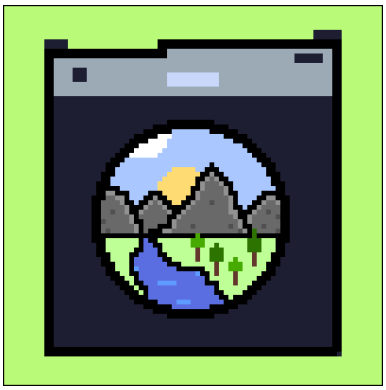
Requirements Specification

12/6/2024

Forest Frames



Sponsored by Dr. Camille Gaillard, Dr. Chris Doughty, Dr. Duan Biggs, and Dr. Jenna Keany

Mentor: Scott Larocca

Members: Dalton Tippings, Daniel Austin, and Nicholas Greco

Version 1.1

Client Signature:

Team Signature:

Table of Contents

1. Introduction

In recent years, citizen science has emerged as a powerful tool for environmental conservation, leveraging the power of crowdsourced data to support wildlife monitoring, biodiversity studies, and ecosystem health assessments. This approach has been particularly impactful in countries with vast, hard-to-monitor wilderness areas. However, in many rural regions—such as those found in Malaysia or Kenya—time and financial incentives are often limited, leaving significant gaps in local ecological data collection. This capstone project addresses these gaps by developing a citizen science app that incentivizes residents in these regions to participate actively in gathering and submitting environmental data.

The primary objective of this application is to make it simple and rewarding for local residents to submit verified data on flora and fauna in their surrounding wilderness areas. Using a combination of machine learning and data integration with established wildlife databases, the app verifies user submission for accuracy. Verified data not only contributes to conservation efforts but also allows for the user to receive benefits, which could range from monetary compensation to other locally relevant rewards. This approach encourages consistent engagement, offering an accessible way for locals to play a crucial role in monitoring and preserving their natural environments.

By aligning with conservation goals, the app has the potential to contribute valuable insights to scientists, environmental organizations, and government agencies that rely on precise, localized data to make informed decisions. Through this platform, the project aims to foster an interconnected community of local citizens and scientists, creating a sustainable, data-driven approach to environmental conservation in underrepresented areas.

2. Problem Statement

Citizen science is a term that refers to initiatives pushing for the general public to participate in scientific research. Citizen science particularly tends to lend itself to people cooperating with scientists on gathering important information and data on the natural world. This allows for the general public to contribute to science easily and allows for scientists to have a significant amount of useful information that they can pull from. Such information gathering can come in a variety of forms, be it non-scientists going to an area that an organization wants to

research and gathering data to send over. It could also be that people live somewhere that scientists can't easily travel to, in order to perform field research. These initiatives can also come with benefits for those who participate in them, be it financial compensation or another form of reward.

However, for many citizen science initiatives, getting information on more less-documented regions of the world can be a significant challenge due to the general public in those areas having limited technological means to contribute. An area with very little way for scientists to consistently communicate can present a challenge for expanding citizen science and its incentives to those regions. Along with that, getting reliable data can be difficult for these regions as data verification can be very limited. The problems that stand out due to this are:

1. **Allow as many people as possible to participate in citizen science** - This problem primarily arises from the fact that in these regions, many of the technologies that people have are either old or outdated. So, when developing a technology for citizen science, we cannot rely on our target market to have the latest technology. A solution has to be to make something that is compatible with old devices along with being able to run efficiently on most devices.
2. **How to ensure we don't need to be in constant contact with participants** - Many areas lack a constant or stable connection, or it would not be feasible to maintain a constant connection. While this is less of a problem in more populated parts of the world, others face this in a large majority of areas.
3. **Let users gather useful data** - All citizen science requires that those who are participating be able to contribute to the wider scientific community, and this primarily consists of gathering data, such as images, video, etc. Without this there is little that a person can do to contribute to any citizen science initiatives.
4. **Verify data is trustworthy and reliable** - Citizen science is useless if any data gained cannot be deemed as a trusted source. If there are incentives to participate, such as financial compensation, there will be participants who don't participate in good faith and seek to abuse this in order to receive more rewards without contributing.

3. Solution Vision

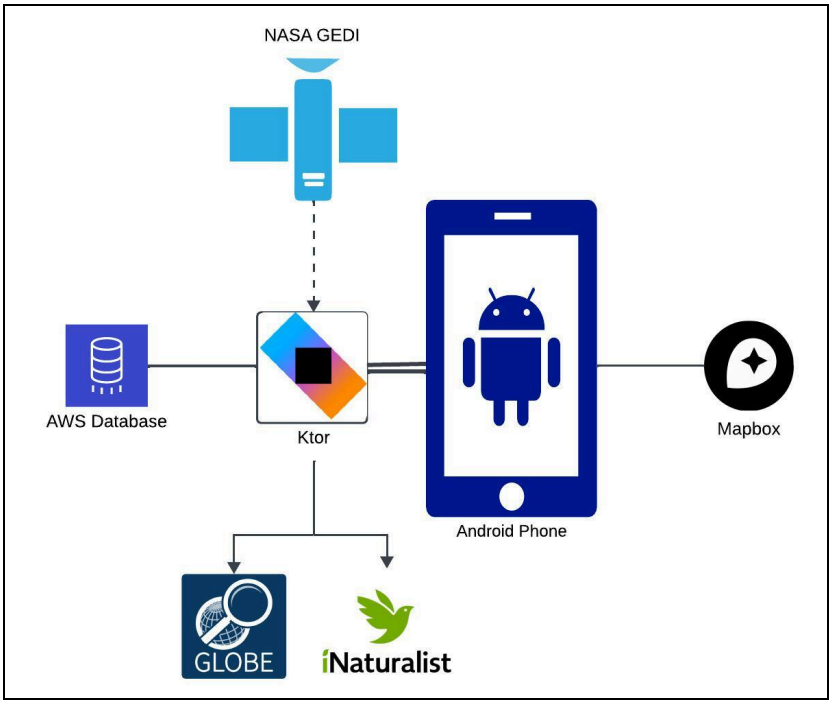
In order to rectify these problems we will be developing an efficient and lightweight mobile application that can run on old devices. This will solve the problems by providing a way for people to participate in citizen science and for us to take said data and verify its validity. This application will be able to connect to a remote server which will handle data processing that would otherwise be too computationally expensive on a phone. This application will seek to facilitate the participation of people in citizen science who would normally be unable to participate in it. We will do this by providing a solution to the problems mentioned above that focus on data collection.

Our backend will consist of both a server and a database. The server will be used to verify the user-collected data using public datasets about the area that the user collected their data in. For multimedia data types, we will use multiple different open-source recognition technologies to extract data that will be comparable with those public datasets. The database will be storing all user-collected data sent to the server as a backlog for future reference to data or retroactive correction. This will help provide a solution to our final fourth concern.

3.1 Key Features

1. To ensure the maximum number of people can participate in citizen science through our application, we are going to focus on creating an **Android application compatible with Android version 12 and up**. This will allow a wide range of devices to use our application without having to worry about the version. Along with this, an important aspect to ensure this works across many devices is to make it as lightweight and efficient as possible. So this includes using well-performing technologies and optimizing our app.
2. While some connection will be required to download and upload data for our application, it will be able to **function offline for large periods of time**. While a user is out gathering data, they don't have to connect to any online services, and will only require a connection when specifically downloading data when needed and uploading gathered data to our servers.
3. Our app will use the Android system services to allow the user to **use their camera and microphone to gather data** that we have requested and store it locally on their phone until they are ready to upload it for verification.
4. **Verification will be conducted** by using a variety of outside APIs and datasets to cross-reference the data and get a "verification score" which we can then use to determine if the data is trustworthy or not.

Figure 1: Project's Architecture



4. Project Requirements

This section describes what particular aspects our project will implement to meet our solution vision. These requirements outline what a particular requirement is, and detailed information on the content of each. There are three requirement sections outlined below. These sections are functional requirements, nonfunctional requirements, and environmental requirements.

- Functional requirements are the technical features that our app will have.
- Nonfunctional requirements are things our system will do using technical features.
- Environmental requirements are the constraints development of our app will have, things that have to be met so that the system can be used by our users.

4.1 Functional Requirements

4.1.1 Multiple Different Collected Data Types

Our system will be able to record and store a variety of different data types.

a. Our system will be able to record photos taken by the user, which will be native to our app.
b. Our system will be able to record and store the data. The recorded data via the app will be stored on the user's device and can be removed once uploaded to our database.

4.1.2 App To Server Data Transfer

User-collected data is able to be transferred from the user's device to our server.

a. The cumulative set of data from the user's device on a specific footprint will be selected and grouped together, which will then make the data ready to be sent to the server.
b. Once the data transfer is initiated and the process finishes, both the user, via the app, and the server are notified that the data was a success or a failure.

4.1.3 Verification Of User Data

User-collected data is verified on our server through one or multiple verification processes.

a. Our server will use the metadata, and other provided supporting information to the given data, to determine which verification process to run.
b. A verification process involves first possibly transforming the given data into data that can be compared to public datasets/databases, and then performing that comparison test to produce a "rating"
c. After all possible verification processes have been run, the overall validity of the data can then be determined.

4.1.4 Store User-Collected Data

User-collected data will be stored on a separate database for later retrieval.

a. Once data is received on the server side, automatically store data and its attributes: user identification, data type, etc.. This data being stored will be the first action when the data arrives in case the user does remove this data from their local storage, as permitted by individual prompts after upload or based on their predefined settings.
b. Be able to modify database entries to update attributes: verification result, organizations data has been sent to, user incentivized for data, etc..
c. As a result, multiple calls to add, modify, and retrieve from the database will be performed as data is sent to the user, verified, and sent back out.

4.1.5 Providing Data To Other Organizations

Package collected data and send them to other organizations' databases

a. Data being sent is retrieved from our database and is formatted in a way so it is ready to be transferred.
b. Data can be transferred from our server to organizations' databases (GLOBE mainly) and the server is notified of that result.

4.2 Non-Functional Requirements

4.2.1 Data Collection While Offline

The user will be able to collect data while they do not have access to the internet or other forms of online communication.

a. The user will be able to open and use the application without being connected to the internet. This includes using previously downloaded data, using the application to gather data, and view data gathered.
b. The application will be able to store any recorded data locally so that the user can later upload it when they are able.

4.2.2 Accuracy of Provided Data

Based on a set threshold, data sent to an organization will be accurate and reliable.

a. Our system will create an accuracy rating for all data uploaded to rate how trustworthy it is. This rating will then be used to determine if a given set of data is reliable.
b. Our system will be able to filter data by how reliable it is and determine if it is eligible to send to outside organizations.

4.3 Environmental Requirements

4.3.1 Android Version 12+ Compatibility

The application will be able to run on Android devices capable of running on Android version 12. This will include ensuring a minimum SDK level of 31.

a. Development for this application will focus on the Android operating system, specifically for their mobile phones. All Android phones that can run version 12 and up will be able to use our application.
b. The application will have a target SDK Level 34, for Android Version 14, which is required for new Android apps.
c. The application will have a minimum SDK Level 31 which corresponds to Android Version 12, to implement compatibility with older devices.

5. Project Risks

All software projects come with various risks when developing a finished product. Many such risks can be avoided by ensuring that all implementation is made correctly, but some risks are unavoidable due to what the system has to do. In particular, there are two primary risks that are unavoidable, one is the verification of data, failure to verify uploaded data and getting the wrong result. The second risk is authenticating users correctly to ensure security.

5.1 Type I & Type II Verification Errors

Since verification of the user-collected data does not result in a clear “yes or no” for whether the data is accurate, and instead gives a percentage for how accurate it is, a threshold must be set to decide whether the data is accurate enough. Ideally, we would want a 100% threshold; however, this is not realistic considering random chance could change that outcome. So, a risk we must assess is deciding on a threshold that doesn’t have too many Type I (false positives) or Type II (false negatives). If a large amount of invalid data is deemed valid, then scientists would be basing what would be valid research on an invalid dataset. On the other hand, if large amounts of valid data is being deemed invalid, then honest users would not be rewarded for their work. Multiple tests, with large datasets of both valid and invalid data, would need to be run to determine an optimal threshold to set, which would minimize the amount of those aforementioned cases.

5.2 Authenticity of App User

Taking into consideration that this project will eventually provide incentives to people for their collected data, not only is verification of their data important, but whether the user using the app is who they say they are. This could happen if that person’s device is stolen and they are still logged into the application, or just someone using that person’s login info. In order to combat stolen devices being used, we could provide users an option to set a PIN, so in order to access and update their personal local data, they would need that PIN number, which acts as a second layer of security. We could also implement a timer on a user’s device that will automatically logout the user after a certain period of inactivity or a set amount of time. As far as actual login security, two-factor authentication could be an option users can toggle. Otherwise, we would suggest that the users themselves create secure passwords. Our app’s local storage data encryption will also help prevent data tampering which could allow someone to retrieve sensitive information if they have gained access to a person’s device.

6. Project Plan

6.1 Get the server up and running

This step focuses on getting the server that will handle the processing and verification of data up and running. The server is meant to handle all the verification of data as well as the transfer of data to and from the user and outside organizations. We will start out by ensuring that the server is able to do basic functions such as get data from an outside source and send data to an outside source.

6.2 Allow app to get/send data to server

This step will focus on the creation of our app and allowing it to send data to our server. The app will be able to send basic data (text, image, etc.) to our server and receive some data

back. Over the course of development more varied and advanced versions of data will be able to be sent, along with the app being able to receive all needed data.

6.3 Create app UI

This will be a continuous process over development to create and improve the apps UI. The goal is to make it as simple and user friendly as possible, as many of our users will be those that have little to no experience with mobile applications. This will involve creating a simple UI whose navigation is both simple and intuitive for those who have not used mobile applications before. This will involve making navigation buttons explicit and assuming that the user is not familiar with any standard symbols used in many modern applications. This UI will constantly be updated as the project development progresses, showing it to potential users and refining the experience to make it more user-friendly. The end goal is to make an extremely user-friendly experience while having all the functionality previously described in the requirements section.

6.4 Connect server to database to store information

Once the server is running, we will need to make a database that is able to store our users' data and temporarily hold uploaded information. This will involve making a small database that is able to hold some test data to start off with but will expand upon release. Our server will be able to access the database and store/retrieve data as needed.

6.5 Server verifies data

This will be the largest part of the system we are developing. When a user uploads data that needs to be verified, the server will take a piece of the data and find out how it should be analyzed. We plan on using various APIs and outside databases to identify and match data to verify its authenticity. When verifying a piece of data, we will assign it a verification "score" and see if it is reliable enough to then use. This score will be determined over the course of the development process.

6.6 Server sends data to outside sources

Once a piece of data has been verified to a satisfactory level, we will set up a method to send that data to outside organizations, the two decided ones being GLOBE and iNaturalist, that can make use of said data. Our system will handle all of the transfers so that the user does not have to deal with any additional work when their data is being uploaded.

6.7 Field tests

As we develop our functionality, we will conduct field tests for our functionality. While testing will be done throughout the development process, it will mainly be concentrated near the end of development once most of the functionality is implemented. This will involve testing verification effectiveness, user friendliness of the app UI, and processing speeds. Once the field tests show positive results the application will be ready to release.

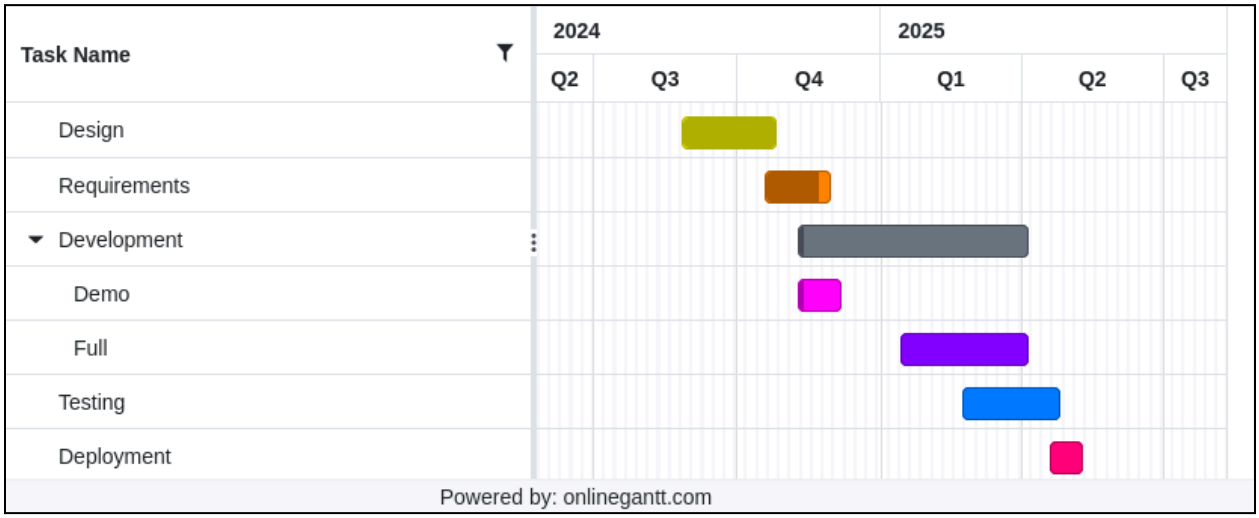
6.8 Optimizations

This will be a constant process to optimize the application to fulfill the requirements for performance. While this will be conducted throughout the development process, it will be a focus once all the functionality is finished alongside the field tests. During the testing phase, we will identify areas that could be improved upon and focus on improving the speed and performance of the application until it reaches an optimal level.

6.a Schedule Overview

The schedule attached below (Fig. 2) displays our overall project's schedule. We have currently completed our design phase and almost completed our requirements phase. We will be moving to developing our demo which will cover sections 6.1-6.2 of our project plan. Developing our full project will involve initial development, with implementing various tests alongside as development continues and becomes more advanced. After full system testing concludes, once development also subsides, we will begin deployment of our app, server, and database.

Figure 2: Overall Schedule



7. Conclusion

In conclusion, our capstone project addresses a critical need in the field of environmental conservation by creating a citizen science app tailored for rural regions. The project shows the importance of local engagement in wildlife data collection, particularly in areas where ecological data is sparse, and locals often lack a reason to contribute. By creating a platform that verifies submissions against established wildlife databases and provides rewards, we are taking a big step toward making environmental monitoring both accessible and sustainable.

Our proposed solution not only fills a data gap but also creates a sense of shared responsibility among users, offering them an opportunity to contribute to conservation efforts while receiving tangible results. This document has laid out the requirements and objectives necessary to make this project, establishing a baseline for our development, and creating the tools and strategies we will use to meet our requirements.

Through this approach, we foresee creating a user-friendly, impactful app that generates valuable data for conservation. By allowing for sustained engagement and verified data contributions, our project is poised to make a long-term impact on environmental data collection in regions that need it most.