



Photo by Carolina May

Ecological Monitoring Program

Data Management Plan

March 2026, First Edition

TABLE OF CONTENTS

Introduction

Data

Documentation and Metadata

Ethics and Legal Compliance

Storage and Backup

Selection and Preservation

Data Sharing

Responsibilities and Resources

Appendix

Introduction



The Ecological Monitoring Program within the New Mexico Forest and Watershed Restoration Institute (NMFWRI) collects ecological data from the forested uplands of New Mexico and riparian areas of the Middle Rio Grande watershed. These data are collected by field crews and stored in relational databases for repository, sharing, and analysis purposes.

Ultimately, collected data is analyzed and used to inform adaptive management reports used by land managers and decision-makers, but is also publicly available. This ensures our data collection program is aligned with the goals of the Institute at large, which includes providing easily accessible information to all user groups – including the general public – who wish to understand the efficacy, longevity, and other impacts of restoration treatments and disturbances on our fire-prone landscape.

The present document defines all parts of the data pipeline, including a history of the Institute’s data management processes, data collection methodologies, technical solutions to various problems, ownership of data, and data sharing procedures. This will be updated periodically as changes to the data management process occur.

Data



Description

Monitoring data is collected over a range of ecological zones that broadly constitute the watersheds and forested uplands across New Mexico. Funding sources historically dictated the data collection protocols used in many of our projects, which were implemented under the New Mexico Collaborative Forest Restoration Program (CFRP) for forested upland regions of the state, or via the Greater Rio Grande Watershed Alliance (GRGWA) for bosque regions of the Middle Rio Grande Valley. The CFRP was a program

through USFS that provided grant funding for collaborative forest restoration projects on Federal, Tribal, State, County, or Municipal land across the state of New Mexico₁. GRGWA is a collaboration between Soil and Water Conservation Districts (SWCDs), Pueblos, public agencies, and other stakeholders that aims to restore riparian ecosystems along the Upper and Middle Rio Grande watersheds.

Technicians use a range of tools and ecological knowledge to gather data from monitoring plots, randomly sampled from the project area. For data collection methodologies, we use protocols established by the USFS and scientific literature. A brief description of the methodologies follows; please refer to the monitoring handbook on the NMFWR website for full descriptions of them.

We use our Uplands Protocol for collecting data from forested upland areas of New Mexico which encompass several ecological zones, ranging through piñon-juniper, pine-oak, and mixed conifer belts. Our typical uplands data collection protocol involves a series of restoration treatments followed by a 15-year monitoring period – a requirement established by the aforementioned CFRP program. We then collect data immediately pre- and post-treatment and at 5-year intervals; these data are synthesized into annual reports documenting how well restoration treatments accomplish project goals, how long treatment impacts last, and management implications. Collection methodologies are based on the Common Stand Exam (CSE) – a set of protocols for collecting data on tree stands. Our methods typically use a 1/10-acre plot (1/20-acre plots for some historical projects) to sample the population of individual trees and snags, as well as a 1/100-acre nested plots to sample seedling and sapling regeneration, canopy cover and understory aerial and ground cover, dead woody biomass using Brown's transects to measure 1-, 10-, 100-, and 1000-hour fuels, plus litter, duff, and ladder fuels via planar intercept.

Our Riparian Protocol encompasses the Middle Rio Grande Watershed and its tributaries and focuses on the cottonwood bosque ecosystem of the Rio Grande. The cottonwood bosque is an ecosystem unique to the Rio Grande Valley composed of floodplains forested by the Rio Grande cottonwood and faces unique challenges due to an extensive history of anthropogenic effects. These data are collected in collaboration with the Greater Rio Grande Watershed Alliance (GRGWA) and also typically use 5-year incremental monitoring periods, as well as pre- and post-treatment and post-disturbance.

NMFRI employs CSE₂-based or adapted metrics from the methodologies used in the Uplands Protocol for GRGWA projects in riparian forests by modifying sampling densities and other metrics, and by including additional information such as soil texture. Older

projects use another set of protocols better suited to the Rio Grande cottonwood ecosystem: Montane Riverine New Mexico Rapid Assessment Method (NMRAM), Bosque Ecosystem Monitoring Program (BEMP), and a modified plot-based protocol. These protocols also measure overstory and understory plant community composition, vertical and horizontal structure and diversity, soil health, anthropogenic disturbance, and invasive species presence. GRGWA field coordinators fill out qualitative inspection reports that contain information about the conditions of the project site, noting details of interest in regards to how well the site is responding to restoration efforts and how well treatments were administered.

All protocols collect reference photos from each plot site. Currently, these photos capture the perspective and conditions in each cardinal direction, as well as along any fuels transects; historically, particularly with NMRAM sampling, the photos captured various plant communities within project areas. Photo points along cardinal directions demonstrate landscape changes over time and serve as reference points for future crews who will need to visit the plot.

The Institute also collects and creates geospatial information on our project sites and plots that serve the creation of maps for reports and contribute to our Vegetation Treatment Database. Geospatial information typically constitutes plot locations as points or project boundaries as polygons. The Monitoring Program collaborates with the GIS Program regarding the maintenance of geospatial data.

Data Collection

Monitoring crews collect data on site in the field. Technicians use tools such as measuring tapes for plot establishment and various measurements, diameter tapes for tree/shrub diameter, fuel gauges for classifying and measuring coarse woody debris and fuel depth, clinometers and hypsometers for height-related tree data, spherical crown densimeters for measuring canopy cover, and crew expertise for observational data on soil textures, ground and aerial cover, and landscape health.

In our current workflow, we use Panasonic Toughbooks running Windows 11 for data entry. Technicians enter data into FEAT-FIREMON-Integrated (FFI) in the field, quality control (QC) the data using a checklist before leaving plot to ensure all the data was collected, and do a final QC process in office using an R script, which enforces value constraints. Data are backed up via an XML file – internally termed an admin export, to describe the admin units used for project structure – and a .bak backup file and uploaded to the Institute’s Microsoft SharePoint site. SharePoint is maintained and

backed up by New Mexico Highlands University's IT Services.

To ensure crews collect the highest quality data possible, the Institute uses quality control (QC) checks and a quality assurance (QA) team to ensure that data collection operates within measurement quality objectives (MQOs). These three concepts ensure that our crews gather data as consistently and accurately as possible. The QC process enforces accuracy, while the QA team utilizing MQOs enforces precision.

Crews currently QC data in the field by exporting the dataset as CSV files and then executing a R script, which compares collected values against a list of constraints and serves the affected rows on an HTML page. We are in the process of implementing a Python application using the Streamlit package as a GUI, which will allow us to QC data directly through the backend databases of FFI or entire XML exports. Ultimately, the data entry process will provide enforcement before the data makes it into the database.

For QA, we maintain a small senior-level crew who collect data from a small sample of recently visited plots and compare their data with the collected data. When comparing data, the QA crew looks for values that fail our MQOs. We have established and refined a set of tolerances that data values can exist within in comparison to the values collected by the QA crew. Being able to collect data in a consistent manner across all field crew members is essential for gathering reproducible data.

Technical Solutions

Paper data sheets limit analyses on the data and FFI lacks customizability and usability for end users to perform cross-project analyses or answer novel questions. FFI interfaces with a MSSQL database on the backend, but the database structure is intended for programmatic access only; using SQL queries to access the data is complex and difficult. Additionally, the interface FFI provides for querying data limits the information one can derive from the data. As such, we have established a new database schema for new data entry and migration from FFI formats. This schema uses custom, semantically meaningful identifiers for all tables and utilizes table formats easily understood by end-users.

We used a spreadsheet to define the database tables, their fields and data types, and specify primary and foreign keys for relationships between tables. A Python script translates the tables into a set of SQLAlchemy models, allowing us to perform SQL

operations in any type of relational database. SQLAlchemy is an Object-Relational Mapper (ORM) for Python that allows for SQL dialect-agnostic operations and versioning. This means we can create a database schema with empty tables, as well as a stepwise list of any updates made to the database using any SQL dialect; there is still a possibility that our internal repository will use PostgreSQL rather than MSSQL due to storage limitations, so this allows us to easily create a database in either version using the same code.

Once databases are created, an Extract, Transform, Load (ETL) process model written in Python migrates data stored in the FFI files into the database. The Extract phase of this process reads in the XML files produced by FFI and turns the file into a set of data frames representing the various database tables. The Transform phase consists of custom logic to rearrange the data into the format we use for our database. This process includes creating custom identifiers for each table, joining and pivoting tables into the desired format, creating calculated fields, aggregating fields where needed, and creating new tables for selected columns to better represent the structure. Finally, the Load phase generates MERGE INTO SQL statements that insert the previously transformed data directly into the database, which allow for inline duplicate checks.

The Institute also created a data management system in Microsoft Access, termed the NMFWRJ Ecological Restoration Data System (NERDS). NERDS provide both for the creation of custom database tables and a data entry system which operates on top of the database to allow for offline data entry at multiple levels in the database hierarchy. We created this due to the needs of the GRGWA project to store data gathered via several distinct data collection methodologies and connect them in analyses. The data entry portion enables users to add and modify projects, plots, and sample events plus enter data observations in the field. Additional modules perform various database functions which include: transferring data from Access into a remote database, transferring data from a remote database into Access, updating data values from Access into a remote database, changing small errors in identifiers, deleting all data under a project identifier, and generate a table of database attributes from the Access metadata. All modules were written in VBA.

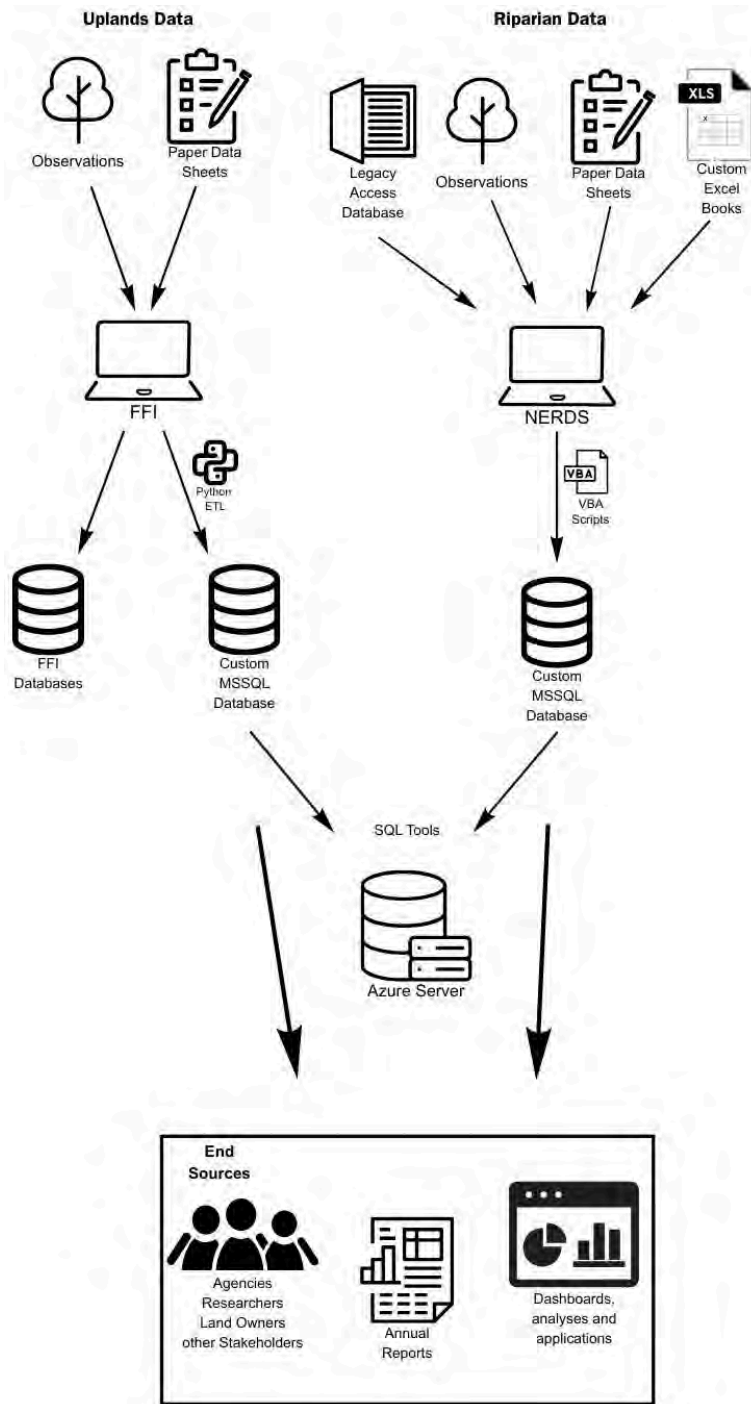


Figure 1. Data Management Workflow

Documentation and Metadata



All data collection protocols have extensive documentation available covering the exact methodologies, the rationale for their use, and their scientific bases; GRGWA and CFRP protocols each have separate documents for field crew training that outlines the specific collection methods and how to perform them. These guides are publicly available on the NMFWRI website.

As previously mentioned, database schemas are included in the Appendix of the data management plan and are provided for both the Upland and Riparian databases. Database schemas describe the tables in a database and how the tables relate to one another; this includes column names and definitions, data types, primary keys, foreign key relationships between tables, and other constraints on what values can be entered into each attribute. Hopefully, these will provide ample introduction and explanation of the data to users who hope to become familiar with it.

We make metadata available for all data collected in the Ecological Monitoring Program. The Environmental Data Initiative (EDI) stores all metadata associated with our Uplands project as well as an archive of the data itself. Metadata associated with our GRGWA project is available in the Appendix section of this document – this data is in the process of being quality controlled and will be made available with all metadata when the process is complete.

Ethics and Legal Compliance



All data produced by the Monitoring Program are collected through field personnel, so all data belong directly to NMFWRI. Some projects are on private or tribal lands and although project/plot locations are publicly available, land managers are made aware of this and implicitly consent to it; on tribal land, we agree to a moratorium on photographs

outside of necessary photo points outlined in monitoring protocols.

The NMFWRRI makes data publicly available in accordance with open data standards³ and licenses datasets under the Creative Commons Attribution 4.0 International license⁴. This licensing allows users to copy, modify and distribute data in any format for any purpose, including commercial use. Users are only obligated to give appropriate credit (attribution) and indicate if they have made any changes, including translations.

Storage and Backup



Currently, the Institute utilizes Windows Server 2016 Standard hosted on a desktop located in the Monitoring Program Office, which runs a Microsoft SQL Server instance. This allows us to remotely access our data. Staff members access the server through a VPN connection secured through the New Mexico Highlands University Information Technology Services. Through the VPN, staff connects directly to the SQL server via Microsoft SQL Server Management Studio or utilizes a remote desktop connection for maintenance.

A Bash script executed by Windows Task Scheduler creates backups via .bak files every Monday morning at 05:00 of both our Riparian and Upland databases. Backups get created and stored in a directory on the Windows Server, where they remain for one year and then are deleted.

Additionally, we maintain our FFI data on this server, although it is stored on different databases. FFI does not support multiple projects in the same database, so we create new a new database for each project and back up the project via both a SQL .bak file and a XML-formatted file exported from the FFI software; we store these files on the Monitoring Program's Field Work SharePoint site. To preserve field-collected data, we scan paper data sheets in the office and upload those to our SharePoint site. Finally, we also use the SharePoint site to store any additional data needed to capture our project goals and research questions, which is maintained using a detailed file plan and file structure. SharePoint automates its backups by versioning the file library every time a modification occurs; this allows for a robust method of rolling back any changes that occur.

Currently, we are limited in our capacity to manage our server by both the NMHU network. As part of our long-term plan, we will ultimately use a cloud service to host our databases such that access is robust and publicly available.

Selection and Preservation



NMFWRI collects all data with the intention of long-term preservation. We will keep all data in perpetuity to maintain our overarching goal of being a central data source for land managers across the Southwest.

Data Sharing



Currently, the data manager – Corey Beinhart – can provide any data we have on an ad hoc basis. He can be contacted at corey@nmhu.edu.

Our sharing capabilities will expand greatly in the future and the current document will be updated to reflect that.

Responsibilities and Resources



The present document should be considered a living document to be updated regularly. The Ecological Monitoring Program Data Manager will update the Data Management Plan as changes are developed. They also maintain the data workflow and the processes behind it, and any questions should be directed to them.

APPENDIX

Databases

NERDS

References

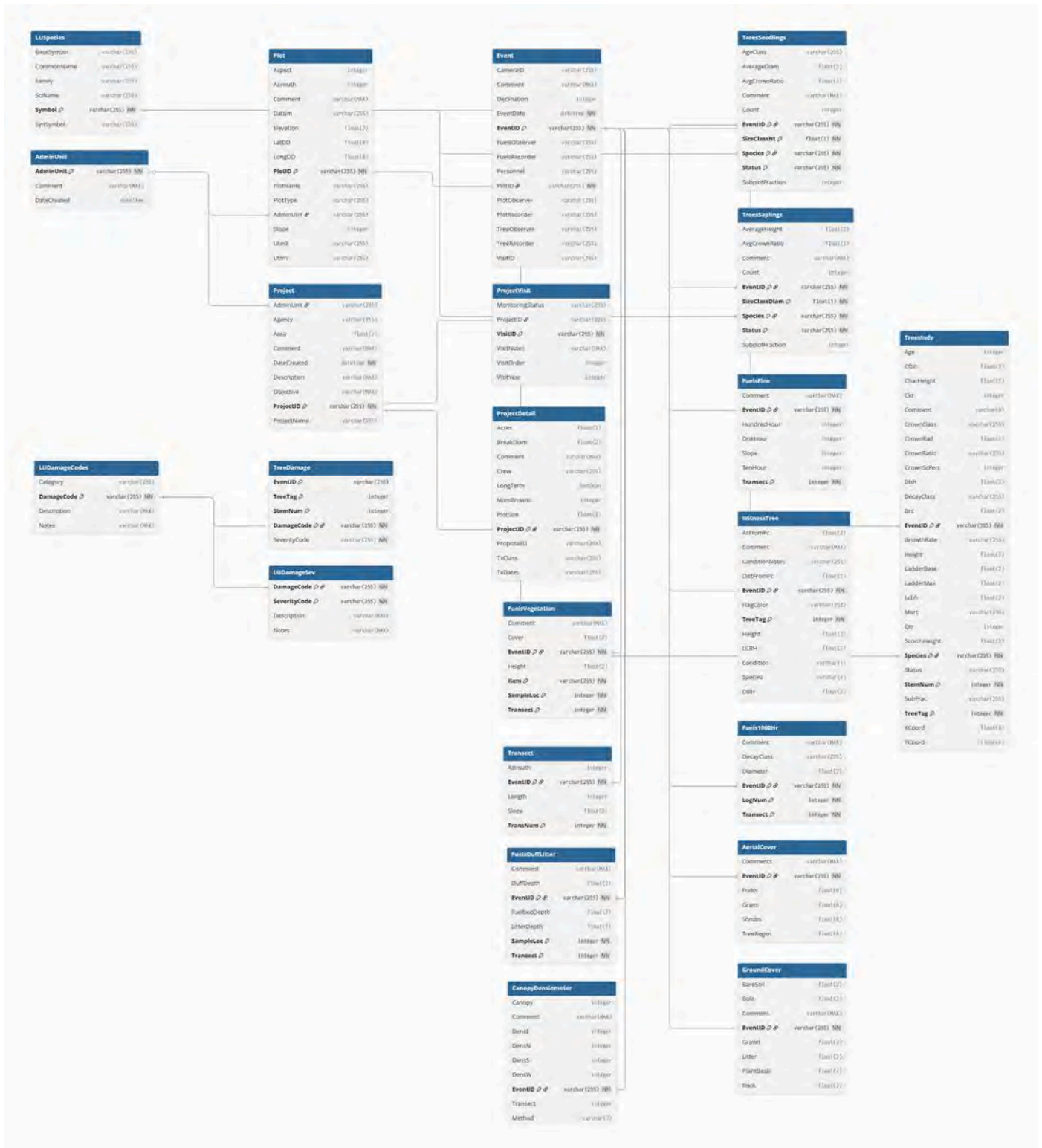


Figure 3. Upland Database Diagram

Entity-relationship diagrams provided here outline all tables in each database, their attributes, and the relationships between tables. We describe all database tables and their attributes in our database documentation, hosted on dbdiagram.io^{5,6}. This tool allows for structured definitions, visualizations, and generated database creation scripts – the Data Manager should be consulted before using these tools to ensure appropriate use. Complete definitions are omitted here due to the cumbersome nature of presenting names and attributes of each database table. Each database description (Uplands and Riparian) is publicly available for viewing via the provided URLs.

NERDS

The NMFWRP Ecological Restoration Data System (NERDS) exists as a solution to aggregating and entering field data for our GRGWA project. Methodologies used to accomplish the goals of this project and the project structure preclude our use of FFI for data entry, so exploring custom solutions was necessary. Ultimately, we settled on the creation of an application using Microsoft Access due to the ubiquitous nature of Microsoft products in natural resources, the flexibility of using a relational database on the backend, the ease of using Visual Basic for backend logic, and the same program housing the interface to create the frontend.

Because Microsoft built Access atop a SQL database flavor akin to SQLServer, we keep the schemas within Access and SQLServer aligned with one another, linked the remote SQLServer database to Access, and built a partial ETL to transfer data between the two databases. VBA support within Access also allowed us to set up additional modules to facilitate the update identifiers, transfer data in and out, update data, and generate table definitions.

Citations

- [1] NMFWRI, “Greater Rio Grande Watershed Alliance,” *New Mexico Forest and Watershed Restoration Institute*, Sep. 03, 2020. <https://nmfwri.org/monitoring/greater-rio-grande-watershed-alliance/>
- [2] K. Mahan, “NMFWRI Riparian CSE-Based Plot Sample Protocols,” *New Mexico Forest and Watershed Restoration Institute*, Jul. 2020. Accessed: Mar. 30, 2026. [Online]. Available: https://nmfwri.org/wp-content/uploads/2020/07/GRGWA_plotprotocols_Instructions_datasheets_withcheatsheets_3.1.2020km.pdf
- [3] Open Knowledge Foundation, “Open Definition 2.1 - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge,” *opendefinition.org*. <https://opendefinition.org/od/2.1/en/>
- [4] Creative Commons, “Creative Commons — Attribution 4.0 International,” *creativecommons.org*, 2016. <https://creativecommons.org/licenses/by/4.0/> (accessed Mar. 30, 2026).
- [5] dbdocs.io, “RiparianStructure,” *Dbdocs.io*, 2026. <https://dbdocs.io/corey-8f848356ef/RiparianStructure> (accessed Mar. 30, 2026).
- [6] dbdocs.io, “UplandsStructure,” *Dbdocs.io*, 2026. <https://dbdocs.io/corey-8f848356ef/UplandsStructure> (accessed Mar. 30, 2026).