Appendix F: Data Management Plan

Numerous conversations with USDA Forest Service representatives and Southwest Ecological Restoration Institute (SWERI) employees, as well as review of previous CFLRP documents, note the complexity of managing data across organizations and over time. The comprehensive data management plan outlines our current approach and will be updated as challenges are identified, and potential solutions are implemented.

NOTE: In progress. To be updated as data sources and management needs are made more clear.

Table of Contents

Tables

Figures

Glossary

Category: Derived from ArcGIS Online, like a folder in a group.

Feature: (layer type) Web feature layers support vector feature querying, visualization, and editing. They are most appropriate for visualizing data on top of basemaps.

Feature layer: A layer that references a set of feature data. Feature data represents geographic entities as points, lines, and polygons.

Geodatabase: A database or file structure used primarily to store, query, and manipulate spatial data. Geodatabases store geometry, a spatial reference system, attributes, and behavioral rules for data. Various types of geographic datasets can be collected within a geodatabase.

Group: In ArcGIS Online, a way to collaborate with other ArcGIS users and to exchange content related to a specific project or common activity.

Hosted feature layer: A hosted feature layer view is similar to a copy of a layer but is more powerful because it allows you to control more than how the layer is displayed. For example, you can create a hosted feature layer view when you need to support different editing capabilities for different groups of people.

Hosted imagery layer: Hosted imagery layers allow access to the imagery or raster data, including the pixel or cell values across multiple bands, and multidimensional data.

Items: Items are the contents made available through ArcGIS Online. Items include content such as files, layers (services), maps, scenes, apps, tools, and templates. View this list of items that can be added to AGOL. Common items stored in this Group will appear with these icons:

Layer: In ArcGIS, a reference to a data source that defines how the data should be shown on a map. Layers can also define additional properties, such as which features from the data source are included.

PII: Personally Identifiable Information. Information that can be used to distinguish or trace an individual's identify, either alone or when combined with other personal or identifying information, that is linked or linkable to a specific individual. PII should be evaluated on a case-by-case basis to determine the specific risk that an individual can be identified (https://www.gsa.gov/reference/gsa-privacy-program/rules-and-policies-protecting-pii-privacy-act).

Shapefile: A vector data storage format for storing the location, shape, and attributes of geographic features. A shapefile is stored in a set of related files and contains one feature class.

Raster tile layer: Raster tile layers support visualization of imagery and raster data, but they do not support analysis.

Service credits: Credits are the currency used across ArcGIS and are consumed for specific transactions and types of storage, such as storing features, performing analytics, and using premium content.

Tag: A keyword used to describe a map in ArcGIS Online.

Tile: (layer type) Web tile layers support fast map visualization using a collection of pre-drawn map images, or tiles. They are appropriate for basemaps that give your maps geographic context.

Vector tile: (layer type) Vector tile layers are a collection of vector tiles and style resources that can adapt to any display resolution and be customized for multiple uses. They are appropriate as operational or basemap layers.

Web map: In ArcGIS Online, a web-based, interactive map that allows you to display and query the layers on the map. A web map contains one or more ArcGIS Server map services that are referenced to ArcGIS Online.

*See the GIS Dictionary for more definitions

Overview

Information sources pertaining to multiparty monitoring (MPM) within the Rio Chama CFLRP will be managed according to USDA Forest Service Records Policy and the Federal Data Strategy, including the integrity, completeness, and accuracy of documents, geospatial data, and non-geospatial data. The MPM plan relies on model runs, observations, and document review. All associated data and final reports will be stored on a Box drive (Pinyon) and ArcGIS Online (AGOL) site co-managed by the USDA Forest Service and non-USDA Forest Service partner organizations (Forest Stewards Guild and Mountain Studies Institute). To maintain consistency and clarity, data edits, uploads, and downloads should go through these entities.

Data Collection

Data will be collected through multiple modeling programs, field assessments, socio-economic surveys, and downloaded from existing databases. Specific collection standards are outlined in the protocols for each monitoring methodology (Appendix B).

Data Storage

Data storage will require coordination between USDA Forest Service personnel and non-USFS Partners. Some data (such as USDA Forest Service corporate data systems) will be housed on the USDA Forest Service internal T-Drive and some data (such as wood industry surveys) will be housed on the Forest Stewards Guild box drive. When possible[1], the data from these two storage systems will be uploaded to a central, shared storage site (Figure 1) – either a USDA Forest Service led Pinyon Box drive and/or a USDA Forest Service led ArcGIS Online (AGOL) site. The Pinyon Box drive will store the majority of project data, including "final versions", raster based spatial data, and data requiring minimal to no edits. AGOL will house frequently edited vector data and/or data for Web Maps to promote data exploration.

Pinyon Box Drive
The Pinyon Box drive will serve as the primary space for data storage and sharing. To limit unintentional data alterations, data should not be used or edited directly from Box. Box should serve as a data library to store up-to-date data (and past versions) for users to download.

ArcGIS Online (AGOL)
AGOL should not be used for general data storage. AGOL should be used for frequently edited data, data requiring geospatial reference locations, or data being used by a communication tool such as a Web Map. When possible, avoid storing shapefiles on AGOL and create Geodatabases and Feature Services instead.

---

[1] Not all USDA Forest Service data can be shared in raw form and USDA Forest Service employees will assist with creating reports and/or downloading shareable versions. In addition, wood industry and contractor surveys contain proprietary information and will only be reported and/or shared in aggregate (to protect PII).
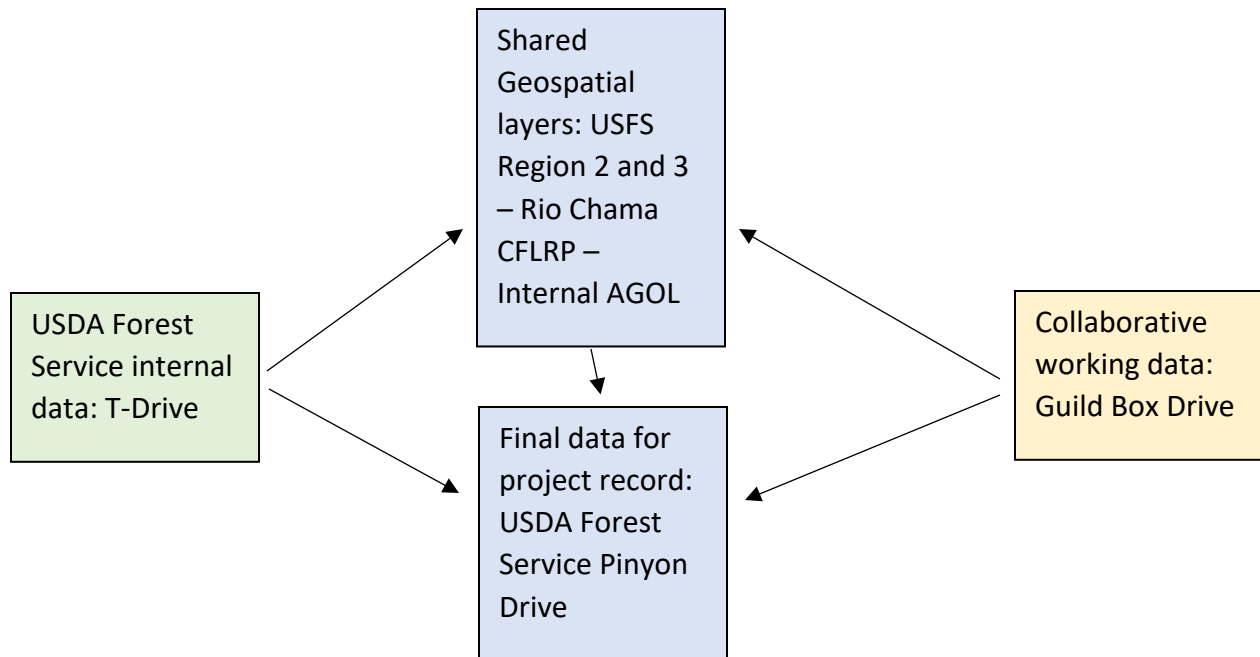
Figure 1. Overview of data storage locations and workflow.

Metadata

AGOL data management will meet the metadata requirements of U.S. Forest Service Handbook 6609.15 and the Federal Geographic Data Committee. In particular, all data should include the following:

- **Summary:** a brief statement describing the purpose and intended use of the dataset.
- **Description:** a detailed description of the dataset, including what the resource is about, why it was created, and who created it. If the data is actively being edited and updated, the description should explain how and when the data will be updated. If the dataset is a snapshot or selection from a larger dataset, the description should provide the date when it was created or copied and the parameters used to select the data.
- **Terms Of Use:** Any special restrictions, disclaimers, terms and conditions, or limitations on using the item's content.
- **Categories:** categories and subcategories used to organize datasets by topic. Each dataset can be assigned one or more category. Categories can be customized within an AGOL group (see table 1).
- **Tags:** Keywords specific to a dataset, used to boost search results and assist in finding items of interest (see Table 2).
- **Credits:** Acknowledgement of the datasets source.

Data Types

This section will be built out once specific protocols are finalized and piloted (Appendix B).

Quality Assurance and Control

This section will be built out once specific protocols are finalized and piloted (Appendix B).

Field Collection

Field data collection will follow set protocols outlined in appendix B. Field Crews and Volunteers will be trained to the appropriate level for a given set of protocols and standardized data collection forms will be used.

Monitoring coordinators from Forest Stewards Guild and Mountain Studies Institute will conduct "hot" (evaluator watching crew collect data) and "cold" (secondary plot measures conducted for comparison) checks of plot data collection as time and resources allow.

Data Upload

Data will be uploaded by USDA Forest Service CFLRP GIS Manager and/or Forest Stewards Guild and Mountain Studies Institute staff following the guidance of this document. All field-based/paper collected data should be uploaded to the Pinyon Box Drive within 7-days of returning to the office (or WiFi enhanced setting). When possible, data sheets and written notes should be input in computer database by a different person than the field recorder to identify potential errors. All hand written notes/data sheets shall be scanned and saved on the Pinyon Box Drive.

Data Backup

Data will be backed-up using a variety of means depending on the data source and type. All data should be stored in at least two ways/locations. Figure 1 outlines how data will have one location on the shared project AGOL or Pinyon Drive and one on the USDA Forest Service T-Drive/Guild Box Drive. Plot data will be input into excel spreadsheets and original data sheets will be scanned and hard copies stored.

Data Review

Annually, a Rio Chama CFLRP data management working group will review the shared Pinyon Box Drive and AGOL sites to remove unnecessary or redundant data, reorganize folders as needed, and mitigate potential data management hiccups.

Data Access

The USDA Forest Service has established an external shared Pinyon Box Drive for Rio Chama shared documents. The agency will manage the permissions of personnel and partner access to the site.

Naming Conventions

1. Separate data by origin (USDA Forest Service, External, Rio Chama CFLRP[2]) with subfolders separated by collection source (e.g., external_data, field_collection, model_outputs, document_review), year, and with additional subfolders as needed.

2. Every folder/subfolder should have an archive folder titled **0_Archive**. Everything but the current version of the dataset goes in the archive folder and a date (**_YYYYMMDD**) should be added at the end of file name.

3. Avoid acronyms and abbreviations in folder, file, and field names (unless very commonly understood).

4. External data shall be kept in original format with original file names. All extracted or filtered datasets or databases should be appended with the filter/extraction parameters (e.g., HUC12**_CFLRP_YYYYMMDD**)

5. Standardized naming conventions will be used for all data originating from the Rio Chama CFLRP (**RC_Category_Subcategory_Name_Status_YYYYMMDD**).

- **RC:** Data may move through individual accounts and this will help users keep their own data organized.
- **Category:** Should match a category outlined in Table 1.
- **Subcategory:** Should match a subcategory outlined in Table 1.
- **Name:** Short and clear description of item (e.g., TrailGulchPlots)
- **YYYYMMDD:** Date of data collection, retrieval, or edit.

Table 1. Overview of data categories, subcategories, and additional information to include in data naming conventions.

Category	Sub-category	Items/additional categories
Basemaps	Hydrography	Watershed boundaries, rivers and streams
	Land ownership	
Boundaries	Administrative	USFS (admin boundaries, ranger districts, regions)
		States and counties
		Tribal
	Collaboratives	2-3-2, SJHFHP, SJCWP
	Initiatives	CFLRP, RGWF
Infrastructure	Transportation	Roads
	Structures	WUI
	Utilities	Power lines, water infrastructure
Projects	USFS	Completed projects: one category for each FY
	All lands	Create one folder for each fiscal year inside All lands
Socioeconomic	TBD	
Ecological	Vegetation	
	Wildlife	
	Water	
	Fire	
PROMOTe	TBD	Rasterized versions of data

---

[2] Rio Chama CFLRP folder contains data created and managed by and for the CFLRP.

ArcGIS Online

In addition to using the naming convention outlined above, data on the AGOL Group site will be tagged for added efficiency (Table 2). Tags provide further description of an item and can be added at anytime. Use only relevant tags to describe an item, and use tags to query searches (link to ArcGIS Blog about using tags effectively).

Table 2. List of tags to be used for AGOL data to provide additional information and ease searching the group site.

Tag Type	Specific Tag
Year:	2022 2023 2024 2025 2026 2027 2028 2029 2030 2031
Managed by:	San Juan National Forest, SJNF Carson National Forest, CANF Santa Fe National Forest, SFNF Rio Grande National Forest, RGNF Bureau of Land Management, BLM Jicarilla Apache Nation, JAN Santa Clara Pueblo, SCP Ohkay Owingeh, OO Southern Ute Indian Tribe, SUIT State of New Mexico, SNM State of Colorado, SCO Private, PRI
State:	Colorado, CO New Mexico, NM
Core restoration treatment type:	Hazardous fuels WUI, FP-FUELS-WUI Hazardous, FP-FUELS-NON-WUI Prescribed fire, FP-FUELS-RX-FIRE Wildlife habitat restoration, HBT-ENH-TERR Stream enhancement, HBT-ENH-STRM Stream crossings, STRM-CROS-MITG-STD Soil improvement, S&W-RSRC-IMP Timber sales, TMBR-SALES-TRT-AC Timber volume, TMBER-VOL-SLD Biomass for bioenergy, BIO-NRG Invasive species, INVPLT-NXWD-FED-AC Forest vegetation improvement, FOR-VEG-IMP Rangeland vegetation improvement, RG-VEG-IMP Road improvement, RD-IMP Trail improvement, TL-IMP

There are two ways to upload data to AGOL:

1. Within AGOL, drag and drop a zipped folder to My Content, or browse for the file on your device.
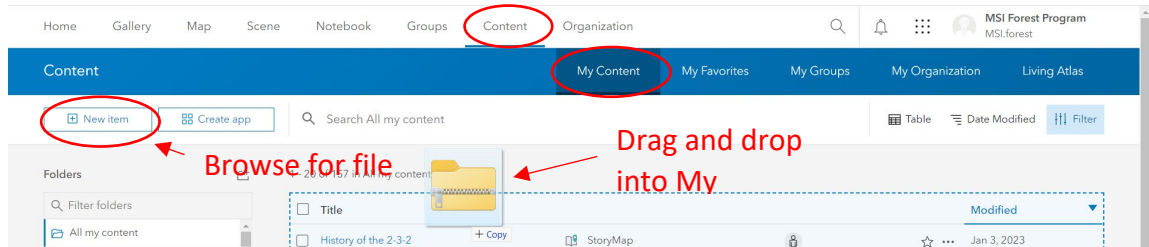
   a. The file must be zipped and contain shapefile(s). **For best practice, each folder should only include one shapefile.** Uploading multiple shapefiles in a single folder will combine them into a single item within AGOL.
      i. **Raster data must be uploaded as a hosted imagery layer or a raster tile layer.** Caution: hosted imagery layers use a lot of credits to upload and maintain.
   b. After identifying the file to upload, identify the file type (likely a Shapefile or Geodatabase) and select **Add _file name_ and create a hosted feature layer.**
   c. This will create a hosted feature layer (for displaying on web maps) and a Shapefile (to download from AGOL).
   d. For more information about adding items to AGOL, view these step-by-step instructions.

2. Publish from ArcGIS Pro
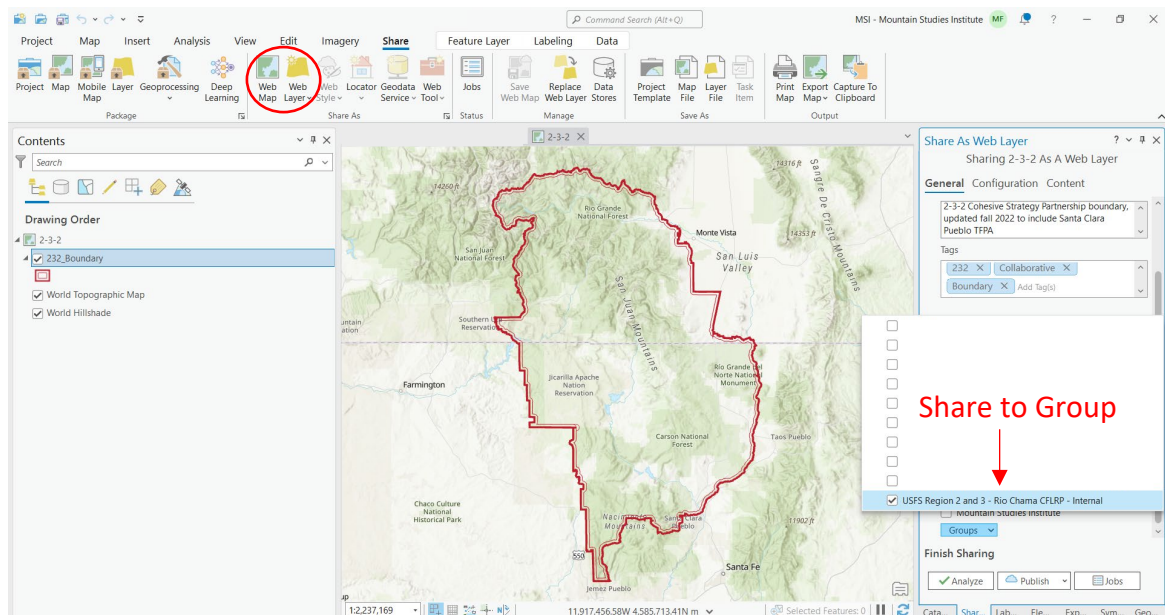
   a. Click on an item in the contents pane, and then choose **publish as web layer**.
   b. Choose a relevant name, provide a description, add tags, and select the layer type.
      i. **Feature:** web feature layers support vector feature querying, visualization, and editing. They are most appropriate for visualizing data on top of basemaps.

        ii.   **Tile:** web tile layers support fast map visualization using a collection of predrawn map images, or tiles. They are appropriate for basemaps that give your maps geographic context.

        iii.   **Vector tile:** vector tile layers are a collection of vector tiles and style resources that can adapt to any display resolution and be customized for multiple uses. They are appropriate as operational or basemap layers.

   c.   **Share to the USFS Region 2 and 3 – Rio Chama CFLRP – Internal Group.**

   d.   *For more information about publishing a web layer from ArcGIS Pro, view these [step-by-step instructions](#).*

Sharing levels

All data shared with the Group will be visible to members of the Group. However, sharing permissions will need to be updated if data is going to be shared outside of the Group. For example, each data layer will need to shared with the Public if you are making a story map or web map that is publicly available. There are three levels of sharing permissions:

Owner: the owner of the item(s) has access

**Organization:** all members of your organization have access (i.e., MSI, Guild or USFS)

Public: people outside your organization have access

Downloading data

There are two ways to access or download data that is stored on AGOL:

1. Within the Group, select an item to download. Click download to save it to your device.
2. Access data from AGOL on ArcGIS Pro.
   a. In the Catalog pane, select Portal. Access data from the Group by selecting the following button:
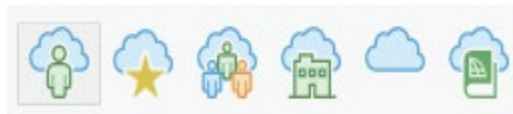
Groups

Data Sharing

Final reports will be shared on the 2-3-2 Partnership website (https://232partnership.org/) with an option to request raw data. Data requests will be reviewed by the 2-3-2 Partnership Executive Committee. Decisions will be made in accordance with USDA Forest Service policy surrounding data restrictions. To maintain consistency and clarity, data edits, uploads, and downloads should go through the USDA Forest Service Rio Chama CFLRP GIS Manager and/or the non-USDA Forest Service Partners active in data management (Forest Stewards Guild and Mountain Studies Institute).

Sensitive Data

Sensitive data will be collected through the surveys used in socioeconomic monitoring. This sensitive data includes Personally Identifiable Information and will only be reported in aggregate. These data sets will be maintained by the Forest Stewards Guild and will not be stored on the USDA Forest Service led AGOL or Pinyon Box Drive.

Sensitive data pertaining to specific plot locations will be handled on a case by case basis through conversations with the 2-3-2 Partnership Monitoring and Executive Committees.

Data Schema

This section will be built out once specific protocols are finalized and piloted (Appendix B).

A conceptual data schema for the 2-3-2 Partnership's Multiparty Monitoring Plan for the Rio Chama CFRLP is presented in Figure 2.

Logical Data Schema Here.

A logical database schema describes the constraints applied to the data and defines fields, tables, relations, views, etc. The rules or constraints that are defined in this logical model help determine how the data in different tables relate to each other.

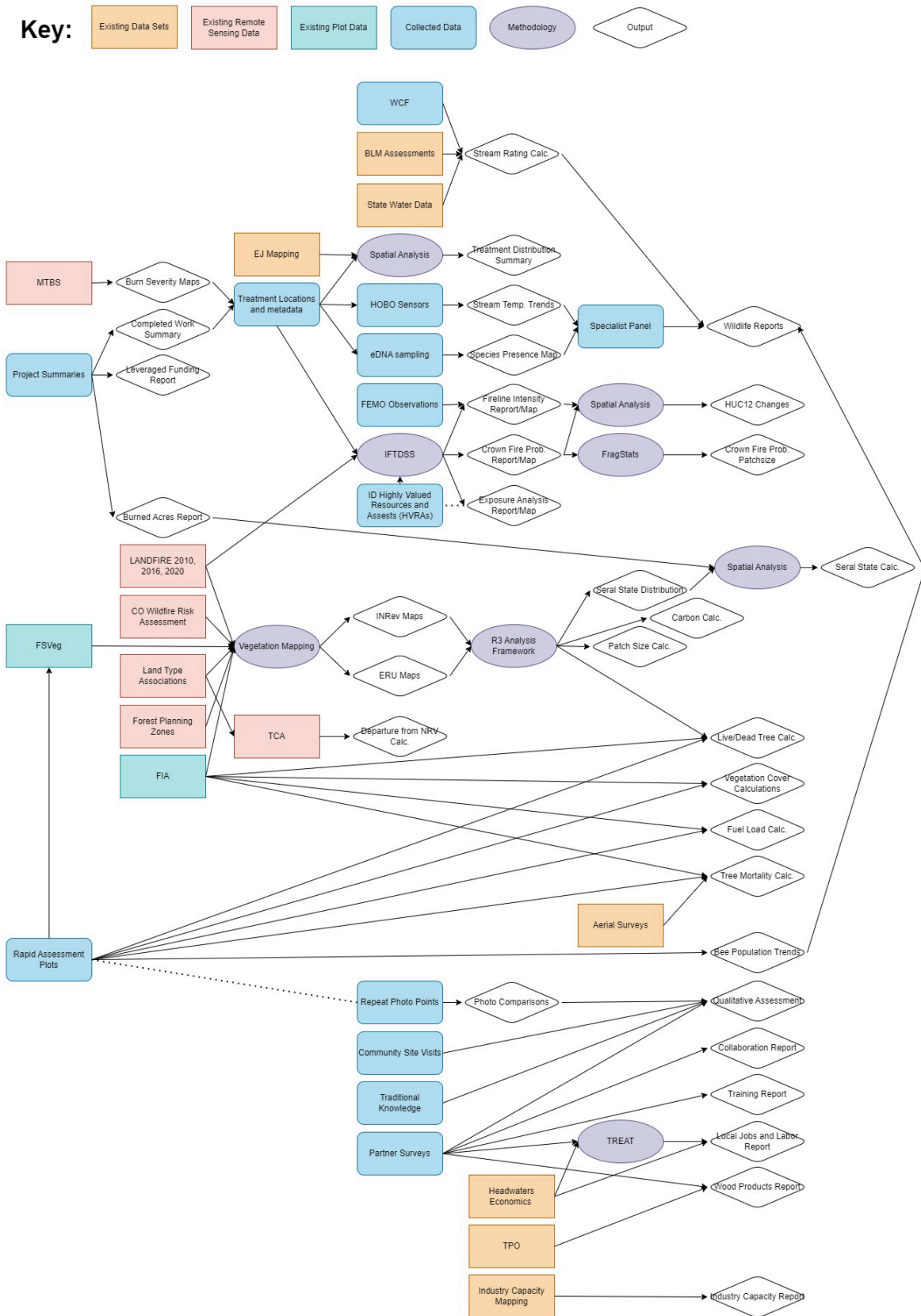Regional Data Dictionaries are being reviewed to inform how data in different formats relate to each other.

Figure 2. Conceptual data schema for the 2-3-2 Partnership's Rio Chama CFLRP multiparty monitoring plan.

Physical Data Schema Here.

A physical database schema describes how data will be stored and the form of storage used (files, key-value pairs, indices, etc.).

Regional Data Dictionaries are being reviewed to inform how data in different formats relate to each other.

Data Preservation

Raw and final data will be stored in project record.