Geology Field Data Collection System

Master's Project Report (Progress Report)

California State University, Chico

Committee Chair: Dr. Edward Roualdes

Committee Member: Dr. Hannah Aird

Committee Member: Dr. Kathy Gray

By

Govardhan Reddy Baddala

December 2025

Project Overview

Introduction

The Geology Field Data Collection System is a web-based solution designed to support geologists, researchers, and students during fieldwork in remote environments where internet connectivity is limited or unavailable. Traditional methods of geological data collection have long relied on paper-based notebooks, which are prone to loss, damage, and transcription errors. While digital tools exist, many fail to meet the fundamental requirement of working reliably in disconnected environments.

This project bridges that gap by implementing an offline-first architecture that ensures data collection can continue uninterrupted regardless of network availability, while automatically synchronizing collected data to a centralized database when connectivity is restored. The application is designed to run across multiple device types, including smartphones and tablets used in the field and desktop computers used for review and analysis.

Problem Statement and Objectives

Field geology work presents unique challenges because researchers and students often operate in remote locations where cloud-based applications are unreliable. Paper notebooks remain functional but create problems such as risk of loss, weather damage, inefficient organization, and time-consuming manual entry into digital systems after fieldwork. Accurate location tracking is also critical, yet manual recording of GPS coordinates is slow and error-prone. This project aims to develop a robust offline-capable data collection system that stores data locally, syncs automatically when online, supports GPS-based location capture, maintains data integrity, and provides an intuitive interface usable under field conditions.

System Summary

The system follows an offline-first model where field entries are stored locally on the device and synchronized to a centralized database when connectivity returns. The system supports structured measurements and notes, emphasizes validation to improve accuracy, and is intended to be deployable and maintainable in academic environments.

Contents

Chapter 1

Introduction

1.1 Overview of Geology Fieldwork

Geological fieldwork is essential for collecting real-world observations that support the interpretation of Earth processes. Field measurements often include sediment and grain characteristics, hydrologic observations, site descriptions, photographs, and geographic location information. Accurate data collection is critical because field measurements typically cannot be repeated easily once researchers leave the site.

1.2 Challenges in Field Data Collection

Many field locations have weak or nonexistent internet access, making cloud-dependent tools unreliable. As a result, students and researchers commonly rely on handwritten notebooks or ad-hoc spreadsheets, which introduces inefficiency and increases the risk of mistakes. These constraints motivate systems that support reliable field data capture in disconnected environments.

1.3 Need for an Offline-First Solution

Offline-first systems prioritize local data capture and persistence on the device, allowing fieldwork to continue uninterrupted without internet access. When a connection becomes available, the system synchronizes data to a central database. This approach reduces the need for manual transcription, improves data quality, and enables more consistent data organization.

1.4 Project Motivation

The motivation for this project comes from the conflict between the need for reliable digital data and the reality of field environments where connectivity is unreliable or non-existent. A well-designed offline-first tool can combine the robustness of paper-based collection with the organizational advantages of structured digital storage.

1.5 Organization of the Report

This report presents the project context, challenges, objectives, proposed methodology, and current progress. It also outlines the remaining work for next semester, a proposed timeline, and the expected outcomes of the completed system.

Chapter 2

Problem Statement

2.1 Introduction to the Problem

Geological field research and education face persistent challenges in data collection due to the mismatch between modern digital workflows and the reality of remote field environments. Most digital tools assume constant internet access, while many field sites have unreliable or no connectivity. Traditional paper-based methods work anywhere, but they introduce other limitations that reduce data quality and efficiency.

2.2 Connectivity Limitations

One of the most significant challenges in geological field data collection is the frequent absence or unreliability of internet connectivity. Geological research often occurs in remote areas such as mountain ranges, deserts, river valleys, and coastal regions where cellular coverage is weak, intermittent, or completely unavailable. Cloud-based applications can fail to save data, time out, or interrupt the workflow due to synchronization failures. Because field schedules depend on site access, weather, and daylight, data must be collected during critical windows regardless of connectivity conditions.

2.3 Paper-Based Workflow Limitations

Paper notebooks remain functional in any environment, but they introduce limitations that compromise reliability. Notebooks are vulnerable to environmental damage such as rain, humidity, and dirt, and field notes can become illegible or lost. Transcribing paper notes into digital systems after fieldwork is time-consuming and error-prone, introducing transcription mistakes and delaying analysis.

2.4 Location Tracking Challenges

Accurate spatial context is fundamental in geology because the location of a measurement can be as important as the measurement itself. Manually recording GPS coordinates is slow and error-prone, and correlating location information with observations requires manual cross-referencing. Researchers may also need coordinates in multiple formats for mapping and reporting, which increases time and error risk.

2.5 Data Organization and Retrieval Issues

Traditional methods create challenges in organizing and retrieving field data collected over multiple days, locations, and participants. Without structured storage and consistent formats, it becomes difficult to query data, compare datasets, or perform comprehensive analysis. Differences in terminology and recording style reduce dataset consistency and overall value.

2.6 Educational Context Challenges

In educational settings, instructors must review data from multiple students, verify field participation, and provide feedback. Traditional workflows make oversight slow, limit timely correction of errors, and reduce time available for analysis-focused learning. Students often spend excessive time transcribing data instead of learning interpretation and analytical methods.

2.7 Need for a Comprehensive Solution

A reliable solution must support offline data entry, structured storage, automatic location capture, and seamless synchronization when connectivity returns. It should remain usable in field conditions and be deployable and maintainable in academic and research settings without extensive technical support.

Chapter 3

Objectives

3.1 Primary Objective

The primary objective is to develop an offline-capable web application that enables geology field data collection without requiring internet connectivity and synchronizes collected data to a centralized database when connectivity becomes available.

3.2 Functional Objectives

1. Support structured entry of common field measurements (e.g., grain size, sediment observations, flow measurements).
2. Capture timestamps automatically for each observation.
3. Capture GPS coordinates automatically when available, with a manual fallback option.
4. Allow users to attach notes and field context to each measurement.

3.3 Technical Objectives

1. Store data locally on the client device using a browser-based database.

2. Synchronize data automatically and reliably to a server database when online.
3. Provide validation and error handling to improve data accuracy.
4. Support a deployment approach that is portable and maintainable.

3.4 Educational and Practical Objectives

The system should be usable by students during field trips and support instructors by providing centralized access to collected datasets for review, feedback, and analysis.

Chapter 4

Background and Related Work

4.1 Offline-First Systems in Other Domains

Offline-first data collection systems are used in domains such as healthcare, environmental monitoring, and humanitarian response. These systems prioritize local persistence and later synchronization to address unreliable connectivity and reduce data loss.

4.2 Advancements in Web-Based Data Storage

Modern browsers support persistent local storage mechanisms (such as IndexedDB), enabling applications to store structured data locally and remain functional offline. This makes offline-capable web applications feasible without requiring native mobile app development.

4.3 Existing Field Data Collection Tools and Limitations

Although digital field tools exist, many rely on continuous internet connectivity or do not support workflows tailored to geology education and field measurement needs.

Tools also vary widely in how they handle structured data entry, validation, synchronization, and multi-user organization.

4.4 Research Gap

Geology fieldwork faces the same connectivity constraints as other disciplines but has fewer widely adopted offline-first solutions. This project addresses the gap by adapting proven offline synchronization patterns to geology field education and research use cases.

Chapter 5

Methodology

5.1 System Overview

The proposed system follows an offline-first architecture with three key components: a client-side web application, local device storage, and a centralized server database used for synchronization. This structure ensures that field data can be collected immediately and preserved even when internet connectivity is unavailable.

5.2 Offline-First Design Approach

The offline-first approach emphasizes local persistence as the default behavior. Users can create, edit, and store measurements without relying on the network. Synchronization is performed automatically when a connection becomes available, reducing manual steps and preventing data loss.

5.3 User Interface Design

The user interface is designed for field conditions, prioritizing clarity and ease of entry on mobile devices. Forms are structured for key geology measurements and support additional notes for context. Validation rules help prevent incomplete entries and reduce common errors in numeric fields and required categories.

5.4 Data Capture Modules

The system is designed to support multiple measurement types commonly collected during field exercises and research. Examples include:

- Grain size and sediment classification entries
- Flow-related measurements (e.g., depth, velocity, distance from bank)
- Field notes and optional image attachments

5.5 Location and Time Metadata

Each entry includes timestamps to preserve a chronological record of field activity. GPS coordinates are captured automatically when available, supporting accurate mapping and spatial analysis. A manual coordinate option is included as a fallback if GPS is unavailable or unreliable.

5.6 Data Storage and Synchronization

All field entries are stored locally using a browser-based database to ensure persistence across sessions. When connectivity returns, the system synchronizes local data to a centralized database. Synchronization is designed to be transparent to users and resilient to intermittent connections.

5.7 Deployment and Maintainability

The system is intended to be deployable in academic and research environments with minimal overhead. A containerized deployment approach supports consistent setup across environments, and a reverse proxy can be used to manage routing and performance for production deployments.

Chapter 6

Current Progress

6.1 Requirements Analysis

Project requirements were identified by examining common workflows in geology field education and research. Key requirements include offline operation, structured data entry, consistent metadata capture, and a workflow that minimizes manual transcription.

6.2 Technology and Architecture Planning

An offline-first architecture was selected to ensure reliability in disconnected environments. The design includes a web-based interface for field entry, local persistence on the device, and synchronization to a centralized database when online.

6.3 Initial Design Decisions

Initial measurement categories and form fields were defined to support common field tasks. The application structure emphasizes consistent data formats, timestamps, and location tagging to improve organization and later analysis.

6.4 Challenges Encountered

Key challenges include ensuring reliable behavior across varying connectivity conditions, designing a user interface suitable for outdoor use, and planning for conflict handling during synchronization across multiple devices.

6.5 Status Summary

At the current stage, implementation and testing are ongoing. The next semester will focus on completing the application, integrating synchronization, and conducting usability and functional testing.

Chapter 7

Proposed Work for Next Semester

7.1 Frontend Completion

Complete the full set of field data forms, improve usability, and finalize validation rules. Ensure mobile responsiveness and consistent input handling for field use.

7.2 Synchronization Implementation

Implement automatic synchronization between local storage and the server database. Ensure synchronization handles intermittent connectivity and preserves data integrity.

7.3 Testing and Validation

Conduct functional testing (offline entry, online sync behavior, error recovery) and usability testing in mock field conditions with student users. Collect feedback and iterate on improvements.

7.4 Deployment and Documentation

Finalize deployment packaging and configuration. Prepare documentation including installation steps, a user guide, and a testing report suitable for academic submission.

Chapter 8

Proposed Timeline

8.1 Short-Term Milestones

- Complete user interface forms and final data model
- Implement local persistence and validation rules

8.2 Mid-Term Milestones

- Implement and test synchronization logic
- Improve error handling, recovery, and conflict behavior

8.3 Final Milestones

- Conduct field-style testing and gather feedback
- Finalize deployment package and documentation
- Complete final report writing and revisions

8.4 Schedule Table

Table 8.1: Proposed Schedule for Next Semester

Time Period	Planned Work
Weeks 1–3	Finalize UI, forms, and validation
Weeks 4–6	Implement local storage + sync integration
Weeks 7–9	Functional testing (offline/online, sync reliability)
Weeks 10–12	Usability testing, feedback, refinements
Weeks 13–15	Deployment packaging, documentation, final report

Chapter 9

Expected Outcomes

9.1 Technical Outcomes

The project will deliver an offline-capable web application that reliably stores field observations locally and synchronizes them to a centralized database when online.

9.2 Educational Impact

The system is expected to reduce the time students spend on manual transcription and improve the quality of collected field data. Instructors benefit from more organized data and improved oversight across student submissions.

9.3 Reliability and Data Quality Improvements

Structured forms, validation, and automatic metadata capture will reduce common errors in field notes and improve consistency across datasets collected by multiple users.

9.4 Scalability and Reusability

The offline-first architecture provides a reusable foundation for future field data collection applications in geology and other disciplines that require data capture in disconnected environments.

Chapter 10

Conclusion

Proposed methodology for an offline-first geology field data collection system. The core challenge is enabling reliable digital data capture in environments without dependable connectivity while avoiding the errors and inefficiencies of paper-based workflows. The proposed approach supports local persistence, structured organization, and seamless synchronization when connectivity becomes available. Future work will complete implementation, testing, deployment, and final evaluation.