UNIVERSITY OF BUEA

Buea, South West Region

Cameroon

P.O. Box 63,

Tel: (237) 3332 21 34/3332 26 90

REPUBLIC OF CAMEROON

PEACE - WORK - FATHERLAND

FACULTY OF ENGINEERING AND TECHNOLOGY

GROUP 7

Task2

REQUIREMENT GATHERING AND DATA CLEANING FOR A SMART ATTENDANCE MONITORING SOLUTION USING FACIAL RECONITION AND GEOFENCING

Course: CEF440: Internet Programming and Mobile Programming

NAME	MATRICULE
FONYUY VERENA MONYUYTA-AH	FE22A220
KENFACK DONJIO ABEL BRUNEL	FE22A380
NSONDO MIRELLE NYISEKINYI	FE22A283
TATA THECLAIRE GHALANYUY	FE22A310
UNJI STEPHEN UKU	FE22A323

2024/2025 ACADERMIC YEAR

Table of Content

Introduction

In today's academic landscape, ensuring accurate and seamless attendance tracking is critical for both administrative efficiency and student accountability. Traditional attendance methods such as manual roll calls or paper sheets are prone to manipulation, inefficiency, and human error. As part of our mission to modernize academic operations, **Group 7** initiated the development of a **Smart Attendance Tracking System** integrating **geofencing** and **facial recognition technologies**.

This report documents our **requirement gathering process**, which served as the foundation for designing a system that is not only technologically robust but also aligned with the real-world needs of students, educators, and administrative staff. To ensure a user-centered approach, we employed a combination of **surveys, interviews, focus groups, brainstorming sessions, reverse engineering,** and **prototyping**. Through this multipronged strategy, we were able to uncover expectations, constraints, and challenges faced by users, which directly informed the design and feature set of the proposed system.

Furthermore, we carried out a thorough **data cleaning** process to prepare the collected information for meaningful analysis. This ensured the accuracy and consistency of insights drawn from the data, allowing us to define user needs and technical requirements with clarity and precision.

I: Stakeholder Identification

1. Introduction

Stakeholder identification is the first and most critical step in the requirements engineering process for any software or mobile application development project. A stakeholder is any person, group, or entity that has an interest, influence, or investment in the outcome of the application. Identifying the right stakeholders ensures that the development process is aligned with user needs, business goals, technical constraints, and regulatory compliance.

Failure to properly identify stakeholders can lead to missed requirements, misunderstandings, project delays, or the development of an application that ultimately fails to meet expectations.

2. What is a Stakeholder?

- Is affected by the project or application.
- Has influence over the project (positively or negatively).

- Provides resources for the project (time, money, technical input).
- Uses or interact with the final product

  In mobile application development, stakeholders come from both technical and non-technical domains, including business owners, developers, end-users, support teams, marketing units, and even regulatory bodies.

3. Importance of Stakeholder Identification

Identifying stakeholders ensures:

- Requirements are collected comprehensively and accurately
- Risks of project failure due to misalignment are reduced
- Stakeholder concerns are addressed proactively
- Higher levels of project acceptance and user satisfaction
- All perspectives are considered in decision-making

4. Tools and Techniques Used

To identify and analyze stakeholders effectively, we used the following techniques:

- **Brainstorming sessions** within our project team to list potential user groups.
- **Stakeholder Mapping** using the Power-Interest Grid to visualize and prioritize stakeholders.
- **RACI Matrix** to define the roles and responsibilities of different groups.

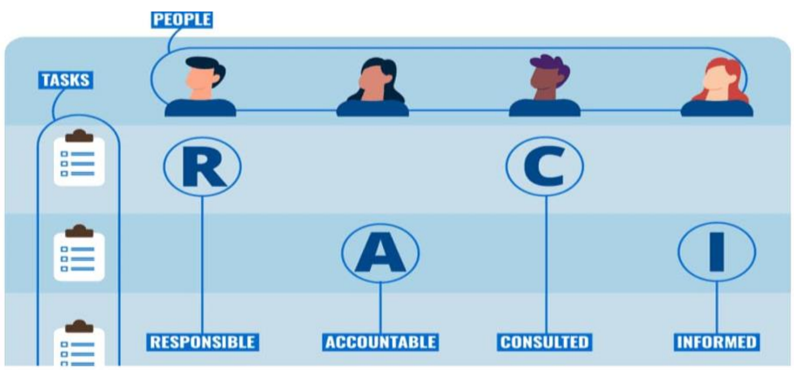


Fig 1: Stakeholder Matrix (RACI)

- **Stakeholder Interviews and Faculty Feedback** to validate assumptions and discover hidden expectations.

We used collaborative tools such as shared spreadsheets and diagrams to maintain a living stakeholder register throughout the development phase.

5. Stakeholder Identification Process

We adopted a structured approach that included the following steps:

- **Analyzing the project scope** to understand the application's users, features, and impact.
- **Listing potential stakeholders** based on institutional structure and system usage scenarios.
- **Categorizing stakeholders** by role, influence, and interest using models such as the Power-Interest Grid.
- **Engaging with stakeholders** through interviews and surveys to understand their expectations and concerns.

  Visual tools to categorize stakeholders based on influence and interest.

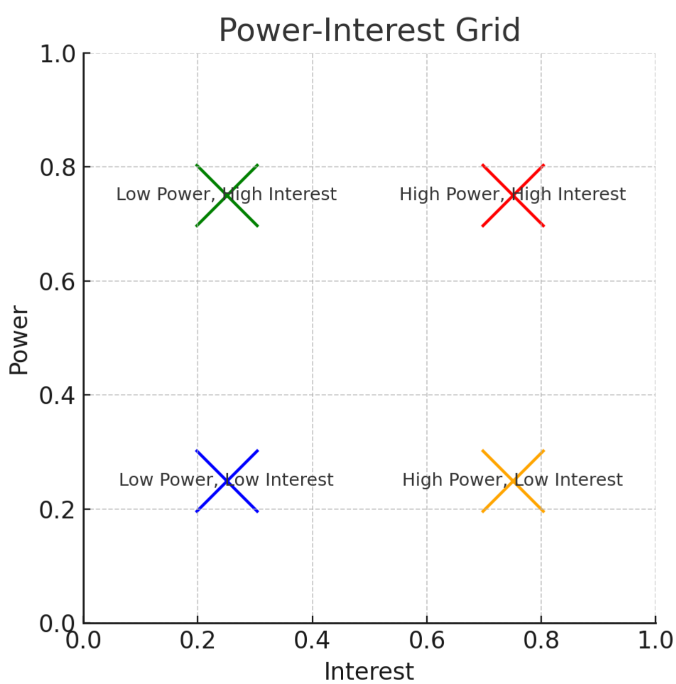


Figure 2: Power-Interest Grid showing stakeholder categorization.

6. Identified Stakeholders and Their Roles 👥

Stakeholder	Role	Interest Level	Responsibilities/Concerns
Students	End users	High	Use the app to check in for classes; ensure face data is secure and app is easy to use.
Lecturers	Secondary users and validators	High	Monitor attendance records, validate unusual entries, and generate attendance reports.
IT Department	Technical stakeholders	High	Responsible for system deployment, backend infrastructure, and integration with school LMS.
University Admin	Key stakeholders/decision makers	High	Approve the system, allocate budgets, ensure legal compliance, and enforce its usage policy.
Parents/Guardians	Indirect stakeholders	Medium	May view attendance reports; expect reliability and transparency.
Data Protection Office	Regulatory and compliance authorities	High	Ensure compliance with privacy laws such as data consent, storage, and usage of biometrics.
Face Recognition API Providers	Third-party services	Medium	Provide face detection and matching services; reliability and uptime are critical.
Geolocation Services	Infrastructure providers	Medium	Provide GPS and geofencing data; crucial for verifying presence in the target location.

7. Challenges Faced in Stakeholder Identification

During the stakeholder identification phase, we encountered a few challenges, including:

- Overlooking less obvious or indirect stakeholders like parents and service API providers.
- Misjudging the level of influence of some users (e.g., the IT department had more influence than initially assumed).
- Communication barriers during early interviews due to technical jargon.
- Resistance from some faculty members who were initially skeptical about using GPS and facial data for attendance.

These challenges were addressed through iterative stakeholder engagement and incorporating feedback into the system design.

8. Best Practices adopted

To ensure effective stakeholder management, we:

- Started stakeholder identification early in the planning phase.
- Used a combination of qualitative (interviews) and quantitative (surveys) methods.
- Updated the stakeholder register continuously as roles and expectations evolved.
- Created stakeholder personas to empathize with their concerns during UX design.
- Maintained regular communication with high-interest stakeholders (faculty and students) during prototype development.

Stakeholder identification was a foundational activity that informed the rest of our requirement gathering process. By actively engaging with stakeholders from diverse categories, we were able to align our system design with real-world expectations and regulatory requirements. This structured approach significantly reduced the risk of feature misalignment and increased the potential for successful adoption upon deployment.

II. Requirement Gathering Techniques

1. introduction

Requirement gathering techniques are the cornerstone of successful system development, providing a structured approach to understanding stakeholder needs, goals, and constraints. These techniques ensure that the final system is not only functional but also aligned with the expectations of its users and stakeholders.

In the context of designing a mobile-based attendance management system leveraging **geofencing** and **facial recognition,** requirement gathering techniques help bridge the gap between abstract ideas and actionable development plans. They facilitate the identification of specific system functionalities, such as real-time location-based attendance verification and secure biometric data handling, while addressing challenges like user acceptance and data accuracy.

Ultimately, effective requirement gathering not only saves time and resources during the development phase but also results in a system that meets the unique needs of its users, fosters trust, and drives successful adoption.

2. Requirement gathering techniques used to gather for this project

2.1 Surveys and Questionnaires

This technique involves distributing surveys or questionnaires to a large audience to collect quantitative and qualitative data about their requirements and preferences. This technique allows for the gathering of quantitative and qualitative data, which can be easily analyzed for insights

Distribution surveys and questionnaires to a broad audience to collect information on a larger scale. This technique is useful for gathering feedback from a diverse set of stakeholders and can be particularly effective in large projects.

Key areas targeted included:

- **Current Challenges** with attendance (e.g., time wastage, proxy marking).
- **Technology Familiarity**: Comfort levels with GPS, face ID, app usage.
- **Willingness to Adopt Biometric Systems**: Acceptability of facial recognition.
- **Data Privacy Concerns**: Understanding of how their data would be handled.

Survey Design Principles used

1. **Clarity**: All questions were clear and concise to avoid confusion.
2. **Relevance**: Questions focused on aspects directly related to the system, such as geofencing, facial recognition, and user interaction questions.
3. **Anonymity**: Respondent anonymity was maintained to encourage honest feedback.
4. **Balance**: Included a mix of quantitative (close-ended) and qualitative (open-ended) questions.
5. **Categories**: Questions were grouped into categories based on stakeholder type (students and administrators) and their interaction with the system.

Survey Distribution

Platforms:

- Online: Google Forms
- Offline: Paper-based surveys distributed during meetings.

Target Audience:

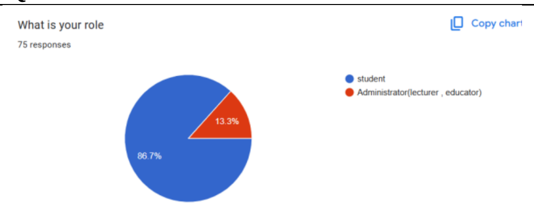
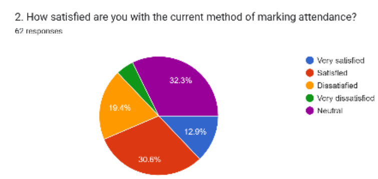
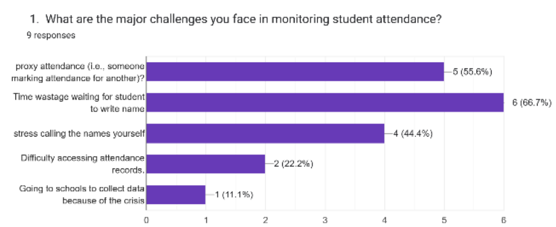
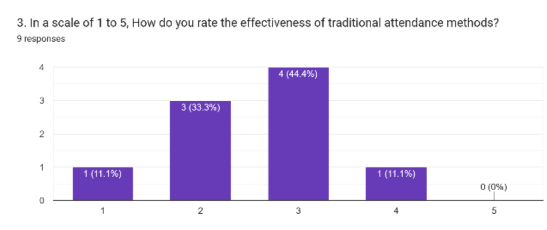
Students, instructors, and administrators across different departments and levels.

Response Collection Period:

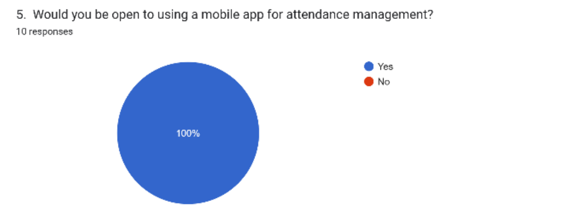
deadline: one weeks to ensure timely responses.

Key Insights

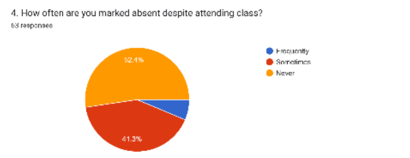
- Over 87.1% supported the concept of mobile-based attendance.
- Major concerns included **privacy**, **location spoofing**, and **biometric accuracy**.
- lecturers requested **reporting**, **manual override**, and **export functions**.

<table>
<tr><td>PARTICIPANTS ROLE</td><td>QUESTION & RESPONDE ANALYSIS</td></tr>
<tr><td>STUDENTS AND EDUCATORS</td><td>
</td></tr>
<tr><td>LECTURERS & EDUCATORS</td><td>
</td></tr>
<tr><td>LECTURERS & EDUCATORS</td><td>
</td></tr>
<tr><td>LECTURERS & EDUCATORS</td><td>
</td></tr>
</table>

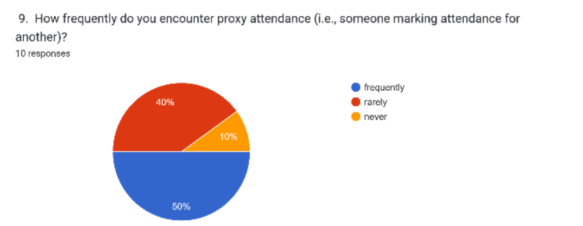
LECTURERS & EDUCATORS



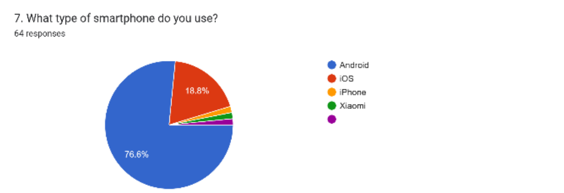
STUDENTS



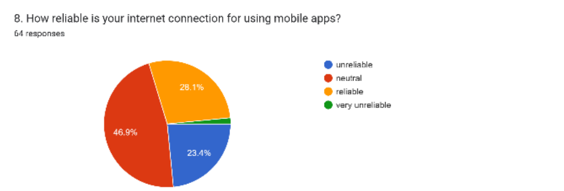
LECTURERS & EDUCATORS



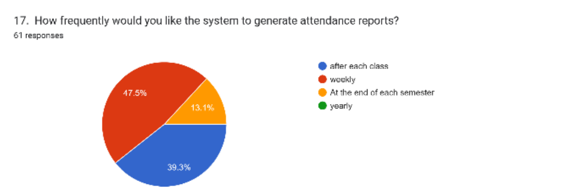
STUDENTS



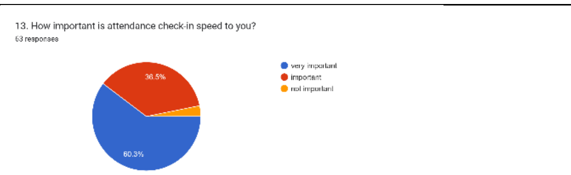
STUDENTS



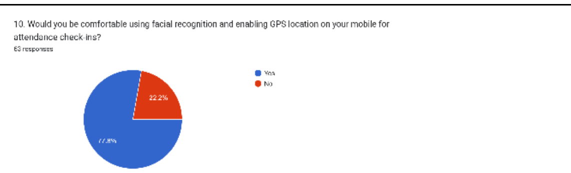
STUDENT

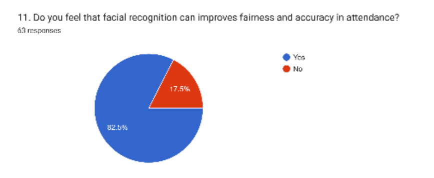


STUDENT



STUDENT



STUDENT	

2.2 Interviews

Technique overview: Interviews involve one-on-one conversations with stakeholders or group interviews to gather detailed information about the project requirements. Preparation is crucial for maximizing the effectiveness of interviews, allowing for the collection of comprehensive and relevant information.

The purpose of the interview is to collect in-depth insights from stakeholders about their needs, challenges, and expectations regarding the mobile-based attendance management system. This qualitative method allows for a deep understanding of personal experiences, preferences, and ideas that can guide system design and implementation.

Interview Design Principles used

1. **Specificity**: Tailored questions to the role and interaction of the stakeholder with the system.
2. **Open-Ended Questions**: participants were encouraged to share detailed responses.
3. **Structured Format**: focus was maintained while allowing for flexibility to explore unanticipated insights.
4. **Confidentiality**: Ensure that participants feel comfortable sharing their honest opinions.

Interview Process

1. Pre-Interview Preparation

Setting:

- Interviews were conducted in stakeholder present environment

Materials:

- an interview guide with predefined questions.
- recording devices (with participant consent) to ensure accurate documentation.

Briefing:

- The purpose of the interview and how the data will be used.

2.3 Brainstorming

Technique overview: Brainstorming sessions encourage the generation of a wide range of ideas in a short period, usually involving a group of stakeholders. This technique is ideal for initiating the requirements gathering process.

Purpose of Brainstorming

Brainstorming is a collaborative technique used to generate innovative ideas, identify challenges, and propose solutions. In the context of developing a mobile-based attendance management system with geofencing and facial recognition, brainstorming sessions involve stakeholders and project team members to explore the requirements, features, and potential implementation strategies for the system. This technique fosters creativity and ensures diverse perspectives are considered during the requirement gathering phase.

Steps used in this technique Brainstorming

1. Preparing for the Session

Define Objectives:

- The purpose of was brainstorming clear before session (e.g., identifying system features, addressing user concerns).

Identify Participants:

- We Include key stakeholders such as students, instructors, administrators.
- We Ensure diversity in expertise and perspectives.

Set Ground Rules:

- Encouraged open participation and discourage criticism during idea generation.
- Limited individual contributions to ensure equal participation.

2. Conducting the Brainstorming Session

- **Session Structure**
  - **Introduction (10 minutes)**:
    - participants were briefed on the project and goals of the session.
    - brainstorming process and ground rules were explained to participants.
  - **Idea Generation (30–40 minutes)**:
    - Use of techniques such as mind mapping, silent brainstorming, or round-robin discussions.

        - participants were encouraged to share ideas freely, no matter how unconventional.
    - **Discussion and Refinement (20–30 minutes)**:
        - similar ideas were and evaluate their feasibility and relevance.
        - Identify ideas that address critical requirements or challenges.

Techniques for Idea Generation:

1. **Mind Mapping**:
    - Started with a central theme (e.g., "Attendance System Requirements").
    - sub-themes like geofencing, facial recognition, security, user interface, etc are branched out.
    - Participants were encouraged to expand each branch with specific ideas or challenges.
2. **Silent Brainstorming**:
    - Participants were allowed to write down their ideas individually before sharing them.
    - All ideas are and discuss them collectively to ensure equal representation.
3. **Round-Robin Technique**:
    - Each participant contributes one idea in turn, ensuring that everyone has a chance to share.
    - The round is repeated until no new ideas emerge.

3. Key Questions for Brainstorming

These guiding questions can help steer brainstorming discussions toward actionable insights:

Geofencing Features*:*

- What methods can be used to define virtual classroom boundaries accurately?
- How can the system handle scenarios like weak GPS signals or overlapping geofences?
- What are potential challenges in using geofencing for attendance tracking?

Facial Recognition Features*:*

- What level of accuracy is expected for facial recognition?
- How can privacy concerns related to biometric data be addressed?
- What are potential challenges with facial recognition in different lighting conditions?

User Interface and Experience*:*

- What design elements would make the app intuitive and user-friendly?
- How can the system accommodate users with varying levels of technical proficiency?
- What notifications or reminders should the system include to enhance usability?

Data Privacy and Security:

- What measures should be taken to protect biometric and geolocation data?
- How can the system ensure compliance with data protection regulations?
- What encryption methods can be used to secure attendance data?

General Questions:

- What other features should the system include (e.g., attendance history, performance analysis)?
- What challenges might stakeholders face in adopting the new system?
- How can the system integrate with existing academic tools or infrastructure?

Advantages Brainstorming contributed to the project

- Encouraged creativity and out-of-the-box thinking.
- Promoted collaboration among diverse stakeholders.
- Identified potential challenges and solutions early in the process.

Challenges faced in Brainstorming

- lack focus without effective facilitation.
- Dominant participants overshadow others if rules are not enforced.
- Requires careful documentation to capture all ideas.

2.4 Reverse Engineering for Mobile-Based Attendance Management System

Introduction

Reverse engineering is the process of deconstructing and analyzing existing systems to understand their architecture, functionality, and design principles. In the context of developing a mobile-based attendance management system using geofencing and facial recognition, reverse engineering can provide valuable insights into the strengths, weaknesses, and opportunities for improvement in similar systems. This technique is particularly useful for identifying best practices, optimizing performance, and ensuring security and compliance.

Applications of Reverse Engineering to the project

1. Understanding Existing Systems

To establish a solid foundation, we analyzed existing attendance management systems to understand their architecture and workflows. This included:

- **Analyzing Source Code**: Where legally accessible, we decompiled the source code of similar applications to examine algorithms and data structures.
- **Examining Features**: We identified key features such as geofencing accuracy, facial recognition reliability, and user interface design.
- **Documenting Findings**: Detailed documentation of the systems' architecture, workflows, and functionalities was created.

2. Geofencing Implementation Analysis

Geofencing was a critical component of our system, enabling location-based attendance verification. We analyzed and achieved the following:

- **GPS Data Utilization**: We examined how GPS data was processed to define virtual boundaries effectively.
- **Accuracy Evaluation**: Edge cases such as weak signals or overlapping geofences were studied, ensuring robust solutions.
- **Performance Optimization**: We improved geofence reliability and responsiveness for seamless user experiences.

3. Facial Recognition Technology Analysis

To ensure secure and accurate user identification, we delved into facial recognition technology by:

- **Analyzing Machine Learning Models**: We studied the algorithms and training datasets to optimize facial recognition accuracy.
- **Feature Extraction Analysis**: Techniques for detecting and matching facial features were reviewed and refined.
- **Security Measures Evaluation**: Biometric data encryption methods were carefully assessed to mitigate vulnerabilities.

4. Security Assessment

Ensuring the protection of sensitive data was paramount. Our approach included:

- **Data Encryption Analysis**: Biometric and geolocation data storage methods were scrutinized for compliance with security standards.

- **Spoofing Risk Assessment**: We tested the system's resilience against spoofing attacks.
- **Access Control**: Authentication mechanisms were evaluated to prevent unauthorized access.

5. *Performance Optimization*

We identified and resolved bottlenecks to enhance the system's performance:

- **Concurrent Processing**: The system's ability to handle multiple check-ins and large datasets was improved.
- **Module Integration**: Seamless interaction between geofencing and facial recognition modules was achieved.

6. *User Experience (UX) Enhancement*

Our focus extended to improving the system's usability:

- **Interface Design**: Pain points in the user interface were addressed through redesign and testing.
- **Feedback Mechanisms**: Incorporation of user feedback into updates ensured the system remained user-friendly.

7. *Compliance and Ethical Considerations*

We prioritized adherence to legal and ethical standards:

- **Data Handling Compliance**: Verification of compliance with data protection regulations such as GDPR was conducted.
- **User Consent Management**: Processes for obtaining and managing user consent were refined.

Similar Systems

1. SMART Attendance Management System

Advantages:

- Combines geofencing and facial recognition for accurate attendance tracking.
- Reduces manual intervention and proxy attendance.

Disadvantages:

- May face challenges with weak GPS signals and facial recognition accuracy.

Improvements:

- Implement fallback mechanisms for GPS issues and enhance facial recognition models.

2. GPS-Based Attendance System

Advantages:

- Uses geofencing for automated attendance tracking.
- Eliminates the need for physical sign-ins.

Disadvantages:

- Limited to geolocation data; does not verify identity.

Improvements:

- Integrate facial recognition for identity verification and use hybrid location tracking methods.

3. Location-Based Attendance Tracking with Facial Recognition

Advantages:

- Combines geofencing and facial recognition for seamless attendance tracking.
- Reduces resource expenditure by using mobile devices.

Disadvantages:

- May encounter challenges with lighting conditions and device compatibility.

Improvements:

- Optimize facial recognition algorithms and ensure compatibility across devices

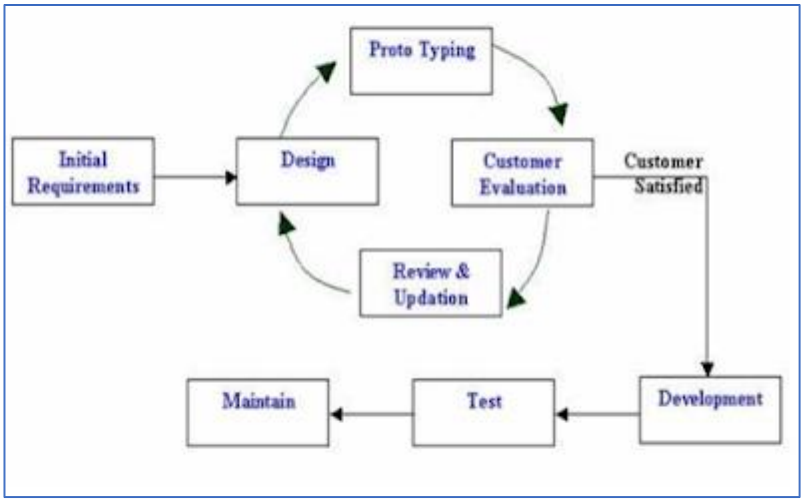


Fig 3: reverse engineering process

Enhancing the Attendance Management System based on systems reviewed using reverse engineering

1. Hybrid Location Tracking

We aim to enhance geofencing accuracy by incorporating:

- **GPS Technology**: For precise outdoor geolocation.
- **Wi-Fi Positioning**: To improve accuracy in urban areas or areas with weak GPS signals.
- **Indoor Positioning Systems**: For accurate tracking within buildings, leveraging technologies like Bluetooth beacons.

2. Advanced Facial Recognition

To ensure reliable identification, our planned improvements include:

- **Leveraging Deep Learning Models**: Training models on diverse datasets to handle a wide range of facial features and expressions.
- **Improving Detection Accuracy**: Addressing potential challenges like low lighting, facial obstructions, and varying angles.
- **Enhancing Model Efficiency**: Striving for a balance between recognition precision and processing time.

3. Robust Security

Given the sensitivity of user data, we are focusing on robust security measures such as:

- **End-to-End Encryption**: To secure data during transmission and storage.
- **Multi-Factor Authentication**: Adding additional layers of security to prevent unauthorized access.
- **Continuous Monitoring**: To detect and mitigate potential security threats proactively.

Insights & Actions

Area	**Lessons from Reverse Engineering**	**Actions Taken**
Geofencing	GPS-only caused issues indoors	Added Wi-Fi & indoor positioning
Facial Recognition	Lighting impacted accuracy	Improved ML model & pre-processing
Security	Weak encryption in some apps	Used AES-256 encryption & MFA
User Experience	Complex menus discouraged usage	Designed streamlined UI
Performance	Slow check-ins in older devices	Added offline mode & batch check-ins

2.5 Focus Groups

Overview

Focus groups involve bringing a **small group of stakeholders** together for a guided discussion. It's similar to interviews but allows **participants to build off each other's ideas**, surfacing deeper insights and shared frustrations or expectations.

Purpose

To gather detailed opinions, priorities, and emotional responses from students, instructors about existing attendance systems and the proposed mobile-based solution.

Process

- Groups of 6–8 participants were formed by role (e.g., "instructor group," "student group").
- A facilitator led discussions using a semi-structured guide.
- Topics included:

- Ease of use
- Trust in biometric systems
- Concerns about tracking/location
- Preferences for notifications and features

Key Takeaways

- Students favored **self-check-in with restrictions** (geofencing/facial recognition) over classroom roll calls.
- Instructors wanted **real-time alerts** for missed attendance or anomalies.

Benefits

- Uncovered **emotional and behavioral factors** (trust, acceptance)
- Useful for refining **user interface** decisions
- Encouraged collaborative idea development

Challenges

- Risk of **groupthink** if one voice dominates
- Requires **skilled facilitation**

5. Prototyping (Low-Fidelity UI Sketching)

Overview

Prototyping involves creating **early visual representations** of the system, such as screen sketches or mockups, to validate functional requirements and user expectations.

Purpose

To validate **UI/UX requirements** and gather feedback before building the actual app.

Process

- Paper-based and Figma mockups of screens were shown to stakeholders.
- Scenarios were simulated:
    - "How would you check in?"
    - "How would you know your attendance was confirmed?"
- Feedback was documented and converted into actionable requirements.

Findings

- Users wanted **minimal clicks** to complete check-in.
- Suggested addition of **feedback indicators** (“check-in success” screen).
- Instructors requested color-coded attendance indicators.

Benefits

- Immediate feedback on design choices
- **Reduces development rework**
- Makes abstract ideas **tangible for non-technical users**

Challenges

- Needs multiple iterations
- Can cause users to fixate on design instead of functionality

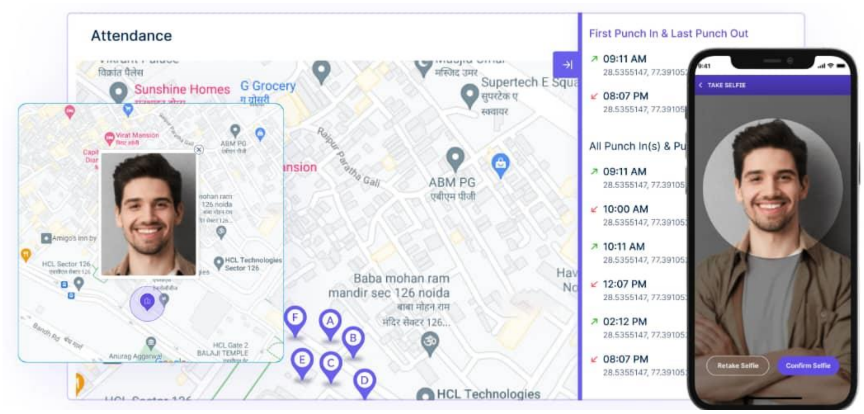


Fig 4: Visual reference prototype

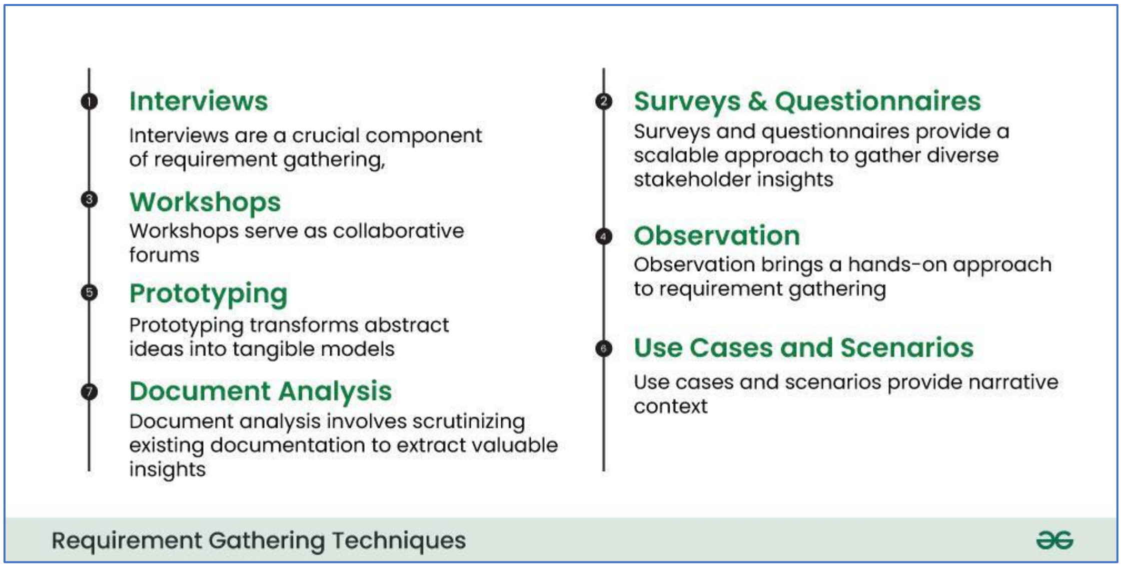


Fig:Requirement gathering technique summary image

Prioritization Matrix (MoSCoW Expanded)

Requirement	**Priority**	**Justification**
Facial recognition check-in	Must Have	Core to the system concept
Geofencing with radius control	Must Have	Prevents remote check-ins
Attendance dashboard	Should Have	Enhances instructor experience
Offline check-in mode	Could Have	Useful, but not critical
Face spoof detection	Must Have	Prevents misuse
App themes (light/dark)	Could Have	UX enhancement
Blockchain-based data storage	Won't Have	Overkill for MVP version

3. Cross-Technique Comparison Table

Technique	**Strengths**	**Contribution to Project**
Surveys	Large-scale data collection	Validated overall interest and identified priority concerns
Interviews	In-depth qualitative insights	Provided role-specific needs and expectations

Technique	Strengths	Contribution to Project
Brainstorming	Collaborative ideation	Generated innovative solutions and anticipated implementation risks
Reverse Engineering	Technical benchmarking	Helped in refining geofencing, facial detection, and security models

The integration of multiple requirement gathering techniques surveys, interviews, brainstorming, and reverse engineering ensured a holistic understanding of user expectations, system constraints, and technical possibilities. Each method contributed uniquely:

- **Surveys** quantified general demand and concerns.
- **Interviews** provided a humanized view of user experience.
- **Brainstorming** fostered collaborative problem-solving and idea refinement.
- **Reverse engineering** ensured competitive analysis, technical robustness, and security awareness.

This comprehensive and well-structured approach lays a solid foundation for the development of a secure, efficient, and user-friendly mobile-based attendance system that harnesses both geofencing and facial recognition technologies.

III. Data Gathering

Data gathering is a critical component in the requirements phase, especially for a system integrating facial recognition and geofencing technologies. This section details the data sources, collection procedures, and structure of the datasets used to support the design and implementation of the mobile-based attendance management system. The primary data components collected include:

- Facial Recognition Training Data
- School Attendance Record History Format
- Geolocation Data for Geofencing (in progress)

1. Facial Recognition Training Data

1.1 Objective:

To collect diverse and realistic image data to train a facial recognition model that can handle varying environmental and physiological conditions encountered in real-world classroom settings.

1.2 Dataset Overview

- Number of Images: Approximately 140
- Number of Subjects: 1 (serves as prototype data)
- Purpose: To simulate variability in real-world student appearance and environmental factors.
- Format: JPEG/PNG images
- Captured With: Smartphone (with front camera), Ring Light, Flashlight
- Angles & Expressions Considered:
    - Neutral face (not smiling), lifted face (smiling)
    - Straight, angled (left, right), tilted up/down
    - Eye-wear variations (with and without lenses)

1.3 Environmental Conditions Simulated:

Lighting Condition	**Description**
Daylight - Outdoor (Clear)	Full face exposure to sunlight
Daylight - Partial Shade	One side shaded by sunlight
Indoor - Bright White/Yellow Light	Artificial ring lights with adjustable hue
Indoor - Dim Light	Natural low-light conditions (early morning beside window)
Indoor - Darkness with Flash	Camera flash used in total darkness

Lighting Condition	Description
Complete Darkness	Captured in pitch-black environment
Low Light with Shadows	Slightly dark room with minimal natural light
Unlit Corner Simulation	A room with just a bit of darkness
Motion Blur Simulation	Shaky hands during timed capture for blurry image variations

clear **blur** **dark**

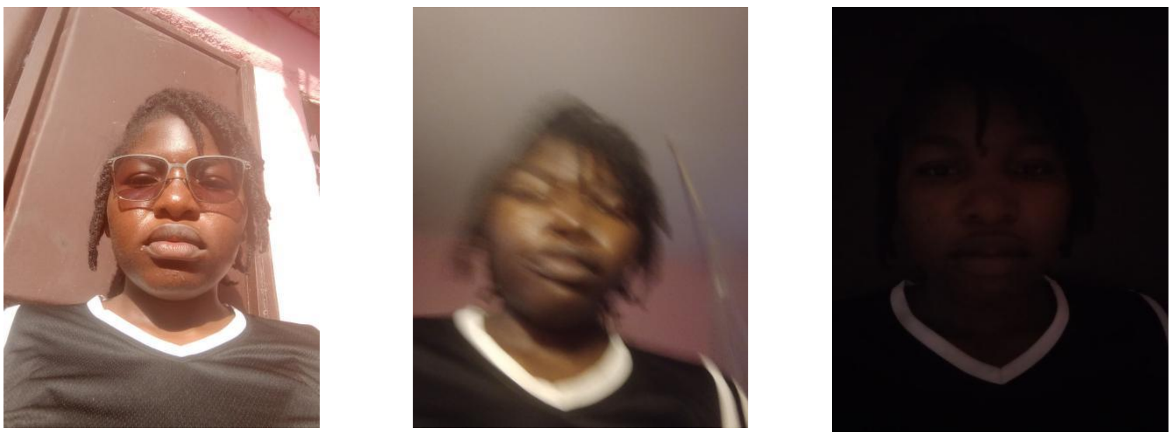


FIG 5: Different lighting conditions image sample

1.4 Data Use Consideration:

This prototype dataset will serve as the basis for training the facial recognition model. The aim is to prepare the system to recognize faces under diverse real-world conditions and improve tolerance for image variability.

2. School Attendance Record History Format

2.1 Objective:

To standardize and model the database schema for recording student attendance, supporting both administrative integrity and app-level functionality such as filtering, querying, and reporting.

2.3 Data Structure:

The table below summarizes the key fields and types involved in the attendance record schema.

Field Name	Data Type	Description
attendance_id	UUID / String	Unique identifier for each attendance record
matric_no	String	Unique matriculation number of the student
student_name	String	Full name of the student
course_code	String	Code of the course or class (e.g., CSC201)
class_date	Date (YYYY-MM-DD)	Date the attendance was taken
check_in_time	Time (HH:MM:SS)	Timestamp of when student checked in
check_out_time	Time (HH:MM:SS)	Timestamp of when student checked out (optional)
status	Enum/String	Attendance status: Present, Absent, Late, Excused
location_verified	Boolean	Was geofence verification successful? true / false
face_verified	Boolean	Was facial recognition successful? true / false
device_id	String	(Optional) Identifier for the device used to check in
remarks	Text	(Optional) Additional notes (e.g., "Face match failed", "Late arrival")

3. Geofencing Data for Geofencing

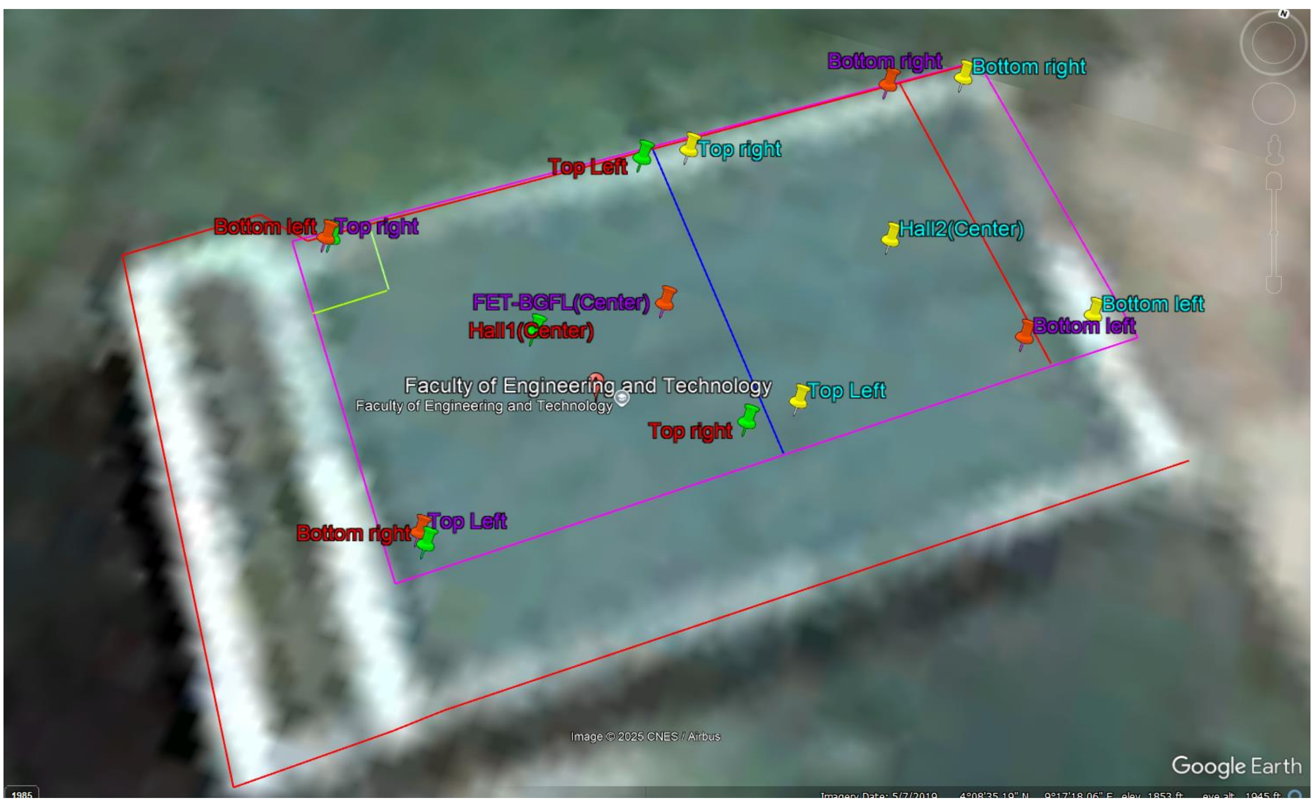
3.1 Objective

The objective is to collect accurate geospatial coordinates for the most frequently used classrooms within the Faculty of Engineering and Technology and the FET Technology Block. These coordinates will serve as reference points for the geofencing system, enabling precise verification of student location during the attendance process.

3.2 Tools and Method Used

- **Measuring Tape:**
  Utilized to measure the physical dimensions of the classrooms to establish scale and verify the relative spatial extent of each area.
- **Google Earth Pro:**
  Used to visually map the approximate structure and location of each classroom. Coordinates were recorded for all corner angles as well as the central point of each room. This ensured that a complete boundary could be delineated for each geofencing zone.

3.3 Collected Geofencing Data



FACULTY OF ENGINEERING AND TECHNOLOGY			
Classrooms	**Class angles**	**Lattitude**	**Longitude**
FET-BGFL	**Top Right**	4° 8'35.21"N	9°17'18.28"E
	Top Left	4° 8'34.99"N	9°17'18.45"E
	Center	4° 8'35.25"N	9°17'18.58"E
	Bottom right	4° 8'35.53"N	9°17'18.70"E
	Bottom Lefft	4° 8'35.32"N	9°17'18.88"E
FET-BFF-HALL 1	**Top Right**	4° 8'35.16"N	9°17'18.68"E
	Top Left	4° 8'35.38"N	9°17'18.51"E
	Center	4° 8'35.19"N	9°17'18.48"E
	Bottom right	4° 8'34.98"N	9°17'18.46"E
	Bottom Lefft	4° 8'35.21"N	9°17'18.28"E
FET-BFF-HALL 2	**Top Right**	4° 8'35.40"N	9°17'18.55"E
	Top Left	4° 8'35.20"N	9°17'18.71"E
	Center	4° 8'35.37"N	9°17'18.75"E
	Bottom right	4° 8'35.56"N	9°17'18.77"E
	Bottom Lefft	4° 8'35.36"N	9°17'18.93"E

FET TECHNOLOGY BLOCK			
Classrooms	**Angles**	**Lattitude**	**Longitude**
TECH 1/TECH 4	**Top Right**	4° 8'34.54"N	9°17'24.96"E
	Top Left	4° 8'34.76"N	9°17'25.18"E
	Center	4° 8'34.84"N	9°17'24.89"E
	Bottom right	4° 8'34.95"N	9°17'24.59"E
	Bottom Lefft	4° 8'35.16"N	9°17'24.81"E
TECH 2/TECH 3	**Top Right**	4° 8'34.67"N	9°17'24.73"E

	Top Left	4° 8'34.90"N	9°17'24.52"E
	Center	4° 8'34.62"N	9°17'24.46"E
	Bottom right	4° 8'34.37"N	9°17'24.37"E
	Bottom Lefft	4° 8'34.59"N	9°17'24.19"E

3.4 Shortcomings

Despite careful planning and execution, several shortcomings and potential challenges have been identified in the current geofencing data collection process:

- **Measurement Accuracy:**

The use of a measuring tape for classroom dimensions may introduce human error, especially in irregularly shaped rooms, which can affect the accuracy of the defined geofence boundaries.

- **Dependence on Google Earth Pro:**

The coordinates obtained from Google Earth Pro are approximate. They may be subject to small discrepancies due to variations in the mapping data and resolution limits.

- **Limited Spatial Resolution:**

Data captured from the corners and center of classrooms provides a basic outline of the geofence; however, potential obstacles (walls, partitions) within the room may not be fully accounted for.

- **Environmental Variability:**

Changes in building structures over time (e.g., renovations) may render the collected coordinates outdated. Periodic re-verification of the geofencing data is necessary.

- **GPS Signal Interference:**

In indoor settings, GPS signals can be weak or prone to multipath errors. This can influence the reliability of real-time geofencing verifications, particularly in densely built environments.

- **Uniformity in Data Collection:**

The manual process could result in inconsistency if the same methods and tools are not used uniformly across all classrooms. Standard operating procedures should be established to mitigate this risk.

Hence, geofencing data gathered serves as the foundational reference for validating classroom locations within our attendance management system. While the current dataset is comprehensive, addressing the aforementioned shortcomings will be crucial for ensuring the long-term reliability and accuracy of the geofencing component. Future work should include periodic data validation and potentially the integration of more advanced geolocation technologies to overcome GPS limitations in indoor environments.

IV Data Cleaning

Data cleaning is a critical step in ensuring the reliability and accuracy of the data collected through surveys, interviews, focus groups, and other techniques. It involves identifying and correcting issues such as missing values, inconsistent formats, incorrect entries, and unstructured responses. For our **smart attendance management system project**, effective data cleaning was essential to derive meaningful insights, build reliable dashboards, and feed the cleaned dataset into downstream processes like reporting and machine learning models.

1. Overview of the Data Cleaning Process

After collecting responses from stakeholders (students, instructors, and administrators), we observed that the raw data included several inconsistencies such as:

- Inconsistent naming of institutions and cities.
- Typos in free-form text fields.
- Varying capitalization (e.g., “UB” vs “ub”).
- Blank or malformed timestamps.
- Responses in non-English languages.
- Mixed formats for rating scales.

To resolve these issues and prepare the data for analysis, we followed a structured cleaning workflow.

2. Methods and Processes Used

Method	Description
Trimming and Formatting	Removed extra spaces, corrected typos, and standardized capitalization.
Deduplication	Detected and removed repeated entries based on identical timestamps and answers.
Standardization	Unified terms for institutions, roles, devices (e.g., "android" → "Android").
Translation & Harmonization	Translated French responses into English and standardized vocabulary.
Data Type Conversion	Transformed categorical ratings to numerical format; parsed dates as datetime.
Handling Missing Values	Filled blanks with defaults (e.g., "Unknown") or dropped incomplete rows.
Encoding Fixes	Ensured UTF-8 encoding to prevent character corruption.
Normalization	Re-aligned nested location or institution data into consistent fields.

3. Before vs After Cleaning: Sample Transformations

Field	Before Cleaning	After Cleaning
location institution	"UB, buea and u belong to the university of buea Cameroon"	"University of Buea"
device type	"ios " (extra space), "android"	"iOS", "Android"
anticipated challenges	"low batterie to make a check in"	"Low battery issue when checking in"
biometric concerns	"tout le monde n'est pas constamment connecté..."	"Some students may not be connected..."

Field	Before Cleaning	After Cleaning
role	" administrator(lecturer , educator)"	"Educator"
satisfaction rating	"Very satisfied", "4.0", "Neutral"	"5", "4", "3"
response timestamp	Missing or blank	"2025-01-01 00:00:00" (defaulted)

4. Why Data Cleaning Was Necessary

A. Improve Accuracy of Analysis

- Ensures that entries referring to the same institution (e.g., "UB", "university of buea") are grouped together.
- Avoids duplicated or misclassified feedback.

B. Enhance Understanding of User Experience

- Makes qualitative feedback more readable and interpretable.
- Helps identify common patterns, such as power/battery issues or connectivity challenges.

C. Ensure Consistency Across Reports

- Clean data allows easy aggregation by categories such as location or role.
- Enables consistent analysis and reporting across survey batches.

D. Prepare for Automation & Machine Learning

- Structured and normalized data can be fed into visualization dashboards, statistical models, and automated systems (e.g., predictive attendance anomaly detection).

5. Tools and Technologies Used

Tool / Platform	Use in Cleaning Process
Microsoft Excel	Initial exploration, filtering blanks, removing duplicates in small datasets.
Python (Pandas)	Comprehensive data wrangling: missing value handling, normalization, encoding correction, transformations.
OpenRefine	Clustered messy institution names; harmonized free-text entries using reconciliation features.
R (tidyverse)	Used selectively for reshaping and string manipulation.
SQL	Filtering data subsets, performing joins, and checking constraints in structured datasets.

6. Python Code Sample for Survey Data Cleaning

Below is a Python script we used for cleaning raw TSV data collected via Google Forms and email surveys:

```python
import pandas as pd
import numpy as np
from datetime import datetime
import re
# Load TSV data
df = pd.read_csv("raw_survey_data.tsv", delimiter="\t", encoding="utf-8")

# Trim whitespaces
df = df.applymap(lambda x: x.strip() if isinstance(x, str) else x)
# Normalize roles
df["role"] = df["role"].str.lower().replace({
    "student": "Student",
    " administrator(lecturer , educator)": "Educator",
    "educator": "Educator"
})
# Standardize device types
df["device_type"] = df["device_type"].str.strip().str.capitalize()

# Fill missing timestamps
df["timestamp"] = pd.to_datetime(df["timestamp"], errors='coerce')
df["timestamp"].fillna(pd.Timestamp("2025-01-01"), inplace=True)
# Translate French responses (simplified example)
df["biometric_concerns"] = df["biometric_concerns"].replace({
    "tout le monde n'est pas constamment connecté...": "Some students may not be connected during attendance"
})
# Handle satisfaction ratings
rating_map = {
    "Very satisfied": 5,
    "Satisfied": 4,
    "Neutral": 3,
    "Dissatisfied": 2,
    "Very dissatisfied": 1
}
df["satisfaction_rating"] = df["satisfaction_rating"].replace(rating_map)
# Save cleaned data
df.to_csv("cleaned_survey_data.csv", index=False)
print("Data cleaning complete. Cleaned file saved.")
```

Fig 6: screenshoot of python code used for cleaning

7. Summary and Impact of Data Cleaning

Data cleaning transformed a messy, inconsistent set of survey and feedback responses into a structured, reliable dataset. This process was essential for:

- Accurately identifying trends in user behavior and concerns.

- Feeding high-quality data into our visual dashboards and analysis tools.
- Ensuring the integrity and reliability of conclusions drawn from user research.
- Establishing a foundation for automation and integration in the final attendance system.

V. User Reluctance Assessment

1. Introduction

The adoption of any new technological system is often accompanied by varying degrees of user reluctance, especially when it involves sensitive data such as biometric information and location tracking. As part of the requirement gathering process for the Mobile-Based Attendance Management System, we conducted a comprehensive assessment to identify the factors that may cause hesitation among users and proposed strategies to mitigate these concerns.

This section provides an in-depth analysis of anticipated user reluctance, the reasons behind it, and planned measures to enhance user acceptance and system usability.

2. Objectives of the Assessment

- Identify the root causes of user reluctance across different stakeholder groups.
- Evaluate the severity and potential impact of these concerns on system adoption.
- Develop strategies to address concerns through design, communication, and training.

3. Methodology

To assess user reluctance, the following methods were employed:

- **Targeted Surveys**: Distributed to students and instructors to gather opinions on biometric data usage and location tracking.
- **One-on-One Interviews**: Conducted with selected participants to explore concerns in more depth.
- **Focus Group Discussions**: Facilitated dialogue sessions with small groups to understand the emotional and practical barriers to adoption.
- **Literature Review**: Referenced privacy and usability studies related to facial recognition and GPS-based systems.

4. Key Areas of User Reluctance Identified

Concern Area	Stakeholders Affected	Description
Privacy of Biometric Data	Students, Instructors	Concerns over storage, misuse, or leakage of facial recognition data.
Location Tracking Fears	Students	Discomfort with continuous location access, fear of being monitored outside academic purposes.
Trust in Facial Recognition Accuracy	Students, Instructors	Worries about misidentification, particularly among people with similar facial features or under poor lighting.
Increased Battery and Data Usage	Students	Hesitation to use an app that may drain mobile resources due to GPS and camera use.
Technical Complexity	Non-tech-savvy Instructors	Resistance to using a mobile app due to unfamiliarity with smartphones or reluctance to adopt new tools.
Resistance to Change	All stakeholders	General hesitancy to transition from familiar manual or semi-digital systems.

---

5. Detailed Analysis

5.1 Privacy Concerns (Biometrics and GPS)

Issue: Many students expressed hesitation in sharing their facial images, fearing potential misuse or lack of control over their personal data.
Feedback Sample:

- "What if someone hacks into the database and gets our faces?"
- "I don't like the idea of my face being stored somewhere forever."
  Implications: High reluctance could result in non-compliance or requests for opt-out alternatives.

5.2 Location Tracking Apprehension

Issue: The use of GPS sparked discomfort about being tracked, especially outside class hours.
Feedback Sample:

- "I don't want the app to track me after class."

- ❖ "Can it access my location 24/7 or just during attendance?" **Implications:** Without clear limitations, this could reduce app installs or cause legal complications.

5.3 Facial Recognition Accuracy Doubts

Issue: Participants doubted the ability of the system to correctly identify them every time, especially under real-world conditions like bad lighting, accessories, or facial changes. **Feedback Sample:**

- ❖ "What if I change my hairstyle or wear a mask?"
- ❖ "How accurate is this really? Can it confuse me with someone else?" **Implications:** Low trust in the technology could reduce user satisfaction or increase false negatives.

5.4 Device Resource Concerns

Issue: Students were concerned about the app draining battery or consuming too much mobile data due to GPS and camera activity.
Feedback Sample:

- ❖ "Will this app eat up my battery with constant GPS tracking?"
- ❖ "Do I need to have data on all the time?"
  Implications: Negative impact on app ratings and reduced daily usage.

5.5 Resistance from Non-Tech Users

- ❖ **Issue:** Instructors who are less comfortable with mobile technology may resist switching from manual methods.

  Feedback Sample:

- ❖ "I prefer paper—it's straightforward."
- ❖ "Too many apps already. One more will confuse things."
  Implications: Could result in inconsistent usage across departments.

6. Mitigation Strategies

Concern	**Mitigation Strategy**
Privacy (Biometric & GPS)	- Store all facial data in encrypted format using AES-256. - Include transparent consent forms during registration. - Clearly explain when and how location data is used (e.g., only during class check-in).
Accuracy of Face Recognition	- Use robust deep learning models (e.g., using TensorFlow + OpenCV). - Include fallback mechanisms (e.g., manual verification or secondary checks).
Battery/Data Usage	- Use Google's **Fused Location API** for efficient location detection. - Allow offline check-in with delayed syncing when online.
Technical Complexity	- Provide a guided tutorial within the app. - Offer training workshops or support materials for instructors.
General Resistance to Change	- Conduct demo sessions before rollout. - Highlight benefits (speed, security, transparency). - Gather early adopters to act as advocates.

The user reluctance assessment revealed valid concerns, particularly around **privacy, usability, and accuracy**. However, these challenges can be effectively managed through thoughtful design choices, clear communication, and user empowerment features. Addressing these concerns early in the development process will significantly enhance user acceptance, trust, and long-term success of the system.

Conclusion

The success of any software system hinges on how well it aligns with the needs of its intended users. Through our comprehensive requirement gathering process, we were able to deeply understand the expectations, behaviors, and pain points of stakeholders in academic institutions. The insights gathered have not only guided the functional and non-functional design of our smart attendance system but also revealed critical usability and adoption factors such as connectivity issues, device limitations, and data privacy concerns.

By investing in structured data collection and meticulous data cleaning, we ensured that our decision-making was based on reliable, high-quality information. The result is a user-driven foundation for a system that promises to deliver secure, automated, and transparent attendance tracking through geofencing and facial recognition.

This report serves as the cornerstone for our design and implementation phases, paving the way toward a smart, scalable, and context-aware attendance solution that meets the evolving needs of modern academic environments.

References

1. TechTarget. (n.d.). *What is Stakeholder?*. Retrieved from https://www.techtarget.com/searchcio/definition/stakeholder

2. Smartsheet. (n.d.). *The Ultimate Guide to Stakeholder Management*. Retrieved from https://www.smartsheet.com/stakeholder-management-guide

3. Lucidchart. (n.d.). *How to Create a Stakeholder Map*. Retrieved from https://www.lucidchart.com/blog/how-to-create-a-stakeholder-map

4. Factors of users reluctances: https://www.linkedin.com/advice/1/what-causes-user-frustration-software-design-skills-software-design-6lhqc, https://chatgpt.com/c/67f8d162-e794-8005-84d5-f4b5921aa62e

5. Requirement gathering techniques(April 2025): https://www.tutorialspoint.com/business_analysis/business_analysis_requirement_gathering_techniques.htm

6. 5 best requirement gathering techniques(April 2025): https://agilemania.com/tutorial/requirements-gathering-technique

7. Top 12 Requirement Gathering Techniques in Agile Product Management GeeksforGeeks (April 2025)

8. Requirements Gathering - Introduction, Processes, Benefits and Tools | GeeksforGeeks (April 2025)