Water Quality Exchange Sharing Biological Data

Michael Brennan, IT Team member and ORISE Fellow, U.S. EPA's Monitoring Branch

Dwane Young, Senior Data Advisor, IT Team, U.S. EPA's Monitoring Branch

WQx
WATER QUALITY EXCHANGE

Presentation overview

- WQX Overview

- WQX data submittal

- Retrieving data from the STORET warehouse

# NARS Survey Data submitted to STORET using WQX

The Water Quality Exchange

- WQX defines the framework by which EPA compiles water quality monitoring data (physical, chemical and biological) that are collected by a number of entities via a shared schema.
- Data submitted via WQX is placed in the STORET Data Warehouse

WQX
WATER QUALITY EXCHANGE

Why Share Data Using WQX

- WQX is currently used by
  - 5 federal agencies
  - 175 State Territorial and Local Agencies
  - 90 Tribal Agencies
  - 75 Public Groups
- Data submitted through WQX is stored in EPA's STORET (STOrage and Retrieval) National Data Warehouse.
- STORET is EPA's repository for over 130 million bits of physical, chemical, biological and habitat data.
- STORET partners with US Geological Survey to host the Water Quality Portal, which is essentially an interagency marketplace for monitoring data.
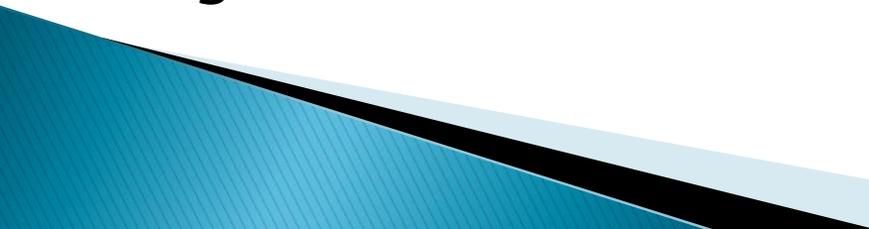
WQX Data Flow
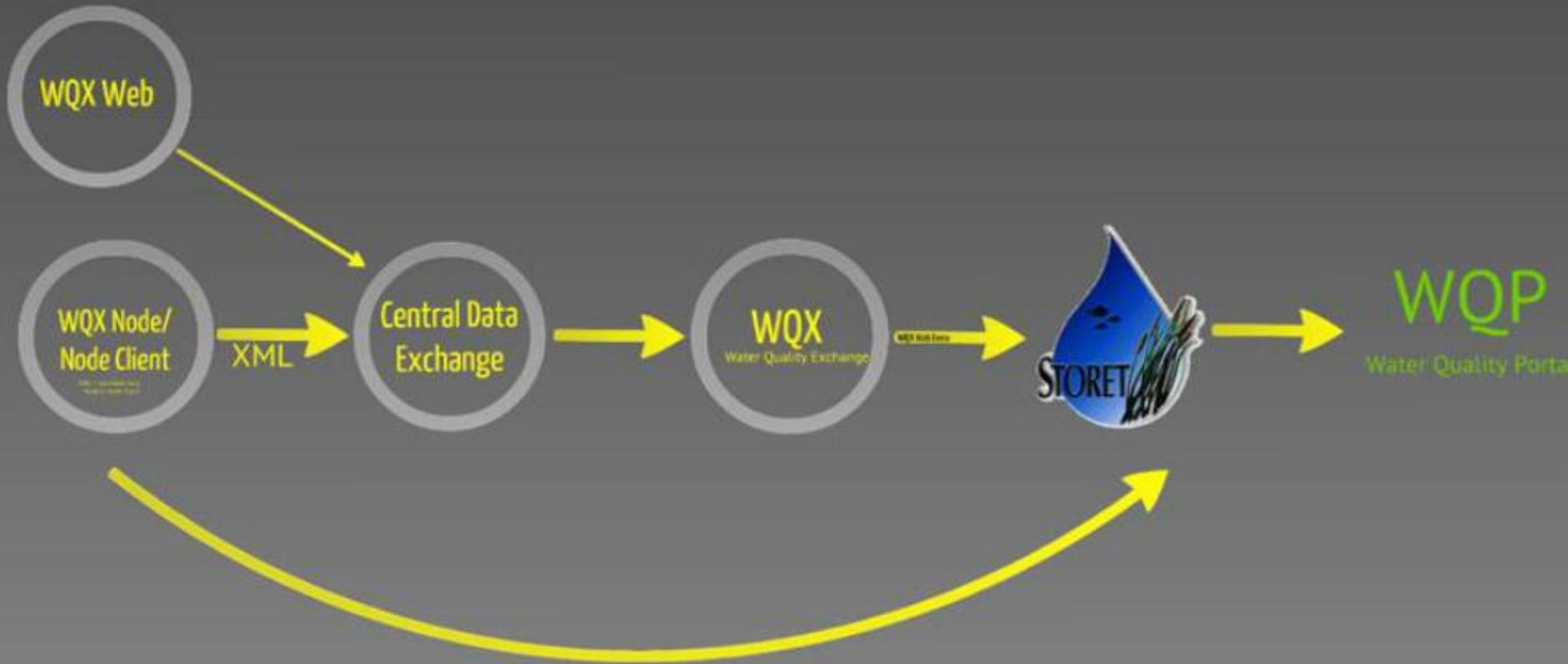
WQX Data Flow
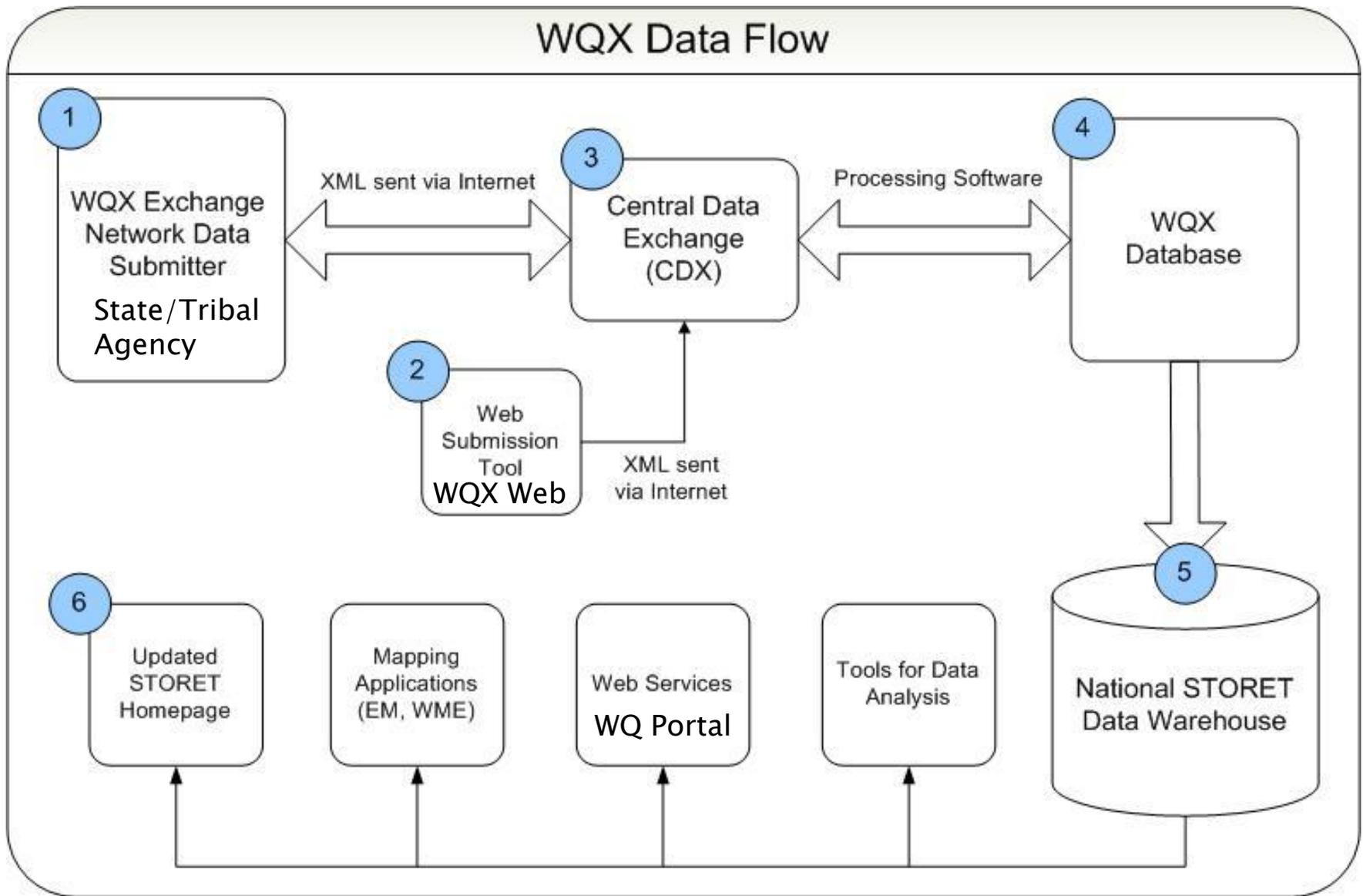
Essentials of WQX Data Content

- What do you have?
- Water Monitoring data consists of:
  - Who collected it
  - What was collected
  - Where sample/ measurement was collected
  - When it was collected
  - Why it was collected
  - How it was collected

- It does not take a lot of information to create a good data set.

WQX Schema

- Common set of data elements, rules and standard names that everyone uses when submitting data
- The schema is organized as follows:
  - **Organizations** (whose data this is)
  - **Projects** (why the data was collected)
  - **Monitoring Locations** (where the data was collected)
  - **Activities/Results** (what the data is, how is was analyzed, etc)

How data is organized in WQX
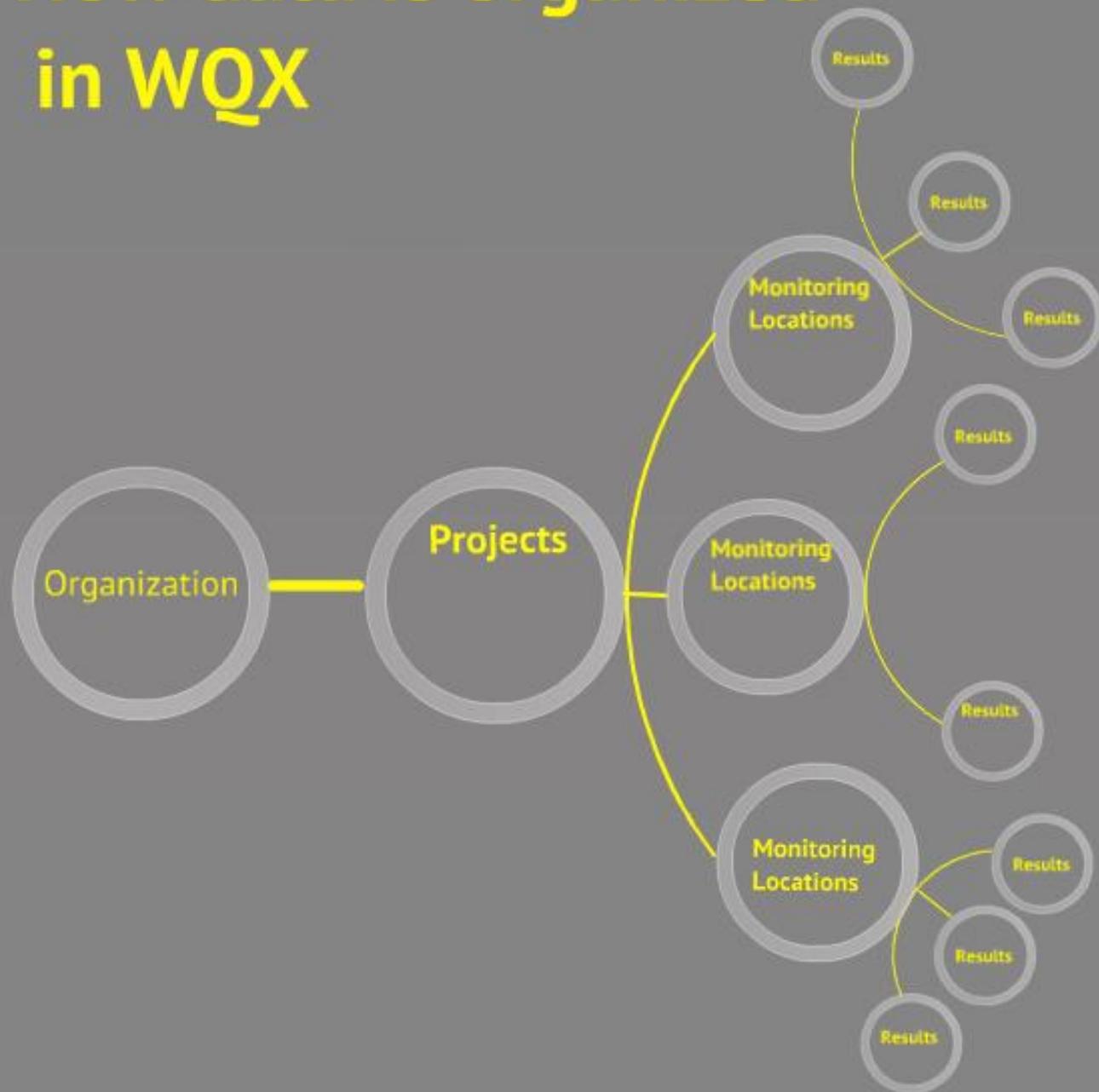
WQX schema contents

- Physical conditions at the time of a site visit
- Water chemistry and bacteriological data
- Fish tissue data
- Biological Taxon Abundance data, including population census, frequency class, group summaries, and individual results
- Reference site information
- Toxicity data
- Habitat Assessment scores and their related metric scores
- Biological Index scores and their related metric scores

The WQX Web Biological Template

About STORET/WQX
Data download
Online Tutorials
Data Submittal
Useful Internet Links
Support
Tools/Web Services
Frequent Questions
Helpdesk
Sitemap

WQX Web Biological Template (ZIP 1.1MB)

The Biological Template is provided to assist in formatting biological results. The template is made up of two Microsoft Excel spreadsheet files that are meant to be used together to assist with data tracking and entry. The WQX Web Template Dictionary file provides guidance on the appropriate use of each data element, highlighting the data elements that are available in the WQX Web Template and showing the additional data elements that a user may add to the template. The dictionary covers data elements for all templates including Physical–Chemical Results, Biological Results, Habitat Results, Activity Metrics and Indices, and Continuous Monitor Results. *Note this template does not include the data elements for submitting biological metrics or indexes; these are handled using a separate WQX Web template file.

This template contains sample data that can be imported into WQX Web using the following import configuration (for more information regarding how to use Import Configuration files in WQX Web, please see the WQX Web tutorials):

- IMPORT BIOLOGICAL RESULTS

•Online tutorials

•Biological Template

•Biological Template Dictionary

http://www.epa.gov/storet/wqx/wqxweb_downloads.html

WQX Web Biological Template

USEPA WQXWeb Biological Template
Version 1.04

Domain values last updated: 03/02/2012 10:38:00 AM

• This template is a data entry spreadsheet that guides data owners through organizing water quality data into a format that meets WQX data validation requirements.

• This template is intended to be paired with the WQXWeb Import Configuration - Import Habitat Results.bin. Changes to the order of columns or the data format in this WQX Web template spreadsheet also need to be applied to the WQXWeb Import Configuration.

• Please refer to the latest version of the "WQXWeb Template Dictionary" for a detailed explanation of the contents within each data entry worksheet, in addition to a complete list of WQX Allowable Values. The dictionary also cotains a list of all the columns available in each Data Entry worksheet.

Worksheets are color coded by function. The single pink tab contains buttons used to export data, the three yellow tabs are used to enter data, and the green tabs are reference lists for data columns that allow only specific values.

Group Name	Use	Worksheet Name	Description
Export	Use buttons on this tab convert Data Entry Worksheets (yellow tabs) to .txt files	Export	The Export tab contains buttons to automatically export data from each of the data entry worksheets into tab delimited files ready to be imported into WQXWeb.
Data Entry Worksheets	A template for submission of water quality monitoring data. Projects, Monitoring Locations and Results templates are provided for users	Projects	The Project tab contains information about the water quality data collection program
		Monitoring Locations	The Monitoring Locations tab contains information about the sites were water quality data is being collected
		Results	The Results tab contains the field and laboratory water quality data collected.
Allowable Values/ Look-up Lists/ Domain Values	Tables of allowable values for specific columns in the Data Entry worksheets. All green-colored cells are the values that can be used in the worksheets. Others cells are included for additional reference.	Allowed Values - Monitoring Locs	This tab contains multiple tables of listing the values that can be entered in particular columns in the Monitoring Locations tab.
		Allowed Values - Results	This tab contains multiple tables of listing the values that can be entered in particular columns in the Results tab.
		Characteristics	This tab contains a table of all Characteristics in STORET that can be used in the Characteristic Name field in the Results tab. The table also has fields to indicate if a particular Characteristic requires a Sample Fraction or Field/Lab Analytical Procedure (or both) for a particular characteristic.
		Subject Taxon	Contains the list of taxons used in STORET.
		Analytical Methods	This tab contains a list of all nationally available result analytical methods. Additional methods can be defined by an organization in the "Analytical & Collection Methods" tab.
		Units of Measure	This tab contains a single table listing all result units of measure available in WQX.
		Analytical & Collection Methods	This tab can be used to record organization specific Result Analytical Methods and Sample Collection Methods. Data entered in this tab is not exported to WQXWeb.

For assistance with using this template, please refer to the US EPA STORET/WQX online resources at http://www.epa.gov/storet/

The most recent copy of this template and corresponding dictionary can be downloaded from http://www.epa.gov/storet/wqx/wqxweb_downloads.html

If you have questions or comments about this template, please send email to the STORET Help Desk at STORET@epa.gov

Instructions / Export / Projects / Monitoring Locations / Results / Allowed Values - Monitoring Loc / Allowed Values - Results / Charac

Mapping Biological Data collected in the Field to WQX

Name of Organism   Size of Organism   Organism Anomalies

| Line # | Common Name | Hy-brid? | Intro-duced? | Tally and Counts | | | | Anom Count | Mortality Count | VOUCHER | | | Voucher Tag # | Vouchers Retained | Fla |
				<150 mm	150-300 mm	300-460 mm	>460 mm			Unk/Rng	QA	Photo			
										○	○	○			
1		○	○												
										○	○	○			

NRSA 20__ __SH COLLECTION

Site ID: _____   Date: ___ / ___ / ___   Reviewed by (initial) _____   Page ___ of ___

Characterizing Biological data

Biological Data

Activity Set

Biological Intent

Assemblage Sampled Name

WQX Activities

‣ Results and Metrics must relate to an Activity

‣ The Activity provides information about when, where, and how the biological samples were collected, as well as the Assemblage Sampled.

‣ Each Activity must include an Activity ID, which is completely unique within your organization.
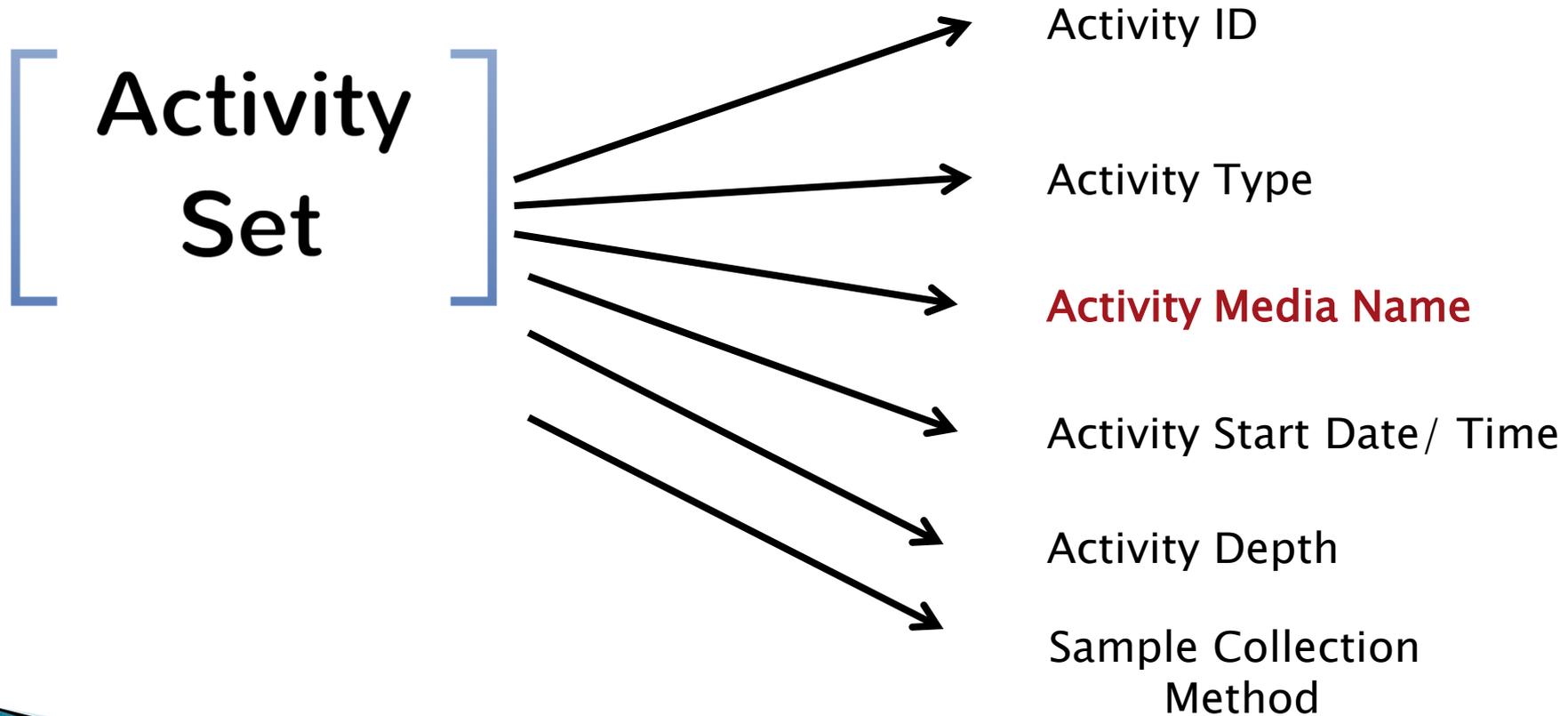
　◦ e.g. "BEARRV23-2013/05/04-FISH2"

Riffle-Kick method

Rock Scrub

Biological Results
Activity Set

Activity Set

→ Activity ID

→ Activity Type

→ Activity Media Name

→ Activity Start Date/ Time

→ Activity Depth

→ Sample Collection Method

Activity Examples

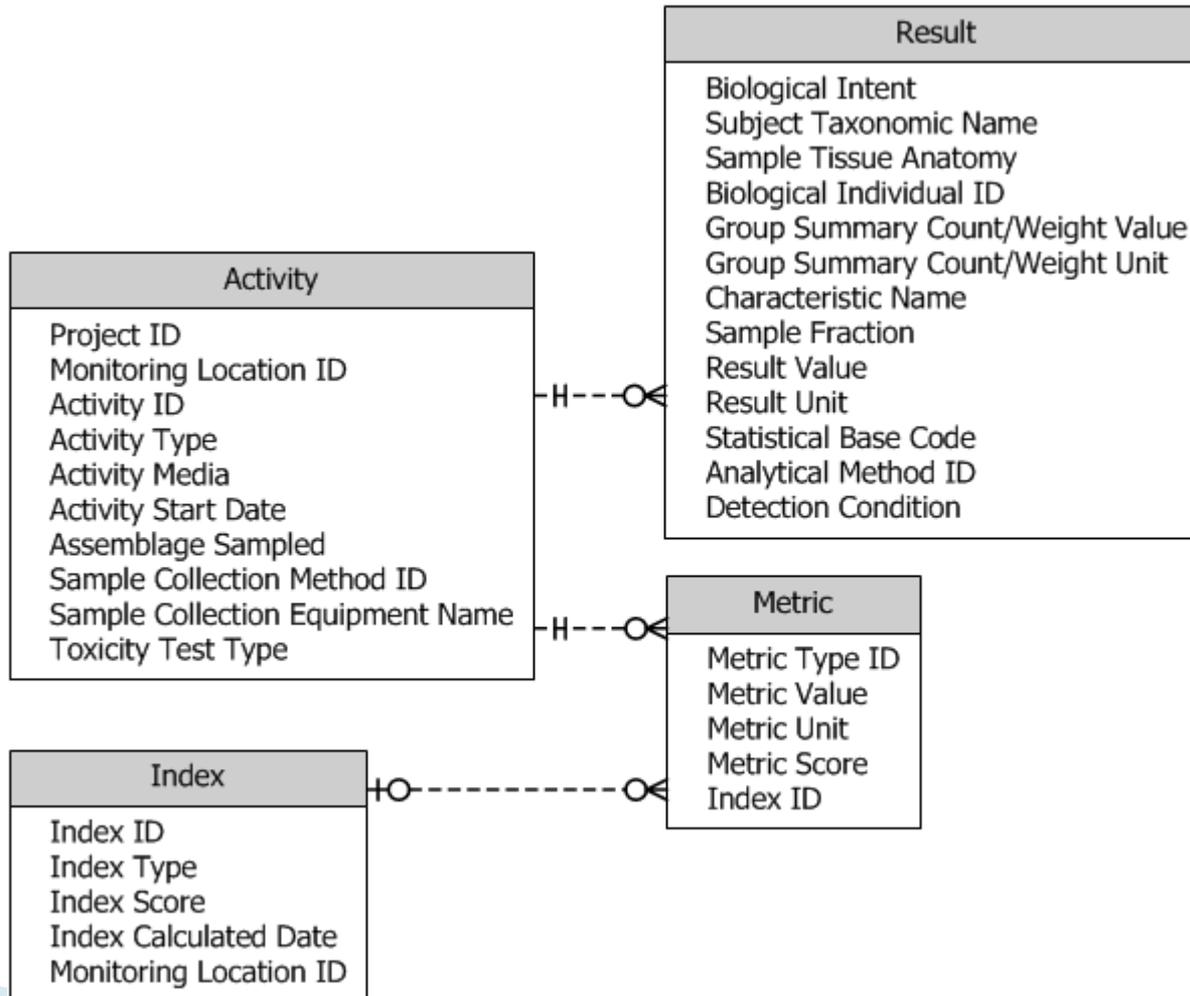|	Activity 1	Activity 2	Activity 3
Project ID	TRIB-2013	LAKES-2013	LAKES-2013
Monitoring Location ID	BEARRV23	BEARLK17	BEARLK12
Activity ID	BEARRV23-2013/05/04-FISH2	BEARLK17-2013/05/04-PZ	BEARLK17-2013/05/04-
Activity Type	Sample-Routine	Sample-Integrated Vertical Profile	Sample-Composite Without Parents
Media	Tissue	Biological	Biological
Activity Start Date	5/4/2013	5/4/2013	5/4/2013
Assemblage Sampled	Fish/Nekton	Phytoplankton/Zooplankton	Benthic Macroinvertebrates
Sample Collection Method ID	GRE:Fish	GRE:Diatoms	GRE:Sediment
Sample Collection Equipment	Backpack Electroshock	Pump/Bailer	Petite Ponar Grab
Toxicity Test Type			Acute

Biological Data in WQX

- There are three entities (i.e. types of records) that can hold biological data:

  - **Results:**
    - The measurements, counts, or lab results relating to a particular species

  - **Metrics:**
    - The individual values and scores that make up a biological assessment (for a specific monitoring location)

  - **Indexes:**
    - The overall biological assessment score (for a specific monitoring location)

Data Flow

Results → [Used To Calculate] → Metrics → [Used To Calculate] → Index

- In general terms…
  - Results are used to calculate Metric Values and Scores
  - Metric Scores are used to calculate an overall Index Score.

- There are relationships between all of these entities in WQX (so they can be linked to each other).

# How These Entities are Related

Result

Biological Intent
Subject Taxonomic Name
Sample Tissue Anatomy
Biological Individual ID
Group Summary Count/Weight Value
Group Summary Count/Weight Unit
Characteristic Name
Sample Fraction
Result Value
Result Unit
Statistical Base Code
Analytical Method ID
Detection Condition

Activity

Project ID
Monitoring Location ID
Activity ID
Activity Type
Activity Media
Activity Start Date
Assemblage Sampled
Sample Collection Method ID
Sample Collection Equipment Name
Toxicity Test Type

Metric

Metric Type ID
Metric Value
Metric Unit
Metric Score
Index ID

Index

Index ID
Index Type
Index Score
Index Calculated Date
Monitoring Location ID

Results

▸ You indicate than an Activity contains biological Results by setting the Activity Media to "Biological" or "Tissue".

▸ Then you provide a Biological Intent (on each Result) to indicate the type of biological result being provided.

▸ When the Activity Media is "Tissue" then the Biological Intent must also be "Tissue"

Biological Results
Biological Intent –Domain Values
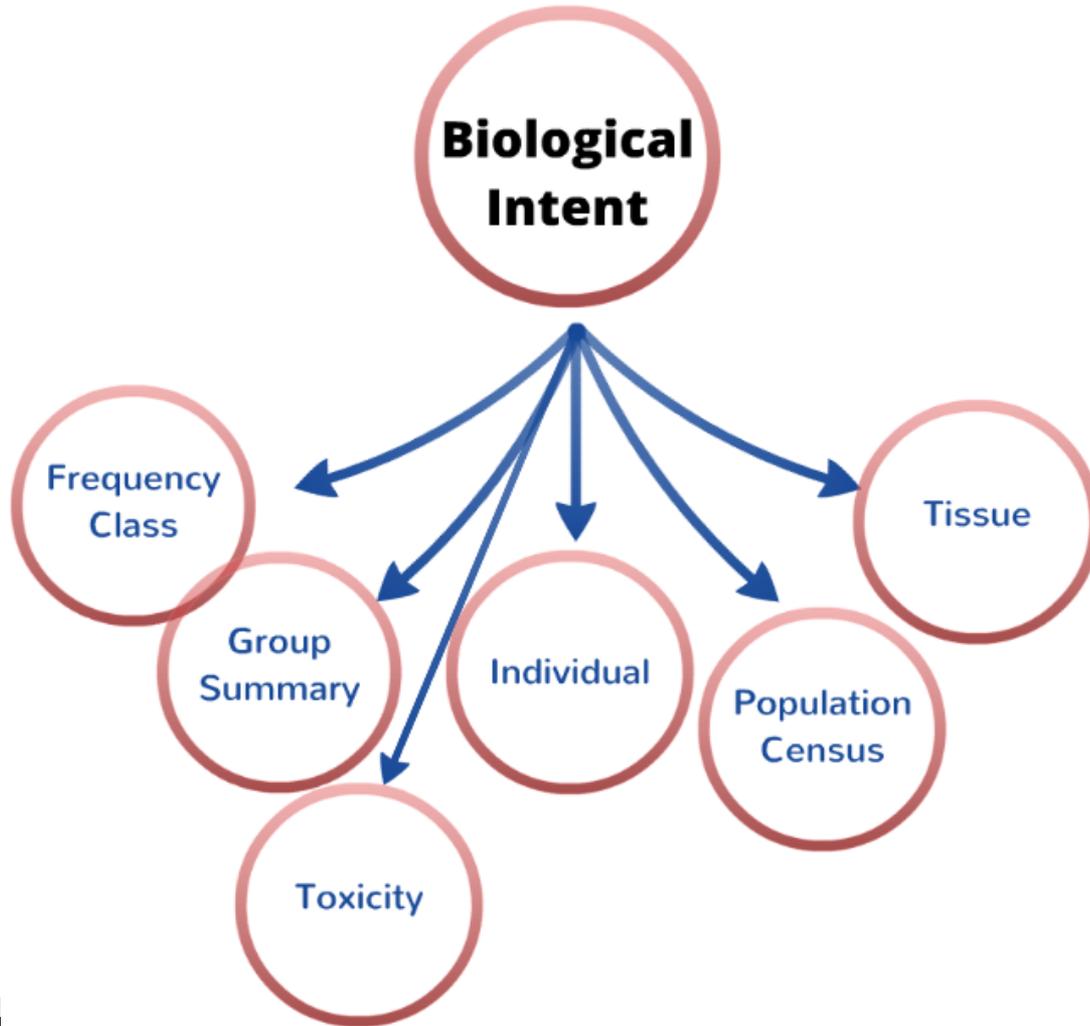
Biological Intent

- "Population Census"
  - for providing the "Count" or "Total Sample Weight" of a particular species found in a sample or collection event.

Population Census Examples

	Result 1	Result 2
Biological Intent	Population Census	Population Census
Taxonomic Name	Salmo trutta	Moxostoma erythrurum
Characteristic	Count	Total Sample Weight
Result Value	231	9728
Result Unit	count	g
Value Type	Actual	Actual

Frequency Class Examples
Measured Characteristics

	Result 1	Result 2	Result 3
Biological Intent	Frequency Class	Frequency Class	Frequency Class
Taxonomic Name	Catostomus commersoni	Catostomus commersoni	Catostomus commersoni
Characteristic	Count	Count	Count
Result Value	7	3	1
Result Unit	count	count	count
Value Type	Actual	Actual	Actual
Frequency Class Descriptor	Length	Length	Length
Frequency Class Descriptor			
Frequency Class Lower Bound	0	20	40
Frequency Class Upper Bound	20	40	60
Frequency Class Unit	cm	cm	cm

Other Biological Intents Supported by WQX

- Individual – For reporting measurements from a single individual. A Biological Individual ID (e.g. #1, #2, …) is provided with each Result to identify which individual.

- Group Summary – For providing summary information (e.g. mean, max) about a particular species that was collected.

- Toxicity –For reporting the results of toxicity testing on a species.

- Tissue – For reporting the results of toxicity testing on a species.
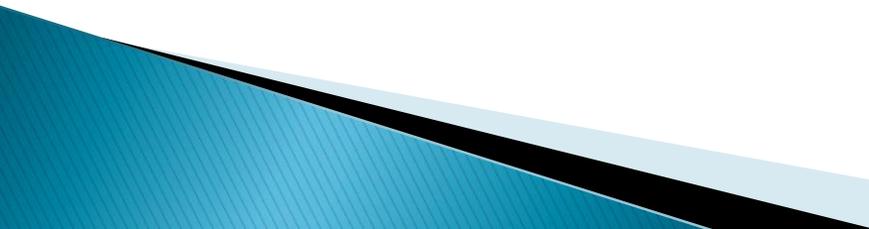
Metrics

- A Metric must be linked to an Activity.

  ◦ The Activity can be the same one that was used for the Biological Results, or it can be a new one (with its own Activity ID)

- A Metric may be linked to an Index

  ◦ If you are using a product like WQX Web, this means your Indexes must be loaded before your Metrics

- A Metric has a "Score" (which is required) and a "Value" (which is optional).

Metric Examples

	Metric 1	Metric 2
Metric Type ID*	2000-DE-%Clingers	Total Taxa Richness
Metric Type Context	21DELAWQ_WQX	USEPA
Metric Score	6	4.482758
Metric Value	59	26
Metric Value Unit	%	count
Index ID		BEARRV23-2013/05/04-FISH2

* Metric Type meta-data can also include citations and documentation on the formula and scale (e.g. 1-100) for the score

Indexes

- An Index is the parent of many Metrics
  - Although, technically, it can stand alone.
- An Index does not relate to an Activity
  - so the "when, where, and how" is part of the Index record
- An Index has a "Score", but does not have a "Value".

Index Examples

	Index 1	Index 2
Monitoring Location ID	VERDE_RIVER_164	BEARRV23
Index Calculated Date	06/05/2001	05/04/2013
Index ID	30841	BEARRV23-2013/05/04-FISH2
Index Type ID*	IBI	Stream Fish Index
Index Type Context	21ARIZ_WQX	IDEQ_WQX
Index Score	75.0369517	82.65292

* Index Type meta-data can also include citations and documentation on the scale (e.g. 1-100) for the score

Retrieving biological data from the STORET Warehouse

Download Biological Results from the STORET Warehouse

Regular

Biological

Habitat

Metric

Index

www.epa.gov/storet/dw_home.html

Interpreting Biological Data from STORET
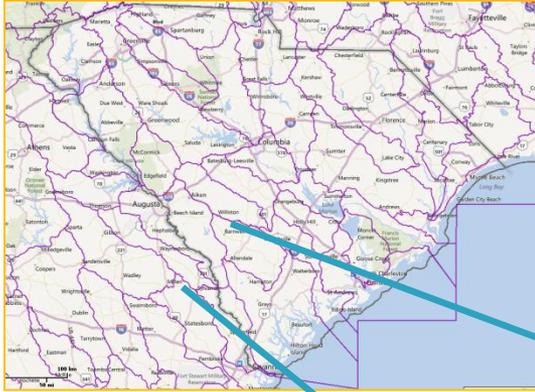
- Geographical reference data
  - HUC, Lat/Long, State County

- Date/ Time of collection

- Sampling collection method
  - Riffle-kick method, rock scrub, electroshocking, seining.

Riffle-Kick method

Rock Scrub

WQX Resources

Watershed Academy Webcasts

October 23, 2012 Webcast Seminar:
"Using the New Water Quality Portal"

http://water.epa.gov/learn/training/wacademy/archives.cfm#w2012102:

March 13, 2013 Webcast Seminar:
"Water Quality Exchange: A Tool for Tribes, Volunteer Monitors and Others to Share WQ Data"

http://water.epa.gov/learn/training/wacademy/archives.cfm#w2012102:

Contact information

▸ STORET WQX Reference Information
  ◦ http://www.epa.gov/storet/wqx/index.html

▸ STORET Helpdesk
  ◦ STORET@EPA.GOV
  ◦ 1-800-424-9067