Goals for Data Organization, Reporting, and Visualization

# Purpose of the Project

To create a centralized, well-structured database that organizes, cleans, and makes accessible our water quality, volunteer, and equipment data, while enabling both internal reporting and public-facing visualizations.

In the first phase of this project, the priority is to compile, clean and flag the existing StreamWatch data and create a functional database in which to store everything.

# Scope of Data

## Water Quality Data

* Data from three StreamWatch teams:
  + Chemical Action Team
  + Biological Action Team
  + Bacterial Action Team
* Historical data spanning 30+ years
* Multiple parameters (pH, dissolved oxygen, E. coli, macroinvertebrate counts, etc.)
* Associated site information (location, sampling frequency, monitoring notes)
* Associated methods metadata (type, detection limits, etc.)

## Volunteer Data

* Contact information and participation history
* Team assignments and sampling site associations
* Training completion records
* Availability and active/inactive status
* Partner/Parent associations

## Equipment Data

* Meter inventory and details (make, model, serial number)
* Calibration records and maintenance logs
* Test results linked to specific meters

# Primary Goals for the Database

## Organize & Manage Data

* Centralized storage for water quality, volunteer, and equipment datasets
  + Map existing data subsets & plan/document migration
* Consistent data formatting and standardized field names
* Processes for cleaning and validating new entries
* Workflows for importing & exporting additional files
  + Survey123
  + Gallery CSVs
  + ArcGIS Dashboards
* Flagging data above or below DEP thresholds

## Facilitate Reporting & Analysis

* Ability to filter, search, and download datasets by date, location, parameter, or other variables
* Export in WQX-compatible format for required submissions
* Tools for trend analysis (seasonal, annual, long-term)
* Auto-flag data issues (meter failed test, volunteer due for training, etc.)

## Enhance Data Visualization & Shareability

* Time-series graphs for parameters over different periods
* Maps showing site locations and parameter values
* Public-facing dashboards and reports
* Potential for individual site pages with photos, historical trends, and downloadable data

# Additional Considerations

* Data backup and version control processes
* Accessibility for both staff and public users
  + User permissions for editing vs. viewing data
* Potential integration with existing tools (ArcGIS, StoryMaps, Excel, Power BI, etc.)