



Image: Amy Puls



COLUMBIA RIVER BASIN
RESTORATION PROGRAM



Columbia River Basin Restoration Program Toxics Monitoring Subgroup Meeting

FEBRUARY 11, 2025 | 1:00 – 1:30 PM PACIFIC

AGENDA

5 MIN

WELCOME & INTRODUCTIONS

60 MIN

DATA MANAGEMENT

- *What is Data Management, and Why is it Important?* **Edyth Hermosillo** (Aquatic Biologist, EPA Region 10, Seattle)
- *Environmental Data Management at WA Ecology – 25 Years of Insight and Tips,* **Chris Neumiller** (EIM Data Manager, WA Ecology)

20 MIN

ROUND ROBIN

- Meeting participants are invited to share 1-2 minute updates about new monitoring on the horizon.

5 MIN

TMS UPDATES

- News about TMS Core Team tasks and the 2025 TMS Workshop

MS TEAMS TIPS

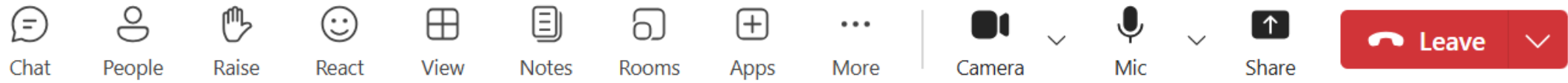
Please turn camera and mic off when not speaking.



Click to turn
video camera
off/on

Click to turn
microphone
off/on

INTRODUCTIONS



Click chat icon and
introduce yourself –
name and affiliation

The background of the slide features a blurred graphic with binary code (0s and 1s) in blue and green, overlaid with orange and red line charts and bar graphs, suggesting a data management or analytics theme.

DATA MANAGEMENT

What is Data Management, and Why is it Important?

- **Edyth Hermosillo** (Aquatic Biologist, EPA Region 10, Seattle)

Environmental Data Management at WA Ecology – 25 Years of Insight and Tips

- **Chris Neumiller** (EIM Data Manager, WA Ecology)



What is data management, and why is it important?

*Edyth Hermosillo
Aquatic Biologist
EPA Region 10, Seattle*



Data Management

- **What is Data Management?**
 - The process of storing, organizing, maintaining and using data created, collected, and/or generated by an organization.
- **What is the goal of data management?**
 - To help organizations optimize the use of data to make informed decisions to maximize the benefit of the organization.

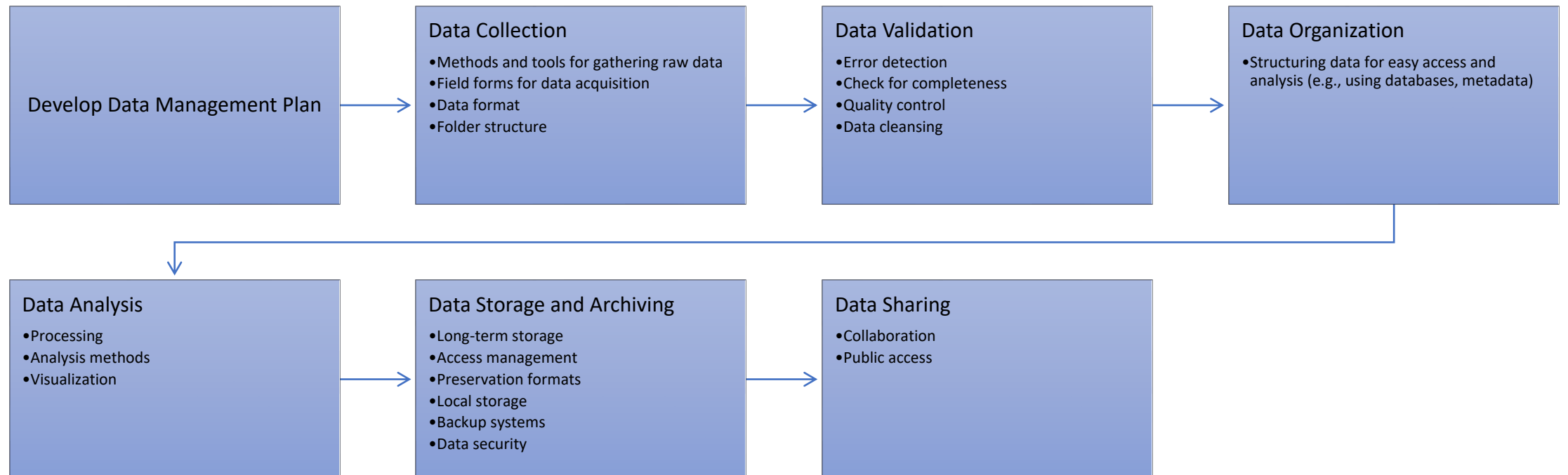


Why is Data Management Important?

Good Data Management:

- Ensures data quality and accuracy
- Facilitates collaborative research and reproducibility
- Promotes data sharing and public access
- Enables compliance with funding agencies, institutions, and legal regulations
- Improves transparency

Data Management Components Flow Chart





Data Management Plans (DMP)

- **What is a DMP?**
 - A document that outlines how data will be managed during and after a research project.
- **Key Elements of a DMP:**
 - Data collection, storage, sharing, preservation, and access.
 - Tools and resources used for managing data.
 - Responsible personnel and timelines.
- **Who implements Data Management Plans?**
 - Federal Government ([USGS](#), [NIH](#), [NIEHS](#), [NOAA](#), [NSF](#), etc)
 - State Government Agencies
 - Universities ([DMPTool](#))
 - EPA policy requires that EPA-funded work have an approved QA Project Plan (QAPP), and data management is one of the elements listed in *EPA requirements for QAPPs (QA/R-5)*.



Data Collection

- **Types of Data:**
 - Laboratory: Environmental results, instrument calibrations
 - Field: Experimental data, survey data, observational data, equipment calibrations
 - Modeling: statistical analysis, predictive simulations
- **Collection Methods:**
 - Laboratory instruments
 - Field sampling and monitoring equipment
 - Computational simulations (such as statistical and predictive models)
- **Best Practices:**
 - Standardized protocols for all data collection methods.
 - Ensure proper documentation of methods and conditions.

Standardized Protocols in Scientific Data Collection

(Bonus Slide
not presented
during live
presentation)

The Importance of Standardization

- **Consistency:** Provides a reliable baseline for data collection
- **Comparability:** Enables meaningful cross-project analysis
- **Quality Control:** Reduces variability and potential errors
- **Efficiency:** Streamlines training and research processes

However, Maintaining Flexibility is Important

- Standardized forms and protocols are starting points, not final boundaries
- Modifications of standard protocols may be needed to fit specific project needs

Balancing Standardization with Flexibility

1. Start with a Standard Template

- Use existing, well-developed protocols as a foundation
- Leverage proven methodologies and best practices

2. Customize as Necessary While Maintaining Core Scientific Integrity

- Allow for project-specific modifications
- Document any deviations from standard protocols
- Justify and explain deviations from standard protocols



Data Validation

Purpose: Ensure data accuracy and quality.

- **Steps:**
 - **Completeness Check:** Ensure no missing values.
 - **Error Detection:** Identify and correct outliers or anomalies.
 - **Data Cleansing:** Remove irrelevant data or duplicates.
 - **Periodic Review:** Reassess data for its relevance and accuracy.
 - **Updates:** Revise data if new information or corrections are made.
- **Tools:** Excel, Python, R.
- **Best Practices:** Maintain version control and track changes.



Data Organization

- **Structure:**
 - Organize data into folders, databases, or spreadsheets.
 - Use consistent naming conventions for files and variables
- **Metadata:**
 - Document the context of data (units, measurement methods).
 - Include timestamps, versioning, and experimental conditions.
- **Best Practices:**
 - Follow standardized formats like CSV, JPEG, or XML.



Data Analysis

- **Processing:** Clean and prepare data for analysis (e.g., normalization, aggregation).
- **Modeling:** Statistical models, machine learning algorithms.
- **Visualization:** Use of graphs and charts to interpret the data.
- **Tools:** Python, R, Excel. Statistical analysis
- Data visualization: Qlik, Tableau, PowerBI
- Machine learning integration platforms

Data Storage and Archiving



Storage Options:

Short-term Storage:

- Local: Hard drives, internal servers.
- Cloud: Google Drive, Dropbox, specialized repositories.

Long-term Storage and Archiving:

- Institutional repositories
- Public data archives
- Specialized research databases



Security:

Data encryption.

Access control measures (passwords, permissions).

Recovery options.



Best Practices:

Use trusted archives that support data preservation standards.

Implement regular backup schedule to prevent data loss

Use proper metadata documentation to ensure metadata is preserved for future use or audits



Data Sharing

- **Collaboration:** Share data with colleagues via cloud storage, repositories (e.g., WQX*), or secure networks.
- **Public Access:** Publish datasets on open-access platforms (e.g., WQP) and provide citation options for proper attribution
- **Best Practices:**
 - Provide clear documentation and metadata to make data understandable.
 - Follow ethical guidelines for data sharing

*EPA grantees collecting water quality data in accordance with an EPA approved QAPP must upload associated data to WQX.

Water Quality Exchange (WQX)



WQX is EPA's repository for data partners to submit water monitoring data to EPA. A standardized data format and submission database provided for sharing water quality monitoring data of various types.



WQX defines a common data model for communicating water quality data (sample data)



Designed to be automated from your existing dataset



The structure of partner data systems don't matter, so long as they can map data to WQX standards



Many ways to prepare and submit data to WQX: including direct submissions, WQXWeb, and 3rd party apps.

Best Practices in Scientific Data Management

- **Data Documentation:**
 - Use of metadata standards and file formats.
 - Clear and complete descriptions of data and methodologies.
- **Data Versioning:**
 - Tracking changes and updates to datasets.
- **Data Sharing Policies:**
 - Open data principles vs. restricted access.
- **Data Backup:**
 - Regular backups to ensure data recovery.
- **Ethical Considerations:**
 - Respecting privacy, consent, and ensure compliance with ethical and legal guidelines.

Conclusion

- Use proper tools and platforms to manage data throughout its lifecycle.
- Effective scientific data management ensures quality, integrity, and accessibility of scientific data.
- Adoption of best practices, tools, and standards is critical for successful data management.
- Overcoming challenges in SDM will enable better research outcomes and innovations.

WQX Basic Resources

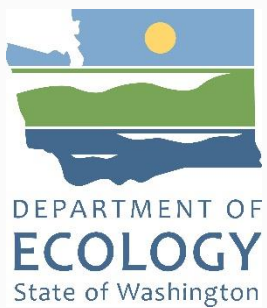
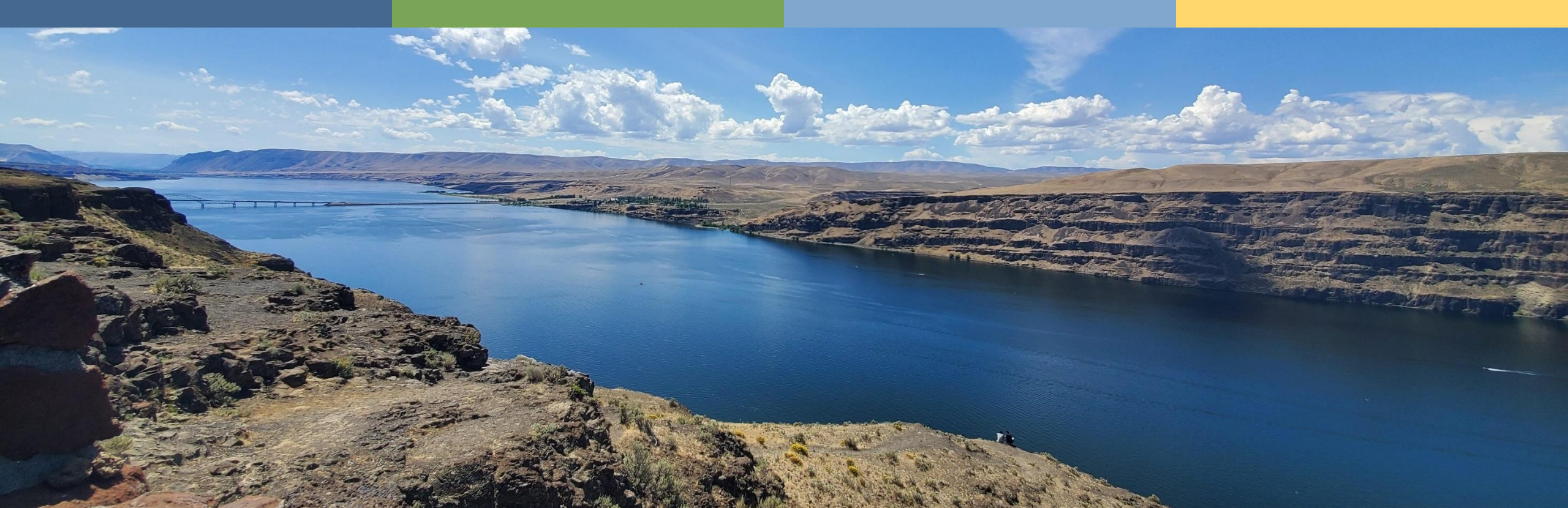
Topic Area	Resource Links
Open Water Data Resources / Links	Water Quality Exchange (WQX) [Data In] The Water Quality Portal (WQP) [Data Out] How's My Waterway (HMW) [Info Out] Central Data Exchange (CDX) [Data Held] Exchange Network (Node Submissions to CDX)
WQX Nuts and Bolts	Upload Resources Page Glossary of Terms Business Rules Data Exchange Template (List of Available Elements) Domain Service (Acceptable Values)
Using WQX – Getting Started	Introduction to WQX, WQX Web, and WQP (5 min video) Quick WQX Web User Guide (PDF Guide) User Guide Version 3.0 for Water Quality Exchange Web Getting Started with WQX Web: How to Gain Access WQX Web Basics – Two-part session--Day 1 WQX Web Basics – Two-part session--Day 2
Using WQX – Templates	Water Quality Exchange Web Template User Guide US EPA Link to Web Templates Web Templates Overview Video
Using WQX – Import Configurations	WQX Web Import Configuration Options Translations via Expert Mode
Using WQX – Best Practices	Best Practices for Sharing Benthics Data WQX Metals Best Practices Guide WQX Nutrients Best Practices Guide
Using WQX - Other	Water Quality eXchange Factsheet Common Errors Resolution WQX Program Information

Other Resources

- [EPA Requirements for Quality Assurance Project Plans, EPA QA/R5: March 2001](#)
- [Guidance for Quality Assurance Project Plans, EPA QA/G-5: December 2002](#)
- [Filename Structure Best Practice](#)



Questions?



Environmental Data Management at WA Ecology - Insights and Tips

Chris Neumiller, EIM Data Manager

February 11, 2025

Agenda

- 1** EIM Overview
- 2** Data Management
- 3** EIM to EPA WQX Data Flow
- 4** Low-Level PCB Congeners

EIM Overview



What is EIM?

- Ecology's main system for environmental monitoring data
- Launched 1997
 - Based in part on EPA STORET and USGS NWIS
- Data types
 - Mostly soil, sediment, surface and groundwater, plant and animal tissue and counts, bioassay, limited air
- Must be a measure of environmental conditions
- End-user, qualified data - no lab QA/QC data*

Now - 2025

Data

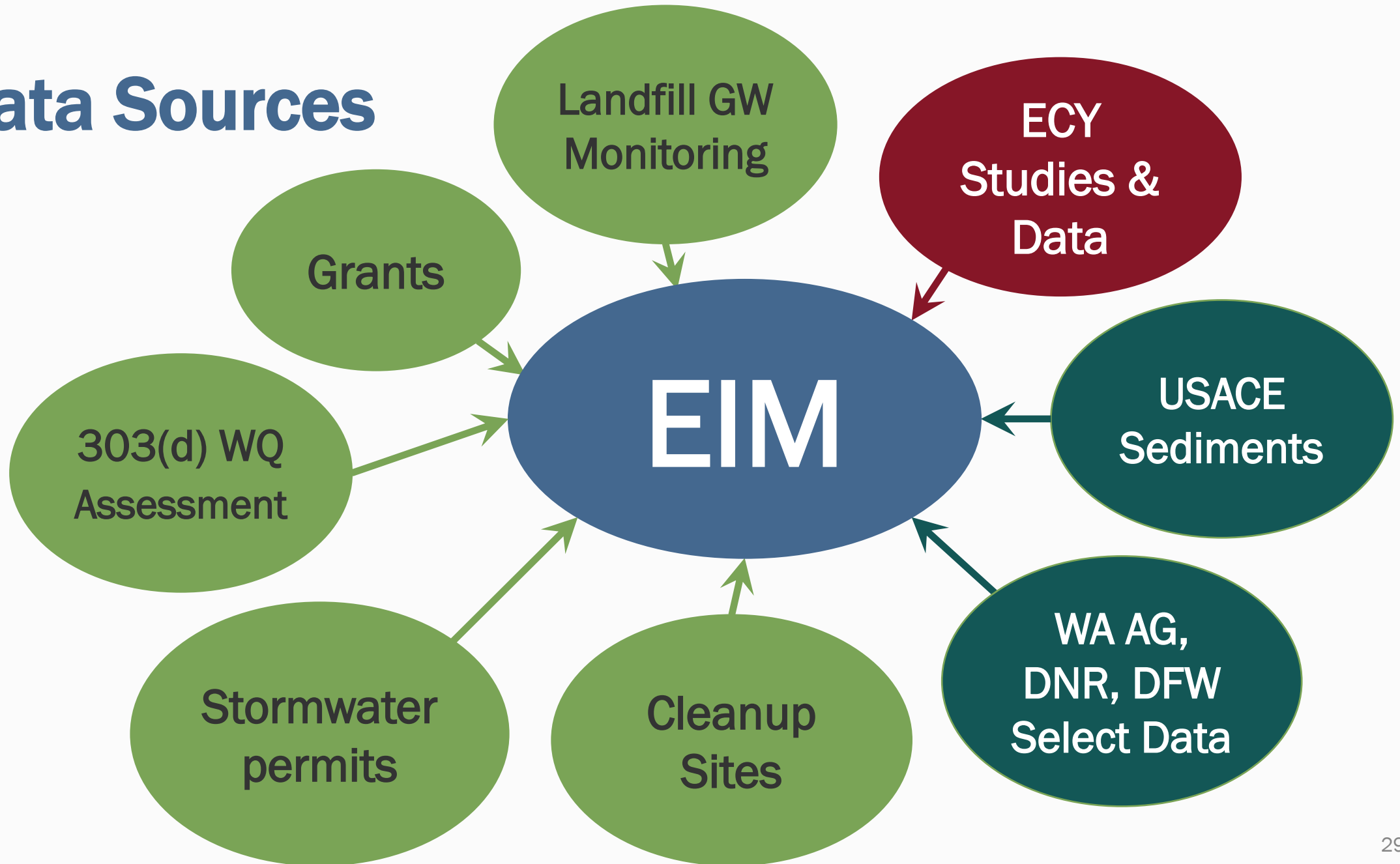
- 239 thousand monitoring locations
- 23 million discrete results
- 198 million time series results
- Auto-calculated time series summaries

Many tools and support

Public access and data sharing

Projects to make it better

Data Sources



Guidelines for Accepting Data

- Needs a predefined reason – Ecology partners
- Must have data coordinator support
- EIM doesn't take data available in other public systems, like USGS NWIS. WQX exception.

EIM Search

- Public tool for data access
- Search multiple ways
- Monitoring Programs
- Download data

Search:

[ALL](#)

Studies + Locations + Results

[Studies](#)

Example: Study ID AODE6815

[Locations](#)

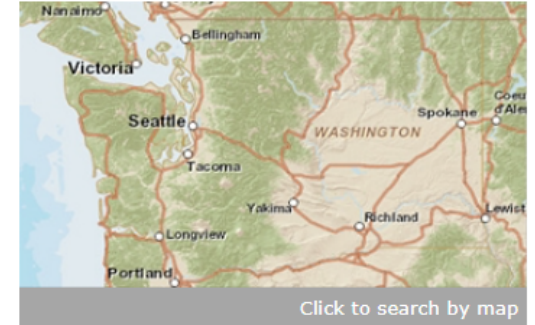
Example: Nooksack River

[Results](#)

Example: Copper

[Bioassay](#)

Example: Neanthes 20 day



Search does not include physical habitat data and metrics - see Watershed Health below.

Search Monitoring Programs within EIM Datasets collected by Ecology and affiliates, with specific monitoring objectives and consistent protocols. Most are long-term and regularly-scheduled. Each Monitoring Program has a custom search form and map to help you find data.

Groundwater

Groundwater quality and level data. Collected by many groups with various objectives and protocols, 1929 to present.

[Get Data](#)  [About groundwater](#)

Freshwater Information Network (FIN)

Monthly river and stream water quality monitoring data and Water Quality Index (WQI), 1960s to present.

[Get Data](#)  [About freshwater information network \(fin\)](#)

Watershed Health Monitoring (WHM)

Physical habitat, biological field, and chemistry data; physical and biological metrics. Includes only EIM Studies with habitat data, 2009 to present.

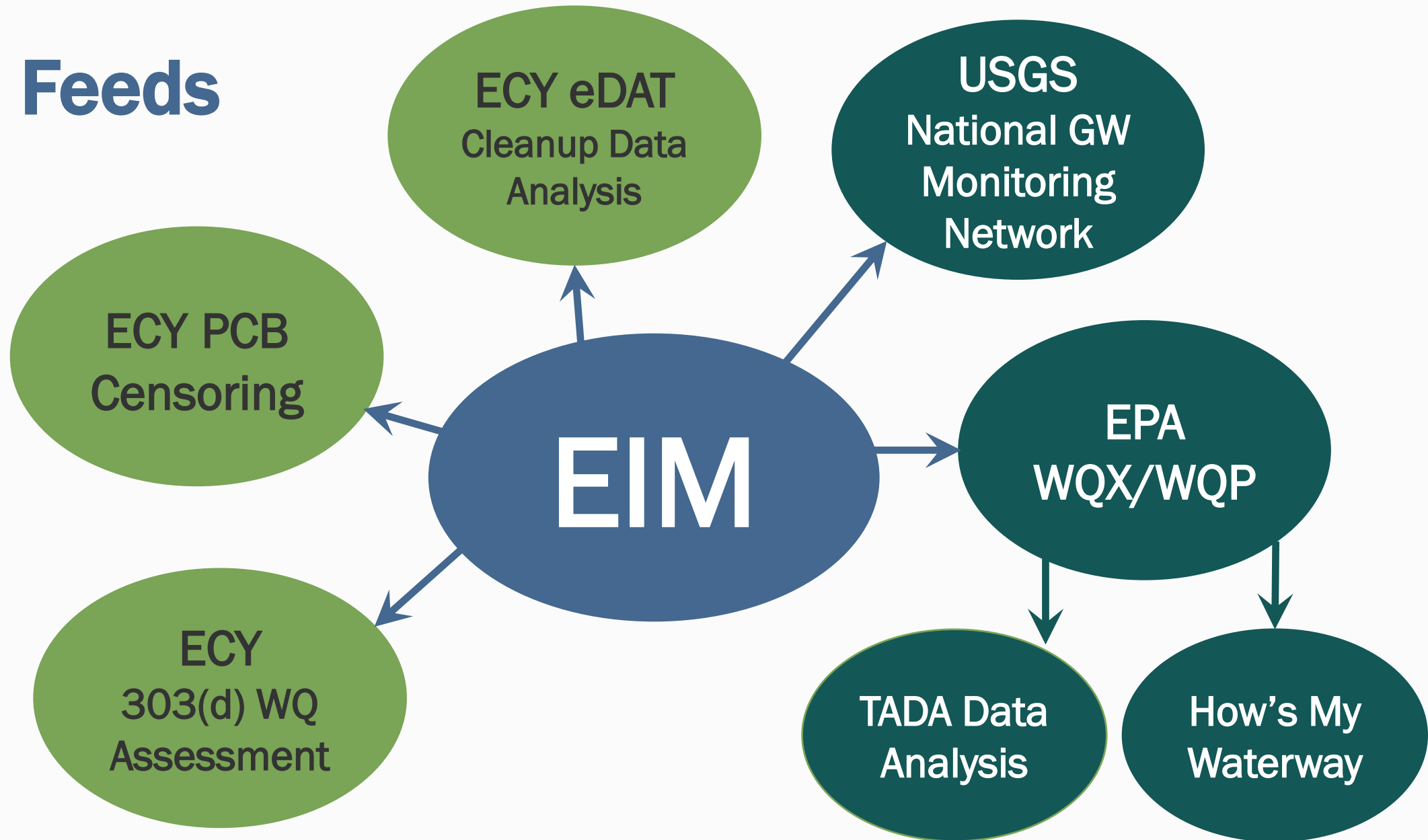
[Get Data](#)  [About watershed health monitoring \(whm\)](#)

BEACH Program

Weekly bacteria data from Puget Sound and coastal beaches, Memorial Day through Labor Day, 2004 to present.

[Get Data](#)  [About BEACH program](#)

Data Feeds



Data Management

Insight and Tips

Data Management is Important!

- Focus on DATA QUALITY
- Standard and consistent fields and metadata
- Standard entry protocols for all data types
- Optimize system and **governance** for **data quality** and shareability
- Reduce data entry errors

Standards and Consistency

- Devise and document naming conventions – files, fields
- Look at established systems
- Use valid values where possible – parameter names, units

Code ↓	Description ↓↑
AXYS-MLA-110-01	Analytical Procedure for the Analysis of Per- and Polyfluoroalkyl Sub MS/MS, Rev 01
AXYS-MLA-110-01-05	Analytical Procedure for the Analysis of Per- and Polyfluoroalkyl Sub MS/MS, Rev 01, Ver 05
AXYS-MLA-110-01-06	Analytical Procedure for the Analysis of Per- and Polyfluoroalkyl Sub MS/MS, Rev 01, Ver 06.
AXYS-MLA-110-01-07	Analytical Procedure for the Analysis of Per- and Polyfluoroalkyl Sub MS/MS, Rev 01, Ver 07.
AXYS-MLA-110-02-08	Analytical Procedure for the Analysis of Per- and Polyfluoroalkyl Sub MS/MS, Rev 02, Ver 08.
AXYS-MLA-110-02-13	Analytical Procedure for the Analysis of Per- and Polyfluoroalkyl Sub Blood/Serums and Solvent Extracts by LC-MS/MS, Rev 02, Ver 13.

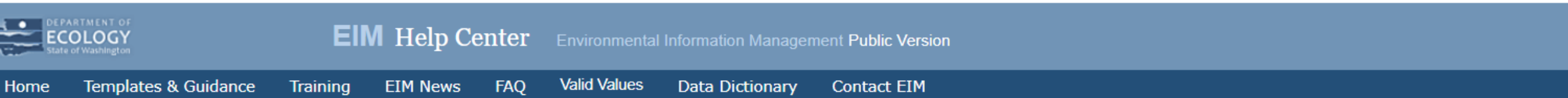
Standardized Data Entry

- Robust documentation
- Field checks for valid values and data types
- Example: [Results Template Help](#)

EIM Help - Results Template, Version 4.4							Page 15
Col	Field Name	Description	Requirements	Type	Size	Valid Values and Conditions	Examples and Guidance
X	Sample Matrix	Describes the general environmental medium which was measured or from which a sample was taken.	Required	Text	14	▪ Air/Gas ▪ Other Liquid ▪ Habitat ▪ Solid/Sediment ▪ Tissue ▪ Water	The matrix for porewater is Solid/Sediment because it's a part of sediment.
Y	Sample Source	Describes the environmental resource which was measured or from which a sample was taken. More specific than Sample Matrix.	Required	Text	30	Must be a valid EIM Sample Source Code. See table of Sample Source valid values (in this document).	▪ Fresh/Surface Water

Organized Documentation

- Assemble your documentation in one, shareable space
- [EIM Help Center](#) example



Templates & Guidance

Data entry documents

Overview

[Data Types We Don't Put Into EIM](#)

[EIM Data Disclaimer and Reliability Statement](#)

[EIM News - Results Template Changes](#)

Changes effective June 29, 2023, including 4 new fields.

[How to Submit Data to EIM \(for public users\)](#)

Steps for making an account and submitting data

[How to Submit Time-Series Data to EIM](#)

Topic-specific guidance

Studies

[How to Link Documents to EIM Studies](#)

Link to electronic publications and documents in Ecology document management systems.

[Naming Conventions for EIM Studies](#)

Locations

[Convert Local Elevation Datums to NAVD88](#)

If you surveyed your wellhead elevations to a local (e.g. 100 ft) datum, convert to our standard datum

Topic-specific guidance, cont.

Groundwater

[NAPL \(Free Product\) Well Data](#)

Water and product levels, physical parameters.

[Spring and Seep Data](#)

[Temporary Environmental Investigation Wells](#)

Data from Direct Push/Geoprobe

[Well Information and Data](#)

How to enter well location information and data for permanent groundwater wells

Help for Users

- EIM Data Coordinators
- Scientists in key roles
- User support and training
- Review and load data
- Help improve EIM
- Cross-agency representation



How we manage our own metadata

- Use authoritative sources for metadata like parameters (chemical names), methods, and taxa
- Cross-check sources

Name	Type	Acroynm	Organization	Website
California Environmental Data Exchange Network	General	CEDEN	Cal EPA	http://ceden.org/CEDEN_Checker/Checker/LookUpLists.php
Water Quality Exchange	General	WQX	USEPA	https://www.epa.gov/waterdata/storage-and-retrieval-and-water-quality-exchange-domain-services-and-downloads
Lab Search	Labs		WA Ecology Lab Accreditation Unit (LAU)	https://fortress.wa.gov/ecy/laboratorysearch/
National Environmental Methods Index	Methods	NEMI	National Water Quality Monitoring Council, USGS, USEPA	https://www.nemi.gov/
SGS-AXYS Analytical Services	Methods		AXYS Analytical Services	https://www.axysanalytical.com/
Standard Methods for the Examination of Water and Wastewater	Methods	Standard Methods	American Public Health Association (APHA), American Water Works Association (AWWA), and Water Environment Federation (WEF)	https://www.standardmethods.org/

Environmental Data Management Resources

- ITRC [EDM Best Practices](#)
- [ICEDM](#) - International Commission on EDM
 - Non-profit, volunteer-based organization, annual meetings
- [EPA WQX](#) white papers on sharing data
 - [Nutrients](#)
 - [Metals](#)
 - [Benthics](#)
- [EIM Help Center](#) valid values guidance

Valid values guidance

Methods, parameters, units, etc

[About EIM Methods](#)

[About EIM Valid Values](#)

Specific values EIM allows in certain fields

EIM to EPA WQX

Data Flow Project

What is WQX?

- EPA's main system for environmental monitoring data
- Formerly known as STORET
- Feeds EPA's [Water Quality Portal](#)
- Anyone can download data or use web services
- Some EPA tools fed by WQP data:
 - [TADA](#) - Tools for Automated Data Analysis
 - [How's My Waterway](#) – provides public with information about the condition of their local waters

WQX Project Background

- 2010 proof of concept data flow from EIM to WQX
- Ecology promised to send federal grant data to WQX
- 2021 Exchange Network Grant to develop robust flow
 - All EIM fields and metadata mapped to WQX 3.0
 - Used WQX Web Service – goes through WQX Web tool
 - Completed August 2024

How it Works for Partner Submitters

- For NEP data, EIM Loader asks if they want data sent to WQX on their behalf
 - We confirm with submitter before sending data
- Send data from an EIM Study to WQX (using our tool)
- Check uploaded data against EIM
- Inform submitter of success

WQX Dashboard (Internal)

WQX Studies

Study Filter Options

Study ID contains

-- Select --

Study Name

Contains

Filter

Clear Search

Delete from WQX

Check Data

Send to WQX

Show 10 entries

	Study ID	Study Name	EPA WQX Organization ID/Name	Locations	Results	Load Start	Load End	WQX Status	
	bcar0003	Manure as Fertilizer Groundwater Study	WA_ECOLOGY/Washington Department of Ecology	8	5092	10/1/2024 7:48:31 AM	10/2/2024 7:03:14 AM	Completed at CDX at Result level	Download

Challenges

- Developing crosswalk logic
 - Mapping fields and values that didn't match up
 - Biologic data – counts and tissue
- Testing
 - Lack of true WQX test environment
 - Help is sparse and hard to find (they are working on it)
 - Figuring out “Inconsistent Data” error messages
- Good thing - WQX staff very responsive and helpful!

Low-Level PCB Congeners

Automated Data Censoring Project

Blank Censoring Results Data

- For **low-level PCB congener data** by EPA1668C
 - Can see down to <1 pg/L
 - Lab blanks can contain up to 20 pg/L
- EIM automatically blank-corrects results data
 - Based on detections in associated lab method blank
- EIM automatically determines if a result should be “censored”
 - Whether considered a non-detect or not

Can Submit Data Two Ways

1. Data that have already been blank censored
2. Data that have not been blank censored

Added 4 New Fields (and Lots of Code!)

QC Blank Type

Method Blank

QC Blank Censor
Factor

Uncensored, 1x, 3x, 5x, 8x, 10x, Factor Unknown

Lab Batch ID

Example: B23C125

Lab GC Column ID

Examples: SPB-OCTYL, DB-1HT

How Blank Censoring Works

EIM auto-calculates results using 1x, 3x, 5x, 8x, and 10x blank censor factors



If Censor Value Threshold is more than Sample Result Value, then...



EIM adds “U” to original qualifier

- J becomes UJ - NJ becomes NUJ

Data Censoring Example

- **Equation:** *Method Blank Result Value x QC Blank Censor Factor = Censor Value Threshold*
- If Censor Value Threshold is more than Sample Result Value, EIM will censor the result by adding a “U” qualifier

Result Parameter Name	Sample Result Value (pg/L)	Method Blank Result Value (pg/L)	QC Blank Censor Factor	Censor Value Threshold (pg/L)	Result Value (final) (pg/L)	Result Data Qualifier
PCB-028	5.0	0.8	3	2.4	5.0	
PCB-028	5.0	0.8	5	4.0	5.0	
PCB-028	5.0	0.8	10	8.0	5.0	U

Download Example

- PCB Auto-Censored Calculations
- Detects become non-detects based on
 - Blank contamination
 - Different blank censor factors
- Censor Value Threshold



		Un-censored				x1				x3		x5		x5		x8		x10		x10	
Result	Result	Result	Result	Blank		Accept		Result		Result		Accept		Result		Accept		Result		Result	
Parameter	Value	Units	Qualifier	Status	Status	1x	Or	Data	3x	3x	Accept	Data	5x	Or	Data	8x	Or	Data	10x	Or	Data
Name	Value	Units	Qualifier	Status	Status	Blank	Reject	Qualifier	Blank	Or	Reject	Qualifier	Blank	Reject	Qualifier	Blank	Reject	Qualifier	Blank	Reject	Qualifier
PCB-017	0.517	pg/L	UJ	Non-detect	Detect	2.15	Reject	UJ	6.45	Reject	UJ	10.75	Reject	UJ	17.2	Reject	UJ	21.5	Reject	UJ	
PCB-023	0.522	pg/L	UJ	Non-detect	Non-detect	0.5	Accept	UJ	1.5	Accept	UJ	2.5	Accept	UJ	4	Accept	UJ	5	Accept	UJ	
PCB-037	2.28	pg/L	J	Detect	Detect	1.89	Accept	J	5.67	Reject	UJ	9.45	Reject	UJ	15.12	Reject	UJ	18.9	Reject	UJ	
PCB-083/099	19	pg/L	J	Detect	Detect	2.69	Accept	J	8.07	Accept	J	13.45	Accept	J	21.52	Reject	UJ	26.9	Reject	UJ	
PCB-044/047/06	12.2	pg/L	J	Detect	Detect	9.7	Accept	J	29.1	Reject	UJ	48.5	Reject	UJ	77.6	Reject	UJ	97	Reject	UJ	
PCB-093/095/09	4.03	pg/L	NJ	Detect	Detect	4.26	Reject	NUJ	12.78	Reject	NUJ	21.3	Reject	NUJ	34.08	Reject	NUJ	42.6	Reject	NUJ	
PCB-040/041/07	1.46	pg/L	NJ	Detect	Detect	1.98	Reject	NUJ	5.94	Reject	NUJ	9.9	Reject	NUJ	15.84	Reject	NUJ	19.8	Reject	NUJ	

How to See the Data

- EIM Search download files
 1. Discrete Results - PCB sample data censored at default of 5x*
 2. PCB Auto-Censored Calculations - Censored results
 3. PCB Blanks - original method blank data

	Category	Record Count
☒	Studies	(1)
☒	Locations	(4)
☐	Well Locations	
☐	Well Measuring Point	
1 ☒	Discrete Results	(3816)
☐	Bioassay Results	
☐	Time-Series Results	
☐	Summarized Time-Series Results	
☐	Water Column Profile Results	
☐	Summarized Water Column Profile Results	
2 ☒	PCB Auto-censored Calculations	(3816)
3 ☒	PCB Blanks	(477)

Queue Download

Help in EIM Help Center

- Templates & Guidance > Topic-specific guidance > Lab and field result data > [Low-Level PCB Congener Data](#)





Thank you! Questions?

Chris Neumiller

ELM Data Manager

Chris.Neumiller@ecy.wa.gov

ROUND ROBIN

- We invite you to share updates about new monitoring on the horizon.
- Raise your hand using the Teams toolbar icon -> 
Raise
- Keep it brief, 1-2 minutes



TMS UPDATES

2025 TMS Tasks

- Screening Values
- Matrix and dashboard updates
- Toxics Science and Monitoring Strategy



TMS Workshop

- In conjunction with the May 2025 CRBRP Working Group Meeting
- Hybrid
- Date and location TBD





THANKS FOR JOINING US!

Questions? Want to join the TMS distribution list?
Email us at gs-crbtoxmon@usgs.gov