Water Quality Portal Strategic Plan

Guiding the Portal over the next five years (2017-2021)
Charles Kovatch, Jim Kreft,

Water Information Strategies Work Group, Mary Skopec, Chair

National Water Quality Monitoring Council

Vision

Be the premiere source for water quality data for everyone, everywhere.

Mission

Provide easy access to all water quality data, facilitate improvements in data quality, and enhance data discovery and data summaries to inform sound water-quality decision making at local, state, regional, and national scales.

Scope

Water quality data collected from discrete samples of ambient surface and groundwater in the United States.

Purpose

This document describes a 5-year strategic plan for the Water Quality Portal (hereafter "Portal"). The plan includes year one to three and year five goals and priorities to enhance Portal functionality and performance, engage new partners, educate partners, identify new data sources, and support the Council's mission to improve the quality and accessibility of the Nation's water quality data.

Background Water Quality Portal

Twenty years ago, the Intergovernmental Task Force on Monitoring Water Quality (ITFM), a subgroup of the Advisory Committee on Water Information (ACWI), published *The Strategy for Improving Water Quality Monitoring in the United States*[1]. The ITFM report recommended a "strategy for nationwide water-quality monitoring and technical monitoring improvements to support sound water quality decision making at all levels of government and in the private sector" and identified the establishment of the National Water Quality Monitoring Council (Council) to guide the overall implementation of the strategy. The report also outlined a number of water monitoring activities that have become key functional roles for the Council[2], of these, functions such as monitoring design, collaboration, communication, assessment, reporting, evaluation, and implementation would not be as effective without a modernized data system that could be used to house, locate, and deliver water-quality data from multiple sources. The data system developed to support these activities has become known as the Water Quality Portal (Portal).

The Portal is a cooperative service sponsored by the United States Geological Survey (USGS), the Environmental Protection Agency (EPA), and the National Water Quality Monitoring Council that integrates publicly available water quality data from the USGS National Water Information System (NWIS), the EPA STOrage and RETrieval (STORET) Data Warehouse, and the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) Sustaining The Earth's Watersheds - Agricultural Research Database System (STEWARDS).

The Portal is able to serve water quality data from NWIS, STORET, and STEWARDS through one seamless location by leveraging the Water Quality eXchange (WQX) data standard. WQX was published through the collaborative efforts of the Council, EPA, USGS, states, and other partners to enable the sharing of physical, chemical, biological, habitat, metric, and index data. There are now more than 400 federal, state and tribal agencies, watershed organizations and others using WQX to share data.

Since being launched in April 2012 with 150 million records, the Portal now exceeds 290 million records, from over 2.3 million sites. In addition, March 2016 marked the first time the website visitation exceeded 3300 visits, a 60% increase over March 2015 and the trend of increasing visitors continues, with monthly visitation in October 2016 10 times that of October 2013. In those visits, users

---

[1] *The Strategy for Improving Water Quality Monitoring in the United States*
http://pubs.usgs.gov/of/1995/0742/report.pdf
[2] The National Water-Quality Monitoring Council (NWQMC) Approved by the Advisory Committee on Water
[2] The National Water-Quality Monitoring Council (NWQMC) Approved by the Advisory Committee on Water Information February 18, 1998. Revisions approved by Council April 11, 2005.
http://acwi.gov/monitoring/about_the_council.html

downloaded over 500 million discrete result records. The Portal is also seeing increased use of automated web services with 162 million result records being downloaded via direct web service calls. In comparison, in June 2014, there were 95 million results records downloaded through the Portal website and only 9 million through direct web service calls. The trend in increased use has continued, with October 2015 setting a record of 41 billion rows of site and result data served.

Monthly Water Quality Portal Visits
October 2013 through October 2016
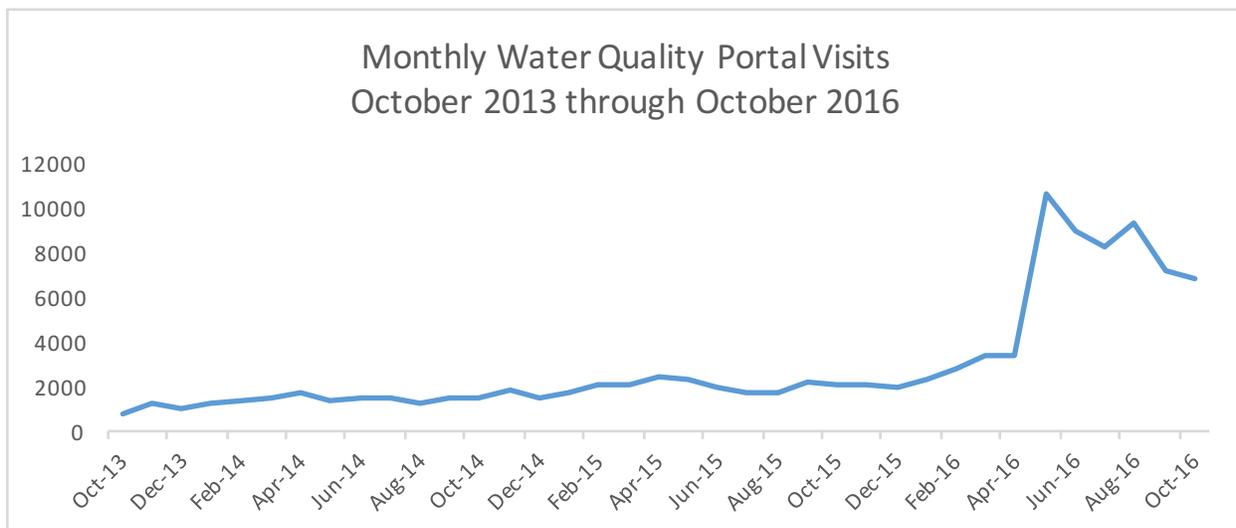
Note: *The jump in visits in May 2016 is due to an increase in visitation from search engine users (e.g. Google) after the launch of descriptive site pages that made the Portal's data more discoverable.*

The Portal's demonstrated success in sharing a large volume of water quality data via web services places it at the forefront of the ACWI Open Water Data Initiative (OWDI)[3] as a successful data sharing and web service delivery model. The OWDI was proposed in 2014 to integrate water information into a connected national water data framework that leverages existing systems, infrastructure, and tools. As part of the OWDI, three pilot projects are exploring ways to improve access to data and develop sustainable data utilization solutions on a National Flood Interoperability Experiment, a Water Supply (Drought) Decision Support System, and Spill response/Water Quality issues. As part of the pilot discussions, the Portal provided technical and strategic examples on data sharing, data utility, and data visualization.

Water Quality Portal Challenges

In its role providing guidance to the Council on water-quality monitoring strategies and communicating the quality of the Nation's water resources, the Water Information Strategies (WIS) workgroup was asked to provide input regarding potential improvements for the Portal. As noted above, the Portal has experienced significant increases in the amount of data the Portal can access and usage of the site by stakeholders and the public. Although members of the WIS workgroup expressed a sense of

---

[3] Open Water Data Initiative http://acwi.gov/spatial/open_water_data_charge_to_fgdc_june23_2014.pdf

accomplishment for reaching these significant milestones, the workgroup also identified a number of needs that should be addressed in the strategic plan being developed by the Portal Team including:

a) A need to improve the Portal system performance to keep pace with the increased usage
b) A need to improve data discovery and mapping tools of the Portal
c) A need to increase education about the Portal among Council members
d) A need to leverage other Council workgroups to promote the Portal
e) A need to increase education to engage water partners in providing and using data
f) A need to identify new data sources and support data sharing
g) A need to identify options to share Portal costs
h) A path to support the OWDI
i) A path to support integration of discrete data in the Portal with a continuous water-quality data portal that would be developed as part of OWDI[4]
j) A path to further support data evaluation tools built by others

The WIS workgroup conducted a survey in December 2014 to request feedback from the Council members regarding a long-term vision for the Portal and priorities for achieving that vision. The 17 respondents indicated:

- Their primary reason for visiting the Portal is to search for data
- They most like the ability to access data and query easily from the Portal
- They also cited challenges in their using the Portal including:
  - A need for staff and funding
  - The query interface is difficult to use
  - The mapping capabilities need to be improved
- The most wanted enhancements to the Portal from respondents were access to more federal partner data and simple data displays

The WIS workgroup then held a discussion on a new Portal vision at the Boise, ID Council meeting in February 2015 to gather additional feedback that could be incorporated into a strategic plan. (See the Appendix for list of comments). Following the Boise meeting, the WIS workgroup and Portal Team developed an outline for the vision and associated strategic plan and presented it to the Council members in May 2015. This document represents a compilation of all received comments on suggested milestones for the Portal's development, engaging new partners, educating more partners, identifying new data sources, and expanding or integrating the Portal's activities among the Council members and functions.

---

[4] Both of these needs are important but software limitations and a desire to focus the year 1-5 strategic plan on adding new data, adding checks to improve current data, and enhancing data discovery and mapping tools mean these needs may have to be addressed outside the Portal.

Water Quality Portal Strategy – Overview

The WIS workgroup developed a Strategy that will enable the Portal to fulfill the vision, meet the needs of internal and external customers, enable growth, and communicate progress to stakeholders. The success of the Strategy is dependent upon the full participation of all Council workgroups, including the WIS, Collaboration and Outreach, Methods and Data Comparability Board, and National Monitoring Network workgroups.

The Strategy is described in 5 sections: System Performance, Data Quantity and Quality, Data Display, Outreach and Strategy and Status.
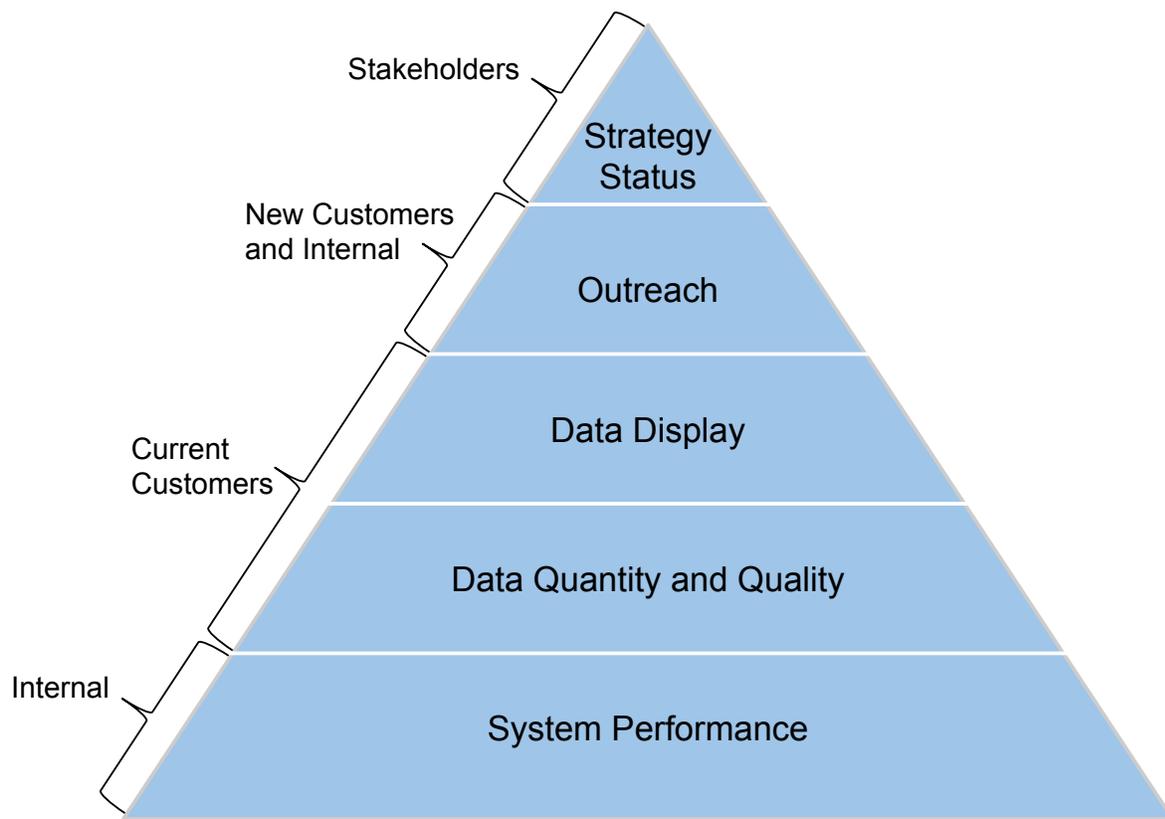
The Portal Strategy Pyramid: Each component of the Portal strategy depends on everything below it. The intended audience for each strategy component is indicated with brackets.

1) System Performance

System Performance is the foundation of the Portal; it includes the development, operation and maintenance of the Portal hardware (e.g. servers) as well as the software necessary to keep system components fully functional and up-to-date. Proposed updates will aim to improve the Portal's ability to deliver high quality data quickly, improve overall system performance, and identify ways to improve efficiency while maintaining or reducing operating costs.

One-Three Year Goals

1.1 Improve the delivery of data from EPA STORET to the Portal. The data presented by the Water Quality Portal does not include the full WQX dataset that is submitted by Water Quality Exchange partners. The Portal Team will explore ways to improve the design of the STORET data warehouse and the loading process to the portal system to increase system efficiencies and reduce data transfer costs.

1.2 Improve the delivery of data from USGS data systems to the Portal. The USGS is modernizing its discrete water quality and biological data management systems. As this modernization process moves forward, the Portal Team will take advantage of the best source of data available, while working to make sure that the Portal can most effectively serve the data access needs of both USGS and outside users.

1.3 Plan Portal platform future. With the growing demands to access more data, deliver a high volume of data quickly, and to incorporate new technologies to the Portal (e.g. Cloud), the Portal Team will evaluate the system platform, and develop baseline metrics to allow evaluation of future portal performance. The platform was updated in 2015 into a scalable, componentized approach that should serve our current data performance needs for 3-5 years. Keeping in mind upcoming infrastructure changes at the USGS Office of Water Information, the Portal Team will set performance benchmarks and identify a schedule for system updates over time for database tools, database operations, and Portal applications. These changes will allow for maintaining or exceeding performance benchmarks, reduce risk and reduce the cost of responding to future needs.

Five-Year Goals

In five years, with dedicated funding, all three activities would be implemented to realize EPA and USGS systems are shared and integrated with the Portal, an updated STORET solution, BioData available, and Portal performing on an updated platform or migration to a Cloud environment.

Goals	Activity 2017-2019
1.1 Improve EPA delivery to Portal	STORET data transfer design plan
1.2 Improve USGS delivery to Portal	Work closely with NWIS modernization team to align the Portal with future USGS system state and requirements
1.3 Assess Portal platform	Portal system platform update plan

2) **Data Quantity and Quality**

Almost every conversation about the Portal quickly identifies an interest in increasing both data quantity and the consistency of data quality. While there are over 250 million physical, chemical and biological data records available, there are still vast amounts of data from federal, state and tribal agencies, watershed monitoring groups and other sources which are not yet available via the Portal. The Portal currently does not serve important biological information such as habitat, metric, or index data. In addition, there are still many inconsistencies in how data and metadata have been recorded by data owners and there is a need for simple data checks to eliminate basic errors (e.g. incorrect units, transcription errors, etc.) prior to the data being added to the Portal.

One-Three Year Goals

2.1 Reduce data-in barriers. Many data owners often don't know that the best way to make data available via the Portal is through EPA's Water Quality Exchange (WQX). To support data owners, the Portal Team will create and provide training webinars aligned with field collection schedule and when possible, on-site training at large draw meetings such as the National Water Monitoring Conferences. In addition, the Portal Team will explore updates to the WQX and WQX Web Tools and other data tools (e.g. an R to WQX) to improve user experience in sharing data.

2.2: All fifty states submit data to WQP. In coordination with the NWQMC, the Portal Team will reach out to remaining states to determine barriers and next steps for submitting data to the WQX system for inclusion in WQP.

2.3 Identify states to pilot input of biological data into the Portal. To date, most data sharing efforts have focused on physical and chemical data. The Portal Team will identify three data owners to work with over the next year to develop an approach for sharing their biological data holdings.

2.4 Improve Portal services. Over the next year, the Portal Team will work to make habitat, metrics and index data available through the Portal. The Portal Team will also stay in close contact with OWDI, EPA, and USGS with respect to continuous data so that future integration of continuous and sensor data as part of the OWDI will be as seamless as possible.

2.5 Explore options to make Legacy STORET data available to Portal. The STORET Legacy data contains over 150 million records in a static database. There have been many requests for these data to be made available to the Portal, however the data represent a wide variety of documented and undocumented records. The Portal Team will explore ways to map the data elements to the WQX schema and make the data available via the Portal so it is clearly marked as Legacy data and not confused as current documented data.

2.6 Improve data quality fields/issues. The Portal delivers data provided by the data owner, along with required and optional metadata or information about the data. There can be different interpretations on how to populate the various metadata fields, such as how samples were collected and analyzed. In addition, some data is lacking basic quality assurance/quality control issues including incorrect units, etc. The Portal Team has started to explore data quality improvements and will work on a plan to begin addressing the data quality issues. The WQP Team will also contribute to the ongoing development of the WQX 3.0 schema and work with the EPA STORET Data Quality Work Group.

2.7 Expand QA Information from the Portal. Explore and develop a plan to serve additional QA information data, such as QA Plans and SOPs through the portal, possibly taking advantage of other NWQMC-sponsored tools such as NEMI.

2.8 Explore error reporting approaches. It is currently unclear for Portal users how to report data issues in data that they download. The WQP team will explore approaches to make it possible for Portal users to contact the data owner for specific data quality issues.

Five-Year Goals

In five years, there will be an even more trained and knowledgeable community in sharing data via WQX, an increase in data sharing due to new data partners and types of data, improved community practice on data quality on (5) issues, data shared via WQX will better support a larger range of data reuse, and the Portal will be recognized as the reliable source for water quality data.

Goals	Activity 2017-2019
2.1 Reduce data-in barriers	Hold (2) training webinars, (1) on-site training, update WQX Web tool
2.2 All fifty states submit data	EPA will reach out to remaining states
2.3 Biological data	Work with (3) states
2.4 Portal Services	Serve habitat, metrics, index data. Generate plan for integrating with continuous water-quality data.
2.5 Legacy STORET	Determine options for moving forward
2.6 Data Quality	Identify data quality issues. Generate data quality plan. Begin addressing (2) issues.
2.7 Expand QA Data from the Portal	Develop a plan to serve additional QA data
2.8 Explore error reporting approaches	Pilot approaches for reporting data issues

3) Data Display and Discovery

The Portal has been successful in enabling users to quickly query and download data from the Portal. There is a demand to make the data from the Portal more useful or more easily incorporated into analytical tools, statistical packages, or water quality models to support water managers to make water resource decisions.

One-Three Year Goals

3.1 Portal Functions. Data can be quickly retrieved via the Portal, and there is growing demand for quick views of data, data quality and data trends and displaying the data on a map. The current mapping features enable quick views of station locations, and incorporating more map features, for example up-stream/down-stream views provided by NHDPlus that make the data more useful. The Portal Team will develop a plan to identify the map views that will support data user needs while maintaining fast Portal performance. The Portal team will also make summary pages of data available at user-relevant scales (e.g. organization, HUC, County) so that users will be able to know that data is available from the Portal without having to download and analyze data.

3.2 Data Tools. The Portal makes data available for users via query download and web services to incorporate into their analytical tools, and the EPA side of the WQP is developing a data discovery tool that runs off of Portal services. Using lessons learned from the data discovery tool, a WQP summary service is planned to further enhance the capability of other data access tools. There is an increasing demand to better guide users in how to incorporate data into their tools that, for example, can help them in their data assessment and data analysis. The Portal Team is pursuing an open source community approach for the development of these tools. Using open source software like 'R' to query and analyze the data from the Portal enables the user community to leverage these tools and adapt them to their specific needs. The Portal Team, with WIS and OWDI, will develop a plan to support a community of practice and a water-quality monitoring marketplace of tools and products. Having an active community will better enable organizations and citizens across the country to share tools that leverage the Portal and make Portal data accessible and relevant to as broad of an audience as possible

3.3 WIS User Focus Group. Work with WIS to put together a focus group to come up with additional ideas that could be evaluated from a cost/benefit perspective to drive development of tools and functions in years 2-5

Five-Year Goals

In five years, the Portal will support a community of practice where the Portal supports sharing tools that enable users in their data discovery, data analysis and data display for water data monitoring trends.

Goal	Activities 2017-18
3.1 Portal Functions	Plan for new functions on data quality, monitoring trends and mapping.
3.2 Data Tools	Plan for (1) new analytical tool for data discovery or assessment.
3.3 WIS User Focus Group	Develop year two development goals

4) Outreach/Education

It is important to communicate the Portal's success in increasing the number of data records downloaded monthly, to both internal and external partners. First the Portal Team needs to provide regular updates about the WQP functions to the Council members, workgroups, and committees, so that the members of those groups can then discuss and promote its use to others. This will require an internal education strategy. Second, the Council needs to conduct external outreach to the water community to communicate the successes of the Portal and explain how they can participate through data sharing, data downloading and analytical tool sharing. The Portal Team needs to provide the tools to make council members communication as easy as possible.

One-Three Year Goals

4.1 Increase Council knowledge. The Portal needs to be part of the daily vocabulary of all Council members so they can understand its basic function, communicate how to become a data partner, and describe how to use it. The Portal Team will work with all Council committees to develop an outreach and communication plan which identifies level of detail, audiences and frequency of "training" materials through factsheets, demonstrations, presentations, and videos for Council members and other external partners.

4.2 Identify data users and customer uses. There is a large gap in understanding between tracking how much and how frequently data are downloaded from the Portal against who is using the Portal to download data and how the data are used. The data download statistics indicate a high interest in the data, though we don't know if the data are being used for monitoring, assessment, compliance or other water management decisions. The Portal Team will work with Council committees to develop a user community framework to improve data user and data use scenario sharing.

4.3 Identify new data suppliers. In support of Goal 2: Data Quality and Quantity, the Portal Team will work with Council committees to attract new federal, state, and tribal data partners to share data. We will emphasize how getting data into the Portal can help with complying with mandates for open, machine-readable data.

4.4 Write a peer reviewed paper about the Portal. The Portal is a great success in data sharing and data integration, and many lessons have been learned, but those lessons have not been shared with the larger data community. A paper that discussed what had been learned and showcases how different stakeholders use the Portal would be both a valuable learning exercise and useful tool for communication.

4.5 Promote the Portal at national scale meetings. Promoting the Portal at national scale meetings, such as the American Geophysical Union, makes it possible to reach new users and new data partners.

Five-Year Goals

Over the five years, there will be more informed water partners and Council members able to communicate the function, role and benefit of the Portal to attract new data partners or create new partnership opportunities. There will also be more educated data owners and data users, who will be leveraging and sharing information and tools in an improved water monitoring and water data sharing community.

Goal	Activities 2017-18
4.1 Increase Council Knowledge	Develop outreach plan, factsheet and conduct (2) webinars
4.2 Identify data users and uses	Develop data user community framework and (2) use case profiles
4.3 Identify new data suppliers	Identify (1) new federal partner and (3) new state data suppliers
4.4 Write a peer reviewed paper about the Portal	Aim for publication in 2017
4.5 Promote the Portal at national scale meetings	Present at NWQMC, AGU, others

5) Strategy and Status

The Portal will be able to support the Council in setting strategic goals, planning, reporting monitoring status and trends, and attracting new collaboration opportunities. The Council will have access to quick displays and views of the Portal's holdings and active monitoring (e.g. in a watershed) to better communicate water-monitoring activities with partners. Thus, through use of the displays, the Council will become a primary user and advocate for the Portal's utility.

One-Three Year Goals

5.1 Identify status of monitoring data/programs. There is a need for the Council to identify the status of water monitoring data and program activities and share this information with partners. The Portal Team, alongside the Council, will work to identify key partners, water data business needs, and technical requirements, (e.g. filters for time, parameters, etc.) to implement the development of maps, displays, and views of the data in the Portal to support data discovery at user-relevant scales, such as watersheds or organization.

5.2 Develop NMN Networks based on WQP Data. The Portal Team will work with the NWQMC to develop a set of water-quality networks from WQP sites. Initial focus could be on surface-water quality networks that consist of sites that meet criteria associated with a specific objective such as reference sites, sites monitoring certain contaminant groups, biologic data, sample media, or trend sites. These networks could base site selection and minimum criteria on the Council's National Monitoring Network framework.

5.3 Identify new collaboration opportunities. As federal, state and tribal agency budgets and resources are strained, it becomes more critical to collaborate and leverage resources to meet water monitoring goals. The Portal can help make coordination and collaboration in a watershed easier by displaying where water monitoring is occurring, who is doing it, and what type of parameters are being monitored. The Portal Team will work with the Council to identify what components will make the displays useful.

Five-Year Goals

Over five years, the water monitoring community will see more water monitoring information available through the aided transparency of water monitoring data holdings. There will also be a recognized role for the Portal to provide water monitoring information in support of water prioritizing, collaboration and decision making.

Goal	Activities 2017-2018
5.1 Identify status of monitoring data/programs	Identify data partners and display needs. Develop (1) view or map display of water monitoring partners and data holdings
5.2 Develop NMN Networks based on WQP Data	Develop surface-water quality network, among others, from WPA sites
5.3 Identify new collaboration opportunities	Display active monitoring to support coordinated efforts