Requirements

November 15th, 2024

Data Integrity and Abuse Prevention for Environmental Preservation

Team Members: Aidan Trujillo, Fady Zaky and Kyle Bambling

Mentor: Scott Larocca

Sponsor: Dr. Camille Gaillard, Dr. Duan Biggs, Dr. Jenna Keany, and Dr. Chris Doughty

Client: ___________________________

Team Lead: ________________________

Table of Contents

1. Introduction

In 2022, 180 of the world's governments came together to decide that they will, by 2030, conserve 30% of the earth's land and sea. This came in an effort to promote biodiversity across the globe. Since this is such a large portion of the world, there are concerns with equity for some communities that live in areas where conservation is taking place, especially with endangered species. The people that live here can find themselves being displaced from the land that they may have lived on for their entire lives. As important as biodiversity and conservation is for the world, the consequences are life changing for some. Combining conservation efforts with benefits to local communities around the world remains a challenge.

Our clients, Dr. Camille Gaillard, Dr. Jenna Keany, Dr. Duan Biggs, and Dr. Christopher Doughty, have noticed the challenges that these communities face. In an effort to give back, that is where our project comes in. Citizen science is a concept where everyday people can contribute to scientific research and data collection. Our clients want to create an application that is going to allow anyone from rural communities to contribute to citizen science so long as they have a mobile phone.

For our application, people in these communities will be able to contribute to validating data from the NASA GEDI satellite. The GEDI satellite collects data around the globe and is able to determine ecosystem biomass and carbon content. This data needs to be verified and that is where citizen scientists can come in. By using our application, users will be able to upload photos of the biodiversity in their area so we know what is actually there along with collecting measurements to assess the accuracy of GEDI satellite data. The end goal is being able to monetize this for our users so that their contributions made to citizen science are rewarded. In turn this should be able to funnel needed funds into rural communities affected by this while also motivating more affected communities to contribute.

In this document, we will define the minimum requirements that the application will fulfill. Meeting these requirements will determine the criteria for a successful application. To start, we will describe the solution vision so that there is a clear outline for what requirements are needed. Next, the functional, performance, and environmental requirements of the project will be laid out. After that, we will identify potential risks that come with the project and the plan for implementation.

This application is in conjunction with another group named Forest Frames. This group will be handling many other requirements that we have not mentioned in regards to security and in regards to the backend of our application. It will be one app in the end promoting the same cause for the same clients.

2. Problem Statement

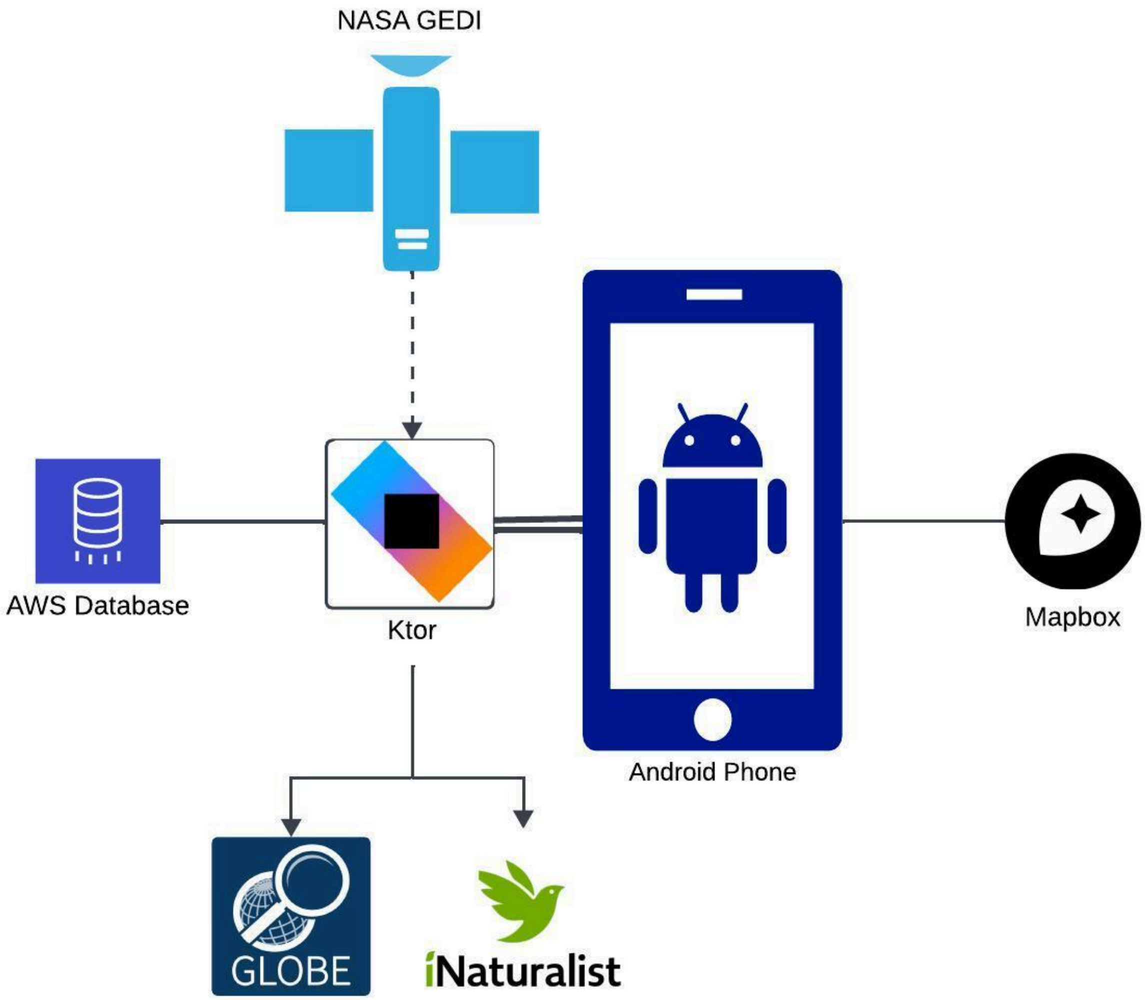
A big part of conservation and biodiversity includes tracking what plant and animal life is actually in the areas being conserved. Our clients have gone to the areas where conservation is the most important and visited the communities most affected. They have contributed to existing community science applications such as Globe (which contributes directly to GEDI data), eBird, and iNaturalist. However, individual researchers are unable to record biodiversity at large scales and across the remote and biodiverse regions proposed to be conserved under the 30 x 30 plan. Being able to rely on GEDI data solely is also not enough since it is still unknown what exists in these areas. That is important for both conservation as well as being able to better understand patterns from the GEDI data. Our clients have been working on the verification of GEDI data as well as working on various other conservation efforts within remote communities. Thus, they are able to understand the needs of these communities as well as meeting them where they are at as far as technology available and current conservation efforts.

Biodiversity and conservation are clearly huge goals for our team as they help to contribute to the promise made by the world government leaders. Creating a biodiverse world takes a community effort to conserve the environment in which it exists. Unfortunately this may include displacing some entire communities from their land. This is a complete disruption of their way of life. Another way these communities are affected is by the existing wildlife in the area. For example in Jenna Keany's studies in Africa, she noticed that elephants may completely destroy entire crop fields just as they move around. Clearly, this is catastrophic for not only the farmers but also the community that they support. Farmers not only lose out on their livelihood, but the community loses out on much of the food that it relies on. This example shows the need for biodiversity conservation in marginalized communities.

Ultimately what our clients want is to give communities most affected by conservation a chance to both contribute and benefit from it. The idea would be to monetize their efforts by paying members that contribute to community science. By doing this our clients are able to funnel funds into communities that are struggling. Ideally, this application would work without any tampering of data, but when money is involved that is not always the case. It is very easy for people to take a picture of something and upload data from another location, a picture they have already sent, or provide the wrong data entirely. This becomes a problem when users are being paid for data that is not only unhelpful, but harmful to the efforts of conservation by providing incorrect data.

The members of this community are going to be mostly working offline and they are going to not always know where the GEDI data is that they need to be at. Additionally, they will likely be using much older versions of Android than current applications support. As a team, this is where we step in to assist in overcoming these problems and verifying where a user is, that they have tools to get there, and that their current technologies will be enough. Through research of different technologies we have come up with what we believe to be a viable solution that encourages community science efforts with integrity.

3. Solution Vision



We are proposing an Android application that has the ability to extract NASA GEDI satellite coordinates and display them on a map for the user to see and navigate to. In order to do this, we would have to extract the GEDI data from NASA's database. This can be done using the RESTful API integrate into Android's Kotlin called **Ktor**.

Our system would then process that data and display them onto the interactive map provided by the **Mapbox** API. Using the interactive map, users would be able to see their location as well as locations of nearby NASA GEDI footprints.

If the user decides to go to one of the displayed coordinates, they would unlock the ability to record data, whether images or audio recordings, from the coordinates area. Our system would then verify that the user is within the coordinates before allowing them to upload the data to our **AWS** database.

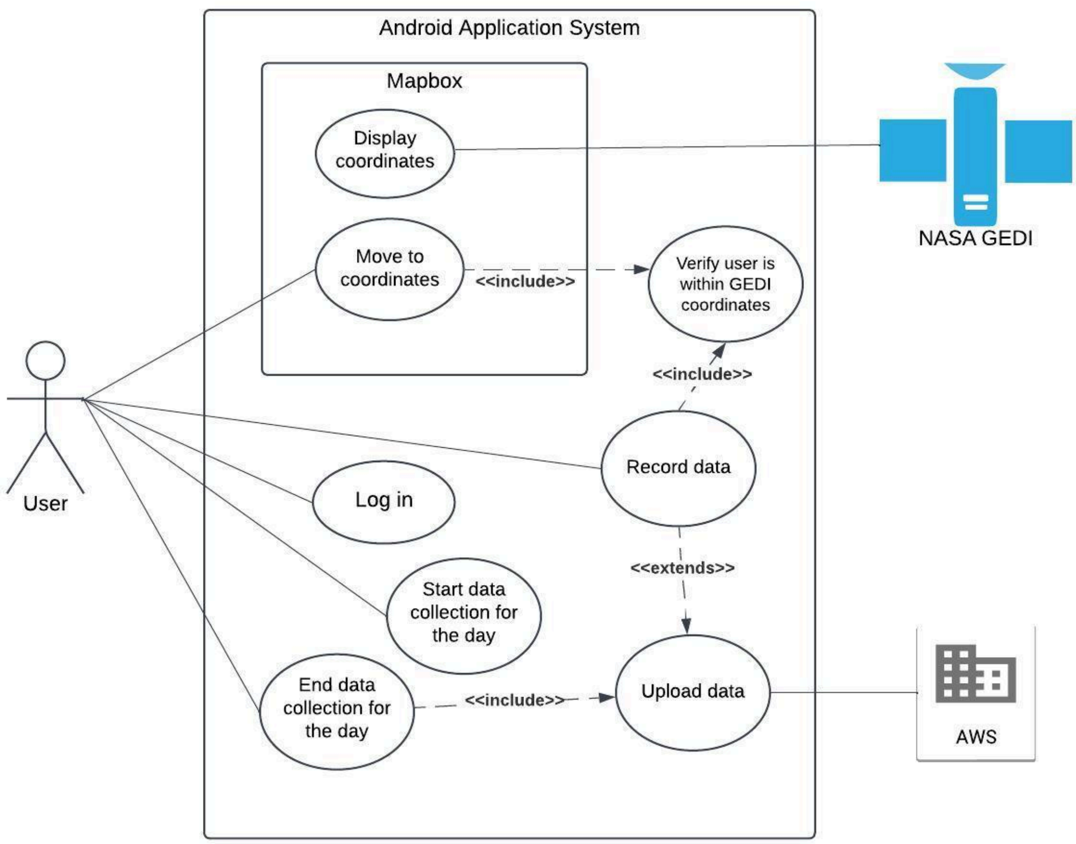
The collected data would then be sent out to *NASA Globe* and/or *iNaturalist* and contribute to their data collection and research.

Our solution will ensure that all the data gathered will help with biodiversity conservation around marginalized areas, benefitting both the wildlife animals as well as the communities that live in said areas.

4. Project Requirements

4.1 Functional Requirements

The following functional requirements will ensure the development of a solution that caters exactly to the requirements of our clients. Each section details the specific implementation of features within the technologies used in our project, ensuring that all objectives are fully addressed. Each is important to promoting an offline application since when they are all provided in the end, a user should be able to collect data after they get a map. As said earlier, these are the requirements for this group's implementation of the project. There will be more requirements in the Forest Frames group that provide a complete functionality when combined with our requirements.



4.1.1 The system allows the user to login with their credentials

It is necessary that a user possesses an account to be able to login keeping track of who should be given funds on successful data verification as well as who may be uploading consistently false data. Password protections are important for managing a users funds and insuring no impersonation of other users.

1. The system asks the user to log into their account
2. If the user does not have an account, the system would ask the user to make a new account
3. The system asks the user if they wish to start collecting data for the day
4. If the user starts the day, the system unlocks the user's ability to use the map to locate GEDI data
5. The system asks the user if they wish to end their day
6. The system uploads relevant data on user activity for the day

4.1.2 The system displays a map of GEDI footprints for the user

GEDI footprints are the main source of data that will be verified with our application, so providing the user with an interface to see where they are is vital. It should display nearby locations that are acceptable for a user to be.

1. The system makes an API call requesting a map
2. The system displays the data returned as a map
3. The system makes a call to the server for some GEDI footprints
4. The system pinpoints GEDI data on the map that is displayed

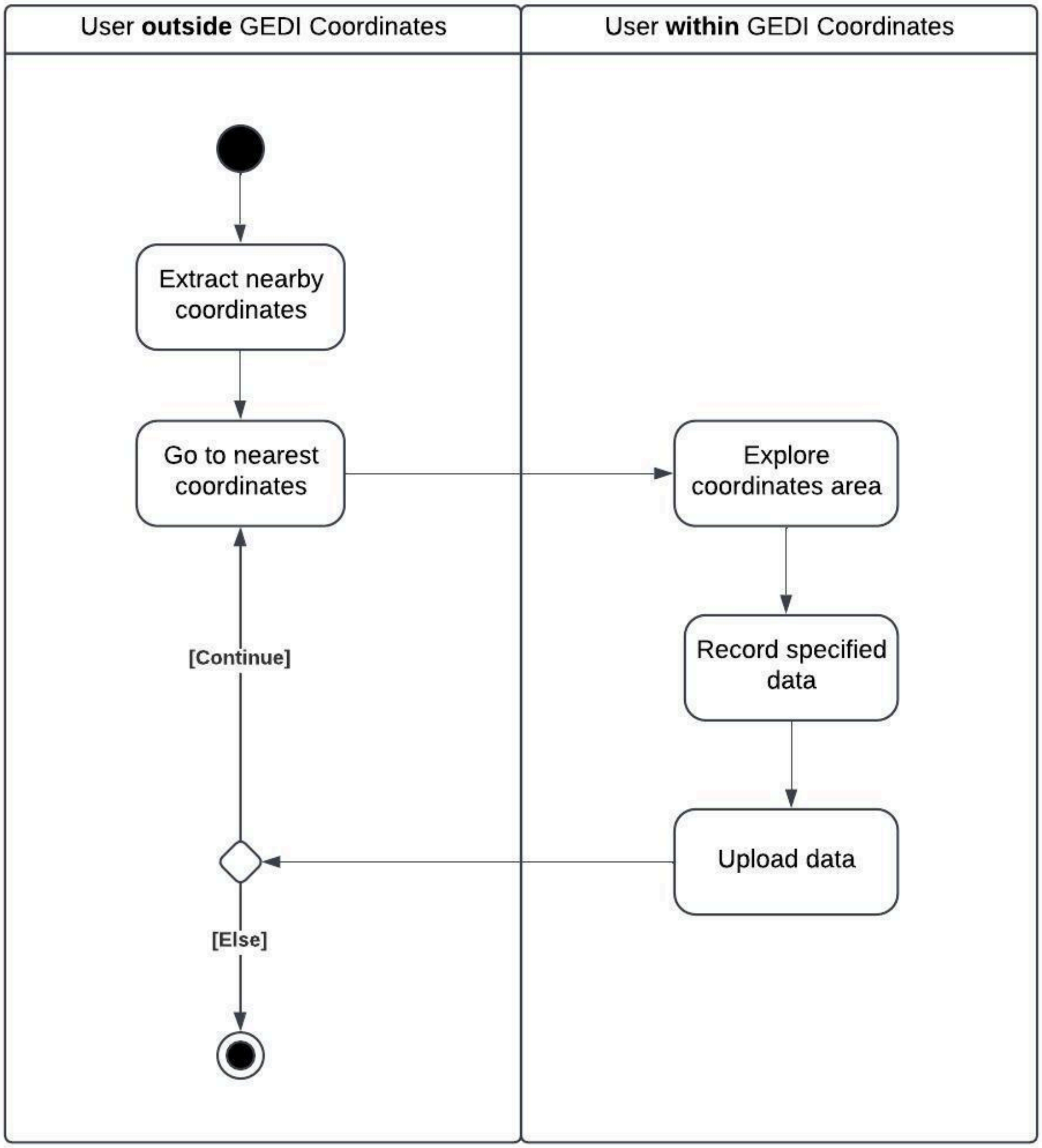
4.1.3 The system shows a route to a GEDI footprint

It is assumed that a user may be going to somewhere that is new to them, so a route will allow the team to do its part by guiding the user to where we want them to be. This is necessary especially for individuals like our client who may be doing research in new areas.

1. The user selects a GEDI footprint
2. The system makes an API call to get a route
3. The system displays a route on the map

4.1.4 The system confirms the user is within a GEDI footprint (+/- error from client)

This is one of the most important pieces of our project since a user must be in a GEDI footprint to have acceptable data. Once a user is within the footprint, they are able to both contribute and receive funding for their efforts. Offline is very important due to the remote nature of many of the GEDI footprints. Many of the prior requirements will be important for the implementation of this requirement.



1. The system will provide a user with a route to a GEDI footprint.
2. The user must choose to download the map if following the route.
3. The system will download the map onto the users phone.
4. The user can begin following the route.
5. The system will track the GPS location on the phone.
6. The user will arrive in the GEDI footprint.
7. The system will confirm arrival within an acceptable error.
8. The system will prompt the user to collect data.

4.2 Performance Requirements

In this section requirements, we will be outlining the efficiency and resource allocation of our system. Performance of our system will be integral, as many users from across the globe will have access to our application. Thus, ensuring that the speed, efficiency and resource allocation of our technologies are up to par would make our users have to worry about one less thing when interacting with our systems.

4.2.1 Android Application

1. **Resource Efficiency:** Many users will be using older versions of Android, thus ensuring that API calls and uploading of data will not be taxing on the user's device is essential.
2. **UI Performance:** Users must be able to navigate through the different interfaces seamlessly, with smooth transitions between UI elements to ensure user-friendliness of our application.
3. **User Experience:** Fast and efficient recording, verification, and upload processes will make the app more user-friendly. Long wait times or slow performance can frustrate users, potentially leading to low adoption or abandonment of the app.
4. **Scalability:** Many users will be using our application simultaneously. Ensuring that no issues arise when making API calls and uploading data is critical to the performance of our application.
5. **Efficiency:** Recording, verifying, and upload of data should not be taxing on the device. Optimizing processes to minimize resource usage and reduce upload time while maintaining data integrity and accuracy is integral.
6. **Compatibility:** Our application should be compatible with older versions of Android, ensuring that backwards compatibility would not render issues in our application.

4.2.2 Mapbox

1. **Responsiveness:** Users should be able to interact with the map with smoothness and no lag. Slow responsiveness can disrupt user experience and hinder with the user's ability to look at points and details on the map.
2. **Scale:** Mapbox should only display nearby coordinates. This would help with system performance as well as making the usage of the map more responsive.

4.2.3 AWS Database System

1. **Scalability:** Our database will be constantly storing new data related to images, audio recordings, and user information. Ensuring that our database can maintain a flow of new data being stored from across the globe is crucial to the success of our system.
2. **Resource Efficiency:** Over time, outdated data will no longer need to be stored in our database. Regularly clearing this data will improve both the efficiency and responsiveness of our database.

4.3 Environmental Requirements

For our environmental requirements, we will be talking about ease of use of our systems as well as other functionalities that, while important, should not be prioritized over our previous requirements. Deploying our app in remote areas would mean that we should expect requirements that would not be typical in other applications that are used in more urban settings.

4.3.1 Android Application

1. **UI:** Many users will have little to no experience using technology proficiently, let alone know how to read a map. Thus, a simple UI that can guide users on what to do will be invaluable.
2. **Usability:** Our app should be easy to use with no prior extensive training required. Since many of our users will have little to no technological proficiency, we want our app to be approachable by people of many different backgrounds.
3. **Data Tampering:** Our data will be valuable to researchers. Thus, ensuring that our data cannot be tampered with is essential. One way to do this is to have data capture only possible using our app, and users cannot upload data from their device storage.
4. **Offline Functionality:** Our app will be used in remote areas with little to no internet connection. Ensuring that our app can still be used to gather data even without internet connection is invaluable.

4.3.2 Mapbox

1. **Range:** User should only be able to see coordinates that are within their range. Users should not have to worry about coordinates that are too far for them to care about.
2. **Accuracy:** Coordinates should be accurate within a margin of error. This level of accuracy is critical for ensuring the validity of geospatial data used in the app, particularly for pinpointing locations and calculating carbon footprints.
3. **Clutter:** The map should not overwhelm the user with unnecessary details. Users should be able to locate a coordinate point with a route and as little details that would obstruct the route as possible.
4. **Offline Functionality:** In remote areas with no internet connection, users should still be able to use the map to locate where they are as well as where to go. By downloading the GEDI coordinates before embarking on their journey, users should still be able to continue using the map offline to know where to go and gather data.
5. **Alternative:** Users who have difficulties reading a map should be able to locate a coordinate area by using a compass instead. This would help inexperienced users be able to use our application without having to rely on reading a map.

4.3.3 AWS Database

1. **Security:** Our database should be secure with access granted only to authorized users, as it will contain sensitive data on user information as well as valuable data collected for research purposes.
2. **Authenticity:** Data collected should be authentic and not tampered with in any way. We want our community science researchers to trust us, and cannot allow fake data to be sent over to them.

5. Potential Risks

After outlining the requirements that need to be fulfilled for a successful application, this next section is going to outline the foreseeable risks that may come with the application or its functions. Being up front about these risks is important in ensuring the clients have a good understanding of the full product. The potential impact from any identified risks will also be discussed to the best of our ability.

The first risk to be discussed will be the device not being configured properly before going offline. This mainly refers to offline navigation, if the sensor is not properly functioning so the distance and/or direction is not correct. An incorrect measurement on either of these technologies will result in the application confirming that the user is inside a GEDI footprint when they are not. Incorrect data will be sent to the NASA Globe app, and false data will be attached to GEDI footprints. That is a huge setback to the way community science is right now. Additionally, the user will be led to a completely separate area than where they are expecting to end up. If users are in remote areas, and try to navigate back to their starting location, they are at risk of getting lost. It will be imperative that the sensors are properly calibrated before use to protect the integrity of data, but much more importantly, to protect our users.

Another risk involved is with our navigation. Since a map will be provided to a GEDI footprint, it is then possible that the route may go through unsafe terrain, locations, or unknown obstacles. In most cases, if these are locals then this may not be a common problem. Although an unsafe route still potentially can put a user in a dangerous situation. Clearly this should be avoided as much as possible. So, it is important to warn them how the route was generated and to receive user feedback so that this can be mitigated.

A risk that may come with the data uploads is that there may be a poor connection that can stop photos from being uploaded. The expected user base will most likely be in an area with very poor service. It will be difficult to test data upload with poor service purposefully without being in those types of areas. If a user is unable to upload their data, or is uploading incomplete photos, they will not get paid for their efforts despite doing the work. It will also be very taxing on their data, if that is what they are using, since that most likely means they will be trying repeated uploads. This will still use their data, but without any success. Our app will then be doing the exact opposite of what our clients want which is costing users money and no advancement in community science.

Another risk that comes with our application may be with poor camera quality. Since this application is made to use on older androids, the camera quality may not always be the best. For data verification this is important to know what the image is. This data is going to be uploaded to NASA for additional data measuring so that they know what is in these areas. Without quality images, it may not add anything helpful and users would be getting paid for images that are unidentifiable. Even if a team was able to review these, it can take quite a while for a group of people to review every single image that is passed through our app. It would be counterproductive to community science.

Continuing with images, another risk comes with not being able to identify duplicate photos accurately. Since it is important to know the uniqueness of these GEDI footprints, repetitions need to be identified to accurately observe the data. This will not only flood our database, but users would be wrongfully getting compensated for work that has not helped at all. Like the precious risk, it would be counterproductive to sit and try to differentiate each individual photo. This risk does not pose too much of a threat, and neither does the previous one, but it is important to weed these out so that there are accurate advancements in the data collection.

6. Project Plan

Below is a Gantt chart showing the estimated completion of the main milestones of the project. The most important milestone and the basis of the project is a map displaying the GEDI footprints and the user. This is the major aspect of the project and every other feature relies on it so it will be the first milestone to be completed. It is expected to take about three weeks to complete the groundwork and then everything else can be built on top. The second milestone is location verification, this is to ensure that the user is at the GEDI footprint and the location data has not been tampered with. After that has been completed, we will move on to implementing data collection. This will mainly be taking pictures or any other form of data collection required. Then, we will implement offline collection. This may be more complicated so the estimated completion time is almost a month. It is important that this works perfectly since most users will be out of range of an internet connection and we must still be able to verify their location and allow them to collect data. Once the app is able to properly collect data, we will ensure that the user has not tampered with it during the time of collection. The next milestone helps verify this by encrypting the locally stored data when offline. This prevents the user from going into the app files and changing improperly collected data to be valid. The last milestone will be uploading the data collected offline to a server to be used. This should not be very difficult but the data will need to be properly formatted and uploaded to the server.

Task Name	Duration	Start	Est. End	2nd Half January	1st Feb	2nd Feb	1st Mar	2nd Mar	1st Apr	2nd Apr	1st May
Functional Map with GEDI Data	21 Days	1/13/2025	2/3/2025								
Location Verification	14 Days	2/3/2025	2/17/2025								
Data Collection	14 Days	2/17/2025	3/3/2025								
Offline Data Collection	28 Days	3/3/2025	3/31/2025								
Data Verification	14 Days	3/31/2025	4/14/2025								
Data Encryption	14 Days	4/14/2025	4/28/2025								
Upload Data	11 Days	4/28/2025	5/9/2025								

7. Conclusion

Conservation efforts, although important, can be damaging to rural communities. The people in these communities can be displaced or have their lives changed in the pursuit of conservation. It is our hope that we can work together with these communities to help conservation as well as the people that live in the area.

We are striving to create an app that can give back money to rural communities as well as collect important data that is accurate. Data verification is so important for this project because we are hoping to pay citizen scientists to go to GEDI footprint coordinates and gather data. Using a map on our app, the user will be directed to the closest GEDI footprint and will be able to record data in the area, even if offline. When connected to the internet, multiple GEDI footprints will be saved from an area around the user. The user can then go into remote areas without service and gather data. Once they are back in an area with service, the app will upload the data to a server where it can be viewed. After a user collects enough data and it is verified, they will be compensated for their work accordingly.

This document lays out what we need to build our finished product. We have researched a solution and various problems that could arise. Our functional, performance, and environmental requirements have been researched and in this document, we have created a plan to build an app that can help both conservation efforts and any communities that have been impacted by those efforts.