predetermined set of results.

## SIGNIFICANCE OF THE PROJECT

The data gotten implementation of this project will be beneficial to the following groups of people:

1. **Advertisers**: Data gathered from this project could be used by advertisers to target a specific audience to advertise their content to.
2. **Researchers**: Researchers are constantly looking for ways to improve the spread of good and reliable information. The data gotten from this project could be of tremendous help to achieve this goal.
3. **Students**: The project provides more information for further works and research on the subject area.
4. **Data analysts**: Data analysts need data sets to analyze, to predict future problems as well as solutions. Data gathered from this project could aid them.

The implementation of this project will also be inline with the sustainable development goal(SDG) 9. Which is to “build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation”.

## MOTIVATION FOR THE STUDY

With the rapid increase in the population of the world today, there is a progressive increase in the demand for good and quality information content. Advancements in technology are gradually bringing about a change in what people engage in as well as their interests. It is clear, that what people listened to some 40 years ago on radio stations is not what they are interested in listening to these days. Thus, the need for some way to accurately determine what people are really interested in listening to. This project tries to measure the listening patterns of radio audiences and predict what people are interested in listening to.

## METHODOLOGY

Two client-side applications were built a web application for administrators and a mobile application for participants. The mobile application sends audio data to a server-side service called firebase. The server-side service handles things like user authentication to allow participants sign up and login. The mobile and web application were built using JavaScript with frameworks like React native and React js respectively. A JavaScript library called ml5 js handles machine learning on the administrators’ dashboard, to predict the kinds of recordings that participants have sent to the server. These results are then analyzed and used to estimate the listening audience.

## PROJECT ORGANISATION

**Chapter 1**: Contains a general overview on the project, the background information, the aim and objectives of the project, what problems the project seeks to solve as well as a brief outline of the methodology.

**Chapter 2**: Gives the literature review, it discusses the past related works on the project subject area. It contains theoretical background and other concepts necessary to make the project well understood.

**Chapter 3**: Contains the methodology of the project, how the project is going to be carried out, design of the project, it contains all components required for the design and block diagrams and software design.

**Chapter 4**: This chapter describes the implementation phase of the project as well as testing. The functional system design will be shown in detail as well as real pictures of the project undergoing testing. The results of the project will be analyzed in this chapter.

**Chapter 5**: This is the last chapter of the project report, and it contains conclusions and recommendations for the project. It also shows the results the project was able to achieve.

# LITERATURE REVIEW

## INTRODUCTION

With the growth of data, the question of what to quantify in terms of radio listeners and their habits has become increasingly crucial. Multiple radios are frequently used at the same time. This necessitates cross-platform measuring. Furthermore, the fact that radio users are adjusting their listening habits at an increasing rate suggests that more effective radio audience measurement techniques need to be employed.

This chapter covers a proper description of concepts in this work, it delves into radio audience measurement systems, its history, evolution and the use of audience measurement systems in various nations and economic settings. Finally, a review of previous research and projects relating to the concept of radio audience research, web-based and machine learning systems is considered. The information garnered from this review will be very vital to the situation of this project.

## DEFINITION OF KEY TERMS

Some key terms related to this project, audience research and radio audience measurement are described below.

1. **Audience:** An audience is a group of individuals who are watching, witnessing or listening to something, such as a television program, a live speaker, or radio program, or it can refer to people who have similar tastes in entertainment.
2. **Audience measurement:** Audience measurement refers to the number of people in a group, usually in terms of radio listeners and television viewers, but also in terms of newspaper and magazine readers and, increasingly, website traffic. Audience measurement is sometimes applied to practices that assist broadcasters and advertisers in determining who is listening rather than how many people are listening. The resulting relative statistics are referred to as audience share in some parts of the world, while market share is used in other locations. Audience research is another term for this broader meaning.
3. **Sampling:** sampling is the process of selecting a sample of people who are representative of the overall population.
4. **Radio:** A Radio is a device that makes use of electromagnetic radiation to transmit electrical signals across long distances without the use of cables, such as in sound transmission, television, and radar. The number of oscillations of electromagnetic radiation per second is referred to as MHz in the context of radio.
5. **Machine Learning:** Machine Learning is a branch of artificial intelligence that deals with computer systems' capacity to solve problems on their own by detecting patterns in databases, enabling computers make successful predictions using past experiences [4]. To put it another way, Machine Learning allows IT systems to discover patterns using current algorithms and data sets and build appropriate solution concepts. As a result, artificial knowledge is developed based on experience in Machine Learning.

## HISTORY AND EVOLUTION OF AUDIENCE MEASUREMENT

### Birth (1950): Phone surveys

Systems have been in place to measure what audiences are listening to since the introduction of commercial radio streaming in 1947 [5]. These ratings systems relied heavily on phone calls to the public and used methodology designed for measuring radio audiences. By devising a survey methodology that solely inquired about what stations respondents were listening to at the time they received the call, as well as demographic information about those who were listening [6], Clark Hooper's method eliminated the prejudice and issues associated with previous random survey methods that required respondents to recall what they had heard. Many of the standard statistics used by television ratings businesses, like as audience shares, were created as a result of these calls, which are known as telephone coincidentals. By 1950, when Nielsen bought Hooper's company and launched the Nielsen Radio Index to track national radio audiences [5], the technique had been much developed.

### Infancy stage (1953) - Diaries and meters

Nielsen largely used metering devices to monitor audiences, which is a technique for tracking radio listeners that was created by Nielsen. These audimeters were used to measure what was being listened to on radio. This system reduced reliance on frequently faulty and untrustworthy sources of information. expensive phone surveys [7], but the system merely gathered data about what was on the radio, not who was listening. To close this gap, Nielsen began collecting more precise data. thanks to a subgroup of the individuals in the sample who Nielsen Diaries maintained track on listening patterns. The data is demographic. As a result, the data collected by the audimeter was supplemented. Nielsen released their rendition of an American in 1971. Data from a Research Bureau meter could be obtained via phone lines, reducing the time it took to prepare data for the market. During the day, this instantaneous audimeter saved data and guaranteed that it was delivered overnight [8]. This newfound quickness ratings and reporting from Nielsen, made Nielsen ratings become increasingly vital, as demographic data of listeners influenced advertising decisions [9]. This system has mostly stayed untouched until 1986, when the peoplemeter was brought into action and first introduced. Individual data gathering from several members of a household was possible with this unique listenership measurement technology; individual users logged their listening patterns as well as demographic information on this device. This significant shift in data collection resulted in a massive, easily accessible database with far more specific information about who was watching what, than had previously been available. Companies could now target their advertising messages more precisely.

### Childhood stage (1986) - VCRs and cable TV and

With the rise of cable television, the types of data collected for audience measurement and how they were used changed even more.

In the 1940s, cable television was introduced to provide television broadcasts to rural areas. Operators collected signals from regions with good reception and transmitted them to subscribers through coaxial cable. Cable systems could handle more stations, and beginning in the 1970s, networks tailored exclusively for cable distribution were formed, with increasingly diverse programming [10]. In 2011, there were around 5300 systems in operation in the United States, with around 60 million members. The ability of the peoplemeter technology to assess tiny, demographically targeted audiences allowed programming content and show development to be tailored to specific populations. At the same time, cable's ad-supported networks could gather the granular information needed to entice niche product advertisers to put advertising specifically targeted at specific demographic groupings. The development of cable networks increased the importance of TV ratings and increased the value of user data for advertisers. They could now

stop catering to the lowest common denominator and focus on the groups most likely to be interested in their products.

In the 1980s, the widespread use of VCRs marked yet another revolution in television viewing habits. People could now record television shows and watch them later, a practice known as time shifting. Time shifting transformed how data is collected and used in programming and advertising decisions substantially after the introduction of digital TV recorders.

### Adolescent stage (2000–2010) - The internet and social media

The early 2000s saw the rise of the Internet, which changed how people listened to radio and incorporated it into other aspects of their lives. Companies realized that clicks, searches, geolocation, tweets, purchases, and demographics could all be measured reasonably easily at a big scale in real time as consumers spent more time online. The public's entry into the new world of the Internet ushered in a new era of data [11]. This might be used to assess the influence of radio streaming and advertising on listeners in terms of attention, what they ‘‘thought" while listening, and what they bought. This can be accomplished in a variety of ways. Since its inception in 1999, comScore has developed to track demographics, clicks, and sales across a variety of platforms (e.g., both home and mobile Internet). Keyword searches can be tracked over time and by geographic region using Google Trends. Despite the fact that customers are more likely to seek for a brand online after hearing it advertised on radio, the earliest online measuring systems were unable to account for radio listenership due to data inaccessibility [12]. This is no longer the case. Since the establishment of a chat site for discussing the show The Prisoner in 199512, a slew of social radio platforms have sprung up, with usage rates skyrocketing. According to specialists from Viacom and Mass Media, users are now more engaged with radio programming since they encourage two-way dialogue.

## REVIEW OF RELATED WORKS

### Review on An Architecture for Real Time Television Audience Measurement

This paper put forward the idea that currently, audience measurement for television reports are only available after a long length of time, such as a daily report [13]. This study presents a system for measuring television audience in real time. Real-time measurement can provide channel owners and advertisers with valuable data that can help them grow their businesses [13].

[13]This study demonstrated how devices that identify the logo of the channel and transfer viewership statistics to a server via the internet can record television viewing. The viewership data is processed by the server and shown in real time on a web-based dashboard. It also offers the ability to provide hourly and location-based viewership trends as well as TRP (Television Rating Points) reports online. Databases that are considered in-memory, reporting and graphing libraries, and a J2EE-based application server make up the server infrastructure.

This study showed that the popularity of a television channel or program is measured in something called Television Rating Points (TRP). TRP is a figure that ranges from 1 to 100, with one rating point equaling 1percent of the population of a television audience that have been targeted. TRPs are usually measured from a target population utilizing statistical sampling procedures [13].

Television broadcasters, media companies, advertising agencies, and advertisers all value TRP ratings. It has a significant impact on ad spending and television show scheduling. Each year, media firms and organizations buy and sell air time that is worth billions of dollars.

Current automated TRP measurement methods rely on the placement of “People Meters” in the houses of the population that has been sampled. The People meters are connected devices that track viewing patterns and provide reports to a backend system on a regular basis. The following are the current methodologies employed by "People Meters."

1. **Matching of Audio** — In this example, the device records the audio content of a television program, compresses it, and sends it to a backend server. The audio samples are compared to stored program audio data on the server, and so viewership data is calculated. This strategy is difficult to deploy and necessitates significant backend system investments.
2. **Measurement of frequency** — This is used with analog transmission systems. Here, the People Meter monitors the frequency of the tuned TV channel and communicates the data to the server. The server associates the acquired frequency with channels and programs, allowing viewers' behavior to be tracked. This method is the most widely utilized in the country, but it is also the most prone to errors and inaccuracies because local cable service providers are not regulated or monitored in terms of transmission frequency.
3. **Watermarking** — Here, watermarks are added in the program feed at the broadcaster's end, and the People Meter detects them. The discovered watermark and timestamp are subsequently sent to the backend by the people Meters. This method has the disadvantage of requiring each aired program to be watermarked and active collaboration from all the related broadcasters.
4. **Visual recognition** — In this situation, the People Meters examines the displayed screen for visual patterns and images in order to determine the program being viewed. This approach is used in the suggested system.

Software like Trumedia [14], assists creators of advertisements in providing tailored adverts by assessing the present audience using technologies like video analytics on the faces of the audience taken by a camera, for example. The audience data is delivered to a dedicated server, which aids in the display of current audience reports. The real-time information about the audience can be utilized to pick the next message to play via TruMedia's interface with Cisco Digital Media Player, providing tailored advertising. Other participants in this industry include Quividi [15] and CognoVision [16].

### Review on Audience Measurement Technologies for user Centric Media

When compared to traditional media, how can you tell the difference between what the consumption is, and what the generation of content is? In a world that is full of new media. In a world where the user of a content could also be a creator/distributor. How appealing is it to the end users? Which business models that are both new and viable may be found in this context, and what is the market and technology's potential evolution? Perhaps it will never be possible to adequately answer these questions, but coming up with new, efficient and robust reference models for audience measurement in a new media world [17], and how they adapt to user-centric media through the use of combined metrics, are some of the most reassuring ways of achieving these goals. This paper outlines a possible system proposal that is based off of the afore mentioned reference model that can be applied in the new media world, which is then applied to user-centric media to provide answers to some of the questions stated above [17].

[2] Suggested that collecting important figures with the use of testbeds or panels of households or individuals to obtain trustworthy figures of media consumption is an important process in the media industry, particularly in the user-centric media. This is used to verify the various impacts and interests of service offerings, modern technological advancements, or even to predict business models that are both new and viable. Many firms may be hesitant to engage in new platforms without trustworthy consumption data, mitigating the development technologies for new media and services [17]. To a great extent not all of the features of audience measurement for traditional media can be applied directly to the user-centric media, where the user is allowed to distribute and consume both audio and visual content, the situation becomes even more complicated when the consumption or creation process occurs in user communities, where individual and collective consumption are equally important. This study proposed a robust system for acquiring the afore mentioned figures on both traditional and user-centric media channels. The following are the components of the system:

1. To begin, the presentation of a model for convergent media consumption in a range of terminals and networks, such as broadcasting to mobile and portable devices, or broadband IPTV distribution, will be done.
2. Second, the model's adaptation to user-centric media and not individual, but collective consumption is discussed.
3. Then, to integrate the results, a set of metrics is formalized.

To address user usage, many measures have been established to address possible program recommendations to users [18]. However, when users are linked in communities, one of the primary issues is the adaption of metrics for measuring audience interest and impact.

### Review on Investigation of Spectral Centroid Magnitude and Frequency that is Used for Speaker Recognition.

Mel-scale filter bank cepstrum coefficients (MFCC), Linear Prediction Cepstrum Coefficient (LPCC), and Perceptual Linear Prediction are some of the most common spectral envelope characterizations used in speaker recognition (PLP). Because of its popularity, the MFCC has become a de facto standard feature that is used for speaker recognition. Alternative features, such as frequency modulation (FM) and subband spectral centroid characteristics, have been proposed to transmit information other than the average subband energy [19]. The process of characterizing subband energy as a two-dimensional feature, consisting of Spectral Centroid Magnitude (SCM) and Spectral Centroid Frequency (SCF), was investigated in this paper. Empirical investigations utilizing SCF, SCM, and their fusion on the NIST 2001 and NIST 2006 databases reveal that the combination of SCM and SCF is somewhat more accurate than traditional MFCC, and that both fuse efficiently with MFCCs. We also show that frame-averaged FM features are fundamentally centroid features, and we present a SCF implementation that enhances both subband spectral centroid and FM feature speaker recognition performance [19].

[19] Suggested that speaker recognition relies on the separation of speaker dependent properties from speech signals, and because of anatomical and behavioral differences between participants, the speaker's vocal tract configuration has been found to be very speaker-dependent [20]. Mel-frequency cepstral coefficients are the most successful vocal tract-related acoustic characteristic (MFCC). However, information on the distribution of energy across the band is not efficiently captured during the MFCC extraction procedure. MFCC conveys the average energy of the subband as a single dimension for a subband speech stream (the overlapped triangular filters capture some information from neighbouring bands, but this can be considered an inter-band rather than an intra-band information). In this study, we look at how to turn this one-dimensional data into two-dimensional data that includes both the average energy and additional information on the energy distribution inside each subband. Phase or frequency related properties may be complimentary to MFCCs, according to research published in [21], [22]. The computational cost of applying frequency modulation (FM) extraction in actual applications is one issue [23]. The efficiency of frame-averaged FM components extracted using the second order all pole approach [21] on speaker recognition, as well as their complimentary nature to magnitude-based information, has recently been established [22]. When these frame-averaged FM components are compared to the deviation of the subband spectral centroid [24] from the subband's center frequency, as illustrated in Figure 1, it is clear that both the subband spectral centroid and the frame-averaged FM components provide identical information. Estimating the subband spectral centroid, on the other hand, is more efficient than estimating frameaveraged FM components. [24] shown that the formant-related information is carried by the spectral centroid frequency. It was also suggested that, while formant locations are resistant to additive noise, formant frequencies should not be employed as features directly due to the difficulty in estimating them accurately. Other features that convey formant related information, such as spectral centroid frequency, can be used to solve this problem, as shown in [24]. The use of subband spectral centroid in recent literature has demonstrated some success in noisy voice identification [. Spectral centroid frequency was previously employed in [24] for speech recognition. In contrast to FM features, spectral centroid frequency has recently been employed to enhance cepstral based features for speaker recognition [25]. The minor advantages over MFCC in speech recognition applications appear to be an oddity, given the similarity with frame-averged FM.This paper studied the efficiency of combining Spectral Centroid Frequency (SCF) and Spectral Centroid Magnitude (SCM) characteristics for speaker recognition, and showed how subband spectral centroid can be enhanced. SCM, like MFCC, conveys magnitude-related information, but SCF carries the SCM's frequency bias. The NIST2001 and NIST2006 speaker recognition datasets will be used to test these functionalities.

### Review on Estimating Audiences: Sampling in Television and Radio Audience Research

This paper put forward the argument that cultural consumption is problematic. It explained various angles to back up its argument and resolved that, certain responses from producers, regulators and observers are contingent on quantitative and qualitative consumption measurements. The data's trustworthiness varies greatly, not least because consumption is undetectable to those who would measure it in some locations, forcing them to come up with estimates that are based on assumptions about methodology and sample practices. Meanwhile, at auditoriums, turnstiles can correctly measure footfall through the doors, and the sale or return of certain sorts of publications drastically inspire inspires greater confidence levels in circulation figures, broadcasters play to intangible audiences who cannot be measured or witnessed en masse [26].

The propriety of sampling methodologies used to gather audience research data for the broadcasting industry, advertisers and programmers who require precise "knowledge" about their viewers, is discussed in this article. It is based on breaking down of Cultural Trends into contexts which looked at the argument over rival approaches for measuring consumption using either innovative technology gadgets or more traditional human recall. For individuals who utilize sampling techniques in the cultural sector, as well as those who would evaluate their results, the paper presents critical considerations [26].

This paper put forward the factors that affect audiences in the cultural sector. The following are some of the factors this paper put forward:

1. This size of the audience.
2. The demography of the audience.
3. The degrees of appreciation of the audience.
4. The nature of any involvement with presentation by the audience.
5. The possibility of members of the audience returning form more on future occasions.

The availability qualitative and quantitative data that are trustworthy, on consumption is critical to the success of many public and private ventures, but the degree of accuracy to which audiences can be measured varies greatly depending on the nature of the work being presented and the context in which the presented work is consumed [26]. When audience size has monetary implications, measurement accuracy is critical for a variety of stakeholders, including investors, producers, performers, exhibitors, and marketers. This paper argued that different people could be counted physically with a relatively high degree of accuracy, the advent of consumption necessitates consumption, that is, audiences entering and exiting premises, even if the personnel or technology required to do so on an automated basis may be too expensive for some organizations to fund from their budgets. The number of feature films that box office provides serves as data for comparative league tables, the contents of which can make or break the reputations of actors and directors. Cinemas are able to make returns to distributors based on the sales of different tickets, and the numerous feature films at the box office provides data for comparative league tables, the contents of which can make or break the reputations of actors and directors [26]. Similarly, sales of newspapers and magazines can be audited to a great degree of accuracy, with returns removed from gross sales statistics and genuine sales differentiated from promotional copies distributed to hotels and airlines. Even hits on a website could be counted physically, and data about the nature of these 'visitors' may be processed and distributed fast. One drawback: hit and footfall counts cannot always distinguish between repeat visitors and first-time visitors, so frequent visitors may skew the results because their characteristics outnumber those of the others.

This paper is divided into the following parts, with each part addressing some form of sampling:

1. Various sampling principles.
2. Using sampling in practice.
3. Right and Wrong uses of estimates from different sample data.
4. Carrying out sampling in crisis.
5. Selecting samples.

The nature of sampling as a method of estimating audiences' invisible consumption is that it is an inexact science. The theory and mathematics of sample orthodoxy may be valid, yet audience survey sampling appears to yield glaring contradictions that manifestly deny certain audience research the credibility to which it aspires. It's impossible to say whether such concerns are generalizable, but given the difficulty of constructing representative samples that report accurately, the probability that carrying out an audience research would give accurate estimates of viewing and listening by the population, as it is supposed to, is really slim [26].

## SUMMARY

In this chapter, a brief overview of audience research, audience measurement and audience related subjects have been discussed to provide a more rounded understanding of the processes involved in carrying out an audience measurement process. Furthermore, previous studies related to different audience monitoring technologies have been reviewed. From the studies reviewed above, it could be inferred that more efficient and less complicated methods need to be employed in carrying out audience measurement. Also, earlier methods used for carrying out audience monitoring were prone to human errors. The general opinion is that the best solutions audience measurement is to rely on more competent computers, thereby reducing human errors and limitations.

# SYSTEM ANALYSIS AND DESIGN

## INTRODUCTION

This chapter addresses the process involved in the development of a web-based radio audience measurement system. The various technologies used in the implementation of this project are discussed. UML diagrams are used where appropriate.

## APPLICATION DESIGN

## CLIENT-SIDE TECHNOLOGIES

The table below shows the various technologies that were used in the implementation of the front-end of the web application and mobile application.

|  |  |
| --- | --- |
| **TECHNOLOGY** | **PLACE USED** |
| JavaScript | Web and Mobile Application |
| HTML | Web Application |
| CSS | Web Application |
| React JS | Web Application |
| React Native | Mobile Application |
| Expo-av | Mobile Application |
| Redux | Web and Mobile Application |

### JAVASCRIPT

JavaScript is a high-level scripting or programming language that enables the implementation of complex features on web pages — whenever a web page does more than simply display static information for you to view — when it displays timely content updates, interactive maps, animated 2D/3D graphics, scrolling video jukeboxes, and so on — JavaScript is involved.

The following are some advantages of JavaScript:

1. According to Stack Overflow, JavaScript is the most used programming language on the planet, which makes it an excellent option for programmers. JavaScript enables the creation of excellent front-end and back-end software via the use of several JavaScript-based frameworks such as jQuery, Node.JS, and others.
2. JavaScript is ubiquitous; it is pre-installed on every contemporary web browser, and therefore no specific environment configuration is needed to learn JavaScript. For instance, Chrome, Mozilla Firefox, and Safari all have JavaScript.
3. JavaScript enables the development of truly standard, scalable, and responsive online applications. It enables the creation of rich web apps with high-quality user interfaces and experiences.
4. JavaScript is currently used in the creation of mobile applications, desktop applications, and games. This has increased the number of apps available for the JavaScript language.
5. JavaScript has a sizable support community, with people from all around the globe continuously working to enhance the language's capabilities.
6. JavaScript is a framework-rich language with a large number of pre-developed frameworks and libraries that may be utilized directly in software development to significantly decrease development time.

The following are some features of JavaScript:

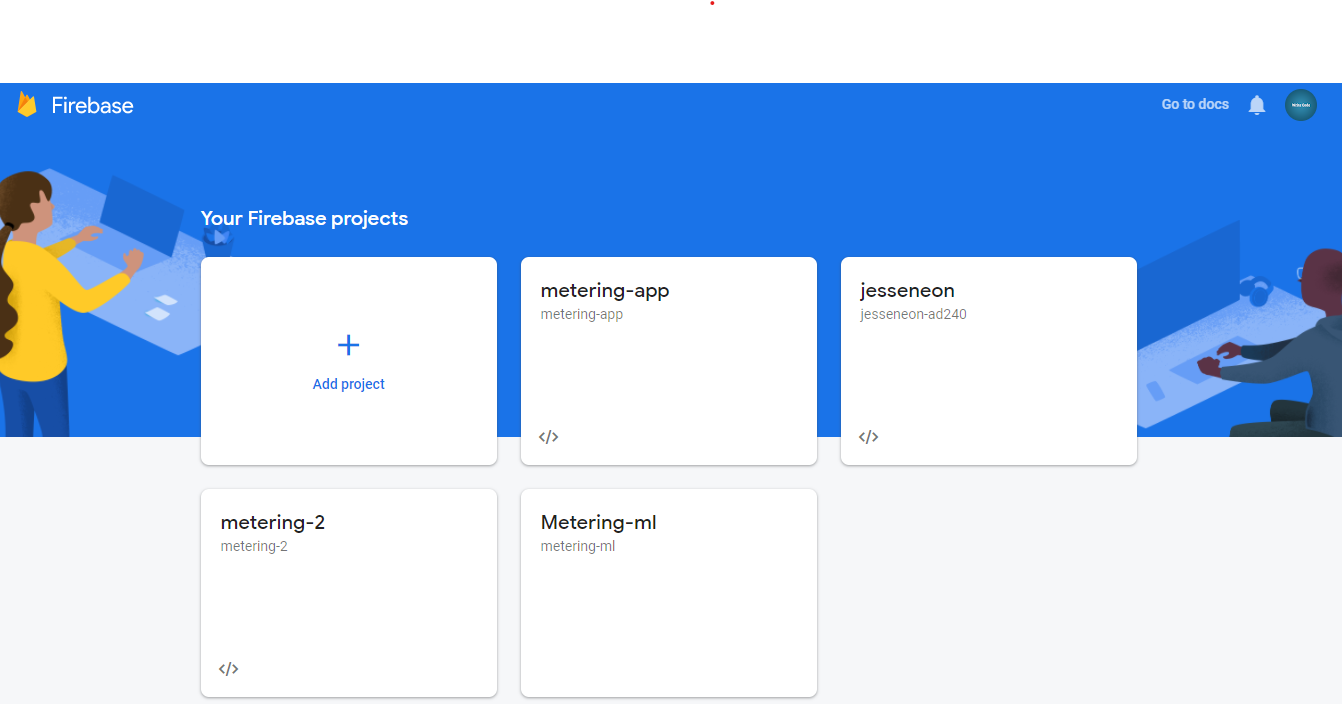
1. JavaScript is a scripting language that is object-oriented.
2. It provides users with more control over the browser.
3. It controls date and time, it does this by detecting the user's browser and operating system.
4. It is very light in weight.
5. JavaScript is a scripting language that is interpreter-based.
6. JavaScript has a case-sensitive syntax.
7. JavaScript includes predefined objects that can be used at any point with a JavaScript file.
8. In JavaScript, each statement must be ended with a semicolon (;).
9. The grammar of the majority of JavaScript control statements is identical to the syntax of control statements in the C language.
10. The ability to define new functions inside scripts is a critical feature of JavaScript. In JavaScript, the function keyword is used to declare a function.

### HYPER TEXT MARK-UP LANGUAGE (HTML)

## SERVER-SIDE TECHNOLOGIES

### FIREBASE

Firebase is a mobile application development platform birthed by Google that supports in build, develop, and advance application. Firebase offers one of the layers of cloud computing, which is Backend-as-a-service. The firebase console is displayed in the figure below.



Firebase is a real-time database, file storage, provides user authentication services, and it can be used to host services for static files and data, the Figure 3.2 below shows a complete firebase suite.

Graphical user interface, website

Description automatically generated

Build Suite

Graphical user interface, application, website

Description automatically generated

Engage Suite

Traditionally, most databases require making HTTP calls to get and synchronize data to gadgets, and they only give data only when queried for it. An app associated with firebase is not interfacing through normal HTTP, but instead, it is interfacing through a WebSocket. WebSockets offer a persistent connection between a client and server that the two players can use to exchange data whenever in real-time. The client sets up a WebSocket connection through a procedure known as the WebSocket handshake. WebSocket API presents an advanced technology that makes it feasible for a two-way interaction between the client's program and a server. With this API, the app can converse with a server and get event-driven responses without surveying the server for an answer. WebSockets is faster than HTTP. The apps do not need to make individual WebSocket calls since one attachment connection is the length. The entirety of the data syncs automatically through that single WebSocket relying upon the quality of the network available. The Figure below shows a graphical explanation.

Diagram

Description automatically generated

One of the advantages of Firebase application is that it remains responsive even when offline because the Firebase Realtime Database SDK stores data to disk. Once connectivity is restored, the client device receives any changes it missed, synchronising it with the current database server state.

For the implementation of this project, Firebase would be used for the following:

1. **Authentication:** This is controls operations like user sign-up and user sign-in.
2. **Data Storage:** With the use of firebase’s cloud fire store service, changes within the database of the system can be rendered on the frontend in real-time. Cloud firestore is a NoSQL database.
3. **Media File Storage:** This uses firebase’s storage to store audio files sent from the mobile application.

Chart, diagram

Description automatically generated

FlowChart for cloud Firestore

## MACHINE LEARNING TECHNOLOGIES

SOMETHING