

Supporting Water Quality Exchange

Improving water
quality data systems



ENVIRONMENTAL POLICY
INNOVATION
CENTER

Part of the Series:

Cracking the Code: Federal
technology innovation to
heal our natural environment



Date Published: **November 13, 2023**

Why This Matters

- The federal government does not have a complete picture of our nation's surface waters, leaving us vulnerable to pollution and climate change impacts like increasing flood events.
- Much of what we do know is passed through EPA's system for ingesting water quality data, Water Quality Exchange (WQX).
- States are required to evaluate "all available data" to inform waterway and watershed protection through state and federal programs—and WQX is a major entry point for this data.
- Challenges with WQX prevent small watershed organizations with significant in-the-stream knowledge from uploading data.

What To Do

- Revamp and modernize WQX's interface and documentation to improve public access and support local assessment of water quality.
- Increase WQX budget for more 1:1 support and proactive outreach.
- Support technology developers building connections between groups and their data sources.
- Look to Mackenzie DataStream for examples of how to support organizations and design an accessible water quality data system.

Special thanks to Blue Water Baltimore, Chesapeake Monitoring Cooperative, and The Commons for making this case study possible – and to Jessica Mahr and Reed Van Beveren for their contributions.

Why This Matters

With over 3.5 million miles of streams and rivers, American waterways could circle the globe 140 times over—without accounting for the 100,000+ lakes and reservoirs. These waterways feed drinking water infrastructure, facilitate billions of dollars in economic activity, and provide crucial habitat for both aquatic and terrestrial wildlife. While we depend on these unsung hydrological heroes for our own survival, we know little about their overall health. An analysis of the EPA’s 303(d) Program—which identifies impaired water bodies across states and territories—found that states regularly used “insufficient water quality information... [and] data.” Lots of this data is collected by the federal government, but evidence from the state of California shows gaps in data collection like out-of-service U.S. Geological Survey (USGS) stream gauges—a pattern found across the country. Ensuring that our nation sufficiently collects robust data—and incorporates it into environmental decision making—is crucial for protecting our waterways.

What we *do* know about waterways comes from government monitoring efforts and a committed contingent of Riverkeepers, watershed organizations, and Tribal Nations. Whether collected by USGS, EPA, or community organizations like Blue Water Baltimore, consistent access to all available data through government portals is key (e.g., EPA/USGS’s Water Quality Portal or How’s My Waterway). These portals contain data that inform key programs like the Clean Water Act’s (CWA) 303(d) process by requiring states to “consider all available information” when submitting a list of waterbodies for restoration and protection. To submit data, entities use the Water Quality Exchange (WQX) system. Yet, as this case study will illustrate, the design and reputation of WQX dissuades submitters despite a robust and flexible data model—in effect, leaving behind troves of key data and ultimately hamstringing protection efforts and public awareness.

Consider Blue Water Baltimore (BWB), a well-established monitoring program with more than 50 sites across the region and testing methodologies for tidal stations meeting EPA’s highest standard for data quality. A Water Quality Portal search for sites in Baltimore with at least one sample between 2018 and 2023 returns fourteen locations—none of which are from BWB. Only adding to the confusion, How’s My Waterway reports significantly more stations for the *same search*—mostly from USGS’s Maryland Water Science Center—but with none from BWB. These inconsistencies beg the question: *Why is data from a federally approved community science source not found in either portal?*

The answer—and solutions—require a look into WQX’s reputation and interface. Without an easy and attractive channel between non-government data collectors and government portals, data languishes in file cabinets and hard drives, inaccessible for evaluation. Originally derived from STORET, an EPA system created in the 1960’s, WQX was released in 2009 allowing for entities to submit monitoring data electronically. Once in WQX, the data would be accessible via public-facing government websites like How’s My Waterway. In practice, however, smaller data collectors are turned off by WQX’s interface, the lengthy setup, and seemingly difficult submission process.

What to do?

Revamp the User Experience

Harkening back to the days of dial-up Internet, WQX's design and user experience is reminiscent of tech from the early 2000's—with an uncanny resemblance to Windows 95. While the WQX team at EPA is well-intentioned—and has released three full updates since WQX's inception—their efforts are hampered. Upstream factors like budgeting, complex federal technology procurement processes, and presumed limitations on user-centered design prevent a clean-slate overhaul.

For instance, after requesting an account over email and gaining access to WQX, data submitters are required to sort through a process data conversion to meet EPA's exacting meta-data standards (e.g., fields like “Vertical Datum Collection Method”). Specifically, groups often have to make modifications to the data schema, additional metadata, and change variable names to match WQX's allowable values. Faced with these daunting tasks, submitters ignore WQX altogether and focus on regional efforts, submitting their data to cooperatives who then (ideally) package multiple groups together for submission to WQX. While this does allow states to consider data in the 303(d) Program (and hence, to inform regional pollution initiatives like the Chesapeake Bay Program), it can significantly delay the reporting time and introduce room for error.

Group Outreach and Developer Support

The heart of the issue with WQX is a *perceived* mismatch between the skillsets community science organizations have around data management, and the level of support EPA provides for organizations to upload their information. Community science organizations are typically staffed with field biologists, community organizers, and outreach specialists—and much less often, data scientists and developers. WQX's system for data submission is robust and flexible but the documentation and user experience presents a challenge that resource-strapped organizations think they don't have time or capacity to solve. They collect their data in the field, often knee-deep in a culvert with a clipboard, for their primary audience: the general public via something like baltimorewaterwatch.org. Yet when it comes time to submit data to WQX, they are stuck navigating the conversion process for a standard wholly different than their own.

To rectify these challenges, EPA has taken significant steps to support organizations. Users can convert their data via pre-created templates or generate an import configuration to map their data to WQX's standards. Data submitters can also access one-on-one consultation to assist in template configuration. Application developers in the environmental space, like The Commons (who manage BWB's public website), have made improvements by creating conversion applications and uploads through an API—but have found integrating with WQX's API structure a struggle given its relational format and sparse documentation. Bolstering staff time to expand user support with more individualized support, webinars, and resources like forums and use-cases will get more data into the system—but requires considerable effort. Beyond staff support, WQX could also use a face-lift to better advertise its capabilities and build trust in the system. Here's a simple analogy: WQX is like an old house—its has great bones and will stand for another hundred years, but is also in desperate need of modern appliances, remodeling, and a fresh coat of paint. Until then, it'll be passed on by first time home-buyers daunted by the DIY.

A Move to Modern Workflows

Returning to Blue Water Baltimore, it's clear that they would rather not navigate challenges with WQX, and instead submit to a regional effort with closer relationships. That effort is run by Chesapeake Monitoring Cooperative (CMC), a group that collates data across regional organizations and hosts it on Chesapeake Data Explorer. In conversations with data managers at CMC and BWB, it became clear that many of the problems with WQX's accessibility cause issues in reporting. Once BWB collects their data, it can often take a week to run it through a conversion process and send to CMC. Once submitted, data managers at CMC will then need to convert the data into WQX's format and upload it on behalf of BWB, further lengthening the timeline. At the time of writing, CMC is planning to upload BWB data to WQX and is currently working with WQX staff to develop import configuration templates.

Instead, focus should be put on limiting the friction between one system to another. If BWB could automatically convert their data to CMC and WQX formatting, and simultaneously submit to both, it would dramatically reduce lag time. To enable this WQX needs resources: with a limited budget and near unlimited responsibilities for one dedicated WQX staff member, it's unlikely that EPA can proactively reach out for custom import configuration setup or host frequent training opportunities to integrate conversions with existing workflows. An extra step is added to the process by CMC being a data stop-over. And while regional cooperatives are incredibly valuable, they shouldn't define data transfer workflows. Like an old house in need of an expensive remodel, WQX is attractive to entities with more resources, like the state of South Carolina and USGS, both of which have moved to collect and submit their data directly to WQX. The scenario gives further credence to the true value of WQX and its underlying data model (which can ingest over 1,000 different data types); but today, the small user still gets left behind.

Follow our Northern Neighbors

Conveniently, the hard work of envisioning and building a modern water quality data submission system has been tackled in Canada. A collaboration between nonprofits and local governments, Mckenzie DataStream, actually copied much of WQX's data schema—but simplified it and built features like a modern API, clear documentation, and capacity to co-develop integrations between existing systems like the one used by BWB and developed by The Commons. The upload template even includes “Tips and Tricks” and calculators to help convert samples into the required formats! For the less technically savvy, the program provides extensive 1:1 support to help watershed organizations clean and upload their data. Moreover, the interface is intuitive and simple, meeting the standards we have come to expect from a modern website enhancing trust and garnering more users. While rebuilding a federal data system and providing hands-on data support are no small tasks, the Canadian model shows how these investments are essential for incorporating hard-earned data into conservation processes. The model also reminds us of this truth: what gets measured gets protected.