


ORIGINAL ARTICLE

Open Access



# Pilot implementation of the US EPA interoperable watershed network

Tad Slawecki<sup>1\*</sup> , Dwane Young<sup>2</sup>, Britt Dean<sup>2</sup>, Brandon Bergenroth<sup>3</sup> and Kimberly Sparks<sup>3</sup>

## Abstract

**Background:** The mission of the United States Environmental Protection Agency (EPA) is to protect human health and the environment, including air, water and land. Understanding the extent of pollution in waters and identifying waters for protection has been based in part on water quality monitoring data collected and shared by parties (federal, state, tribal, and local) throughout the U.S. To date, this monitoring data has been largely represented by data collected as a water quality sample (data collected by a technician in the field or analyzed in a lab). EPA's "STORage and RETrieval" (STORET) and the Water Quality Exchange (WQX) have served as the repository for all this sampling data. However, these tools and systems were not designed to handle today's continuous water quality sensors. EPA has therefore embarked on the Interoperable Watersheds Network (IWN) project, which is focused on identifying a common set of formats and standards for data, and on testing and validating these standards as well as new ways of sharing data and metadata. The completed IWN will greatly expand the sharing of data and its use, thereby streamlining the assessment, restoration, and protection of surface water quality at all levels of government.

**Methods:** Stakeholder workgroups were engaged to assist with developing requirements for the three major project components: required attributes and query capability for a centralized metadata catalog, technological and data requirements for data providers, and desired functionality for a web-based discovery tool that provides access to the catalog services and provider data.

**Results:** The pilot implementation of IWN uses the Open Geospatial Consortium (OGC) Sensor Observation Service (SOS) 2.0 and WaterML2 standards as the foundation for a distributed sensor data sharing network. Data owners in locations across the United States have worked with EPA to publish their continuous sensor data and related metadata either through "data appliances" running the open-source 52° North implementation of SOS or using commercial software like Kisters' KiWIS product.

Metadata are harvested into a centralized catalog that provides a REST Service API for sensor discovery. Users can discover data by querying for specific parameters, or using spatial boundaries such as HUC, county, a buffered point, or a user defined polygon. The sensor results are returned as GeoJSON, which can be used to create maps. The API also provides the service endpoints for the sensors, which can be used to access the continuous data to create charts or download the data for other analysis.

**Conclusion:** The pilot IWN demonstrates that standards-based interoperability can provide a sound basis for a national-scale clearinghouse for continuous sensor data, though scalability of the approach will need further testing. Selected technical detail, lessons learned, and future plans for the IWN are included in the discussion.

**Keywords:** OGC SOS, OGC WaterML, Discoverability, Access, Sensor Data, Water Resources

\* Correspondence: [tslawecki@limno.com](mailto:tslawecki@limno.com)

<sup>1</sup>LimnoTech, 501 Avis Drive, Ann Arbor, MI 48108, USA

Full list of author information is available at the end of the article

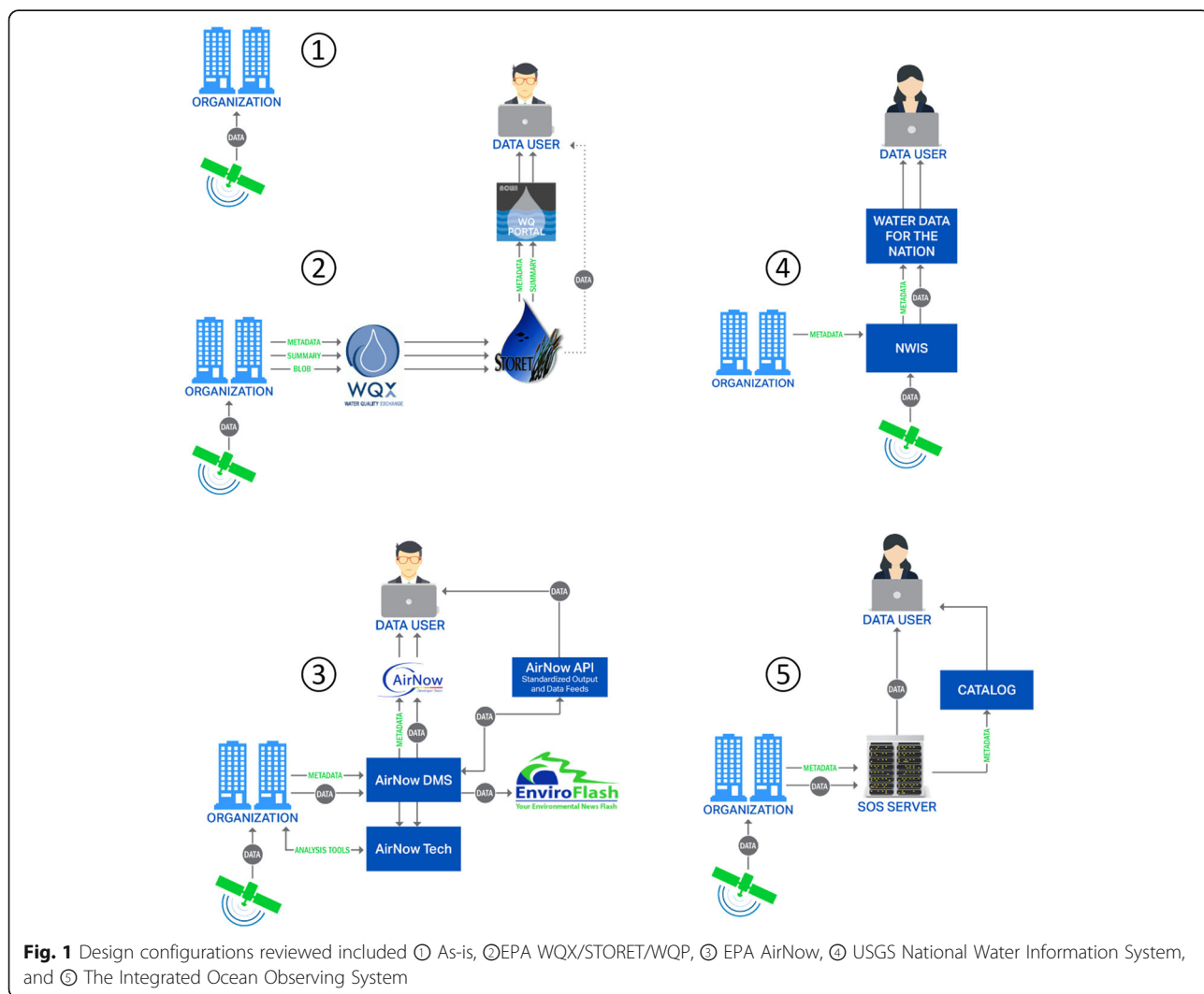
### Background

The United States Environmental Protection Agency (EPA) mission is the protection of human health and the environment, including the waters of the United States. EPA’s “STORage and RETrieval” (STORET) data system [4] has been used to collect and hold millions of water quality sample measurements and associated metadata collected since the 1960s. Additional systems like the Water Quality Exchange (WQX) [5] and the Water Quality Portal (WQP) [11] have facilitated the communications and exchange of water quality sampling data between data providers and promoted discoverability of and access to data across agencies. However, STORET, WQX and WQP emphasize the handling of discretely sampled “grab” data and are not well-suited to manage high-frequency “continuous” data generated by modern, affordable water quality monitoring sensors. The use of these sensors is becoming ubiquitous with a proliferation of this teleme-tered ‘real-time’ data on the internet and development of

new sensor technology for nutrients and other parameters of interest promises to expand and diversify applications.

Recognizing that access to these temporally dense datasets can help water resource managers to make better, effective and timely decisions, EPA funded the development of a draft strategy for sharing continuous monitoring data in partnership with states, tribes, and other federal agencies. Research, interviews, and discussions with a range of stakeholders identified a number of existing designs already in use at different U.S. federal agencies. Figure 1 depicts five design configurations evaluated for the draft strategy:

- ① The worst-case “as-is” scenario is widespread, with collected data not passing beyond the organization and discoverability minimized or non-existent;
- ② EPA’s STORET/WQX/WQP system for water quality data makes centralized grab sample data readily discoverable and accessible, but is not well-structured for handling continuous data;



**Fig. 1** Design configurations reviewed included ① As-is, ② EPA WQX/STORET/WQP, ③ EPA AirNow, ④ USGS National Water Information System, and ⑤ The Integrated Ocean Observing System

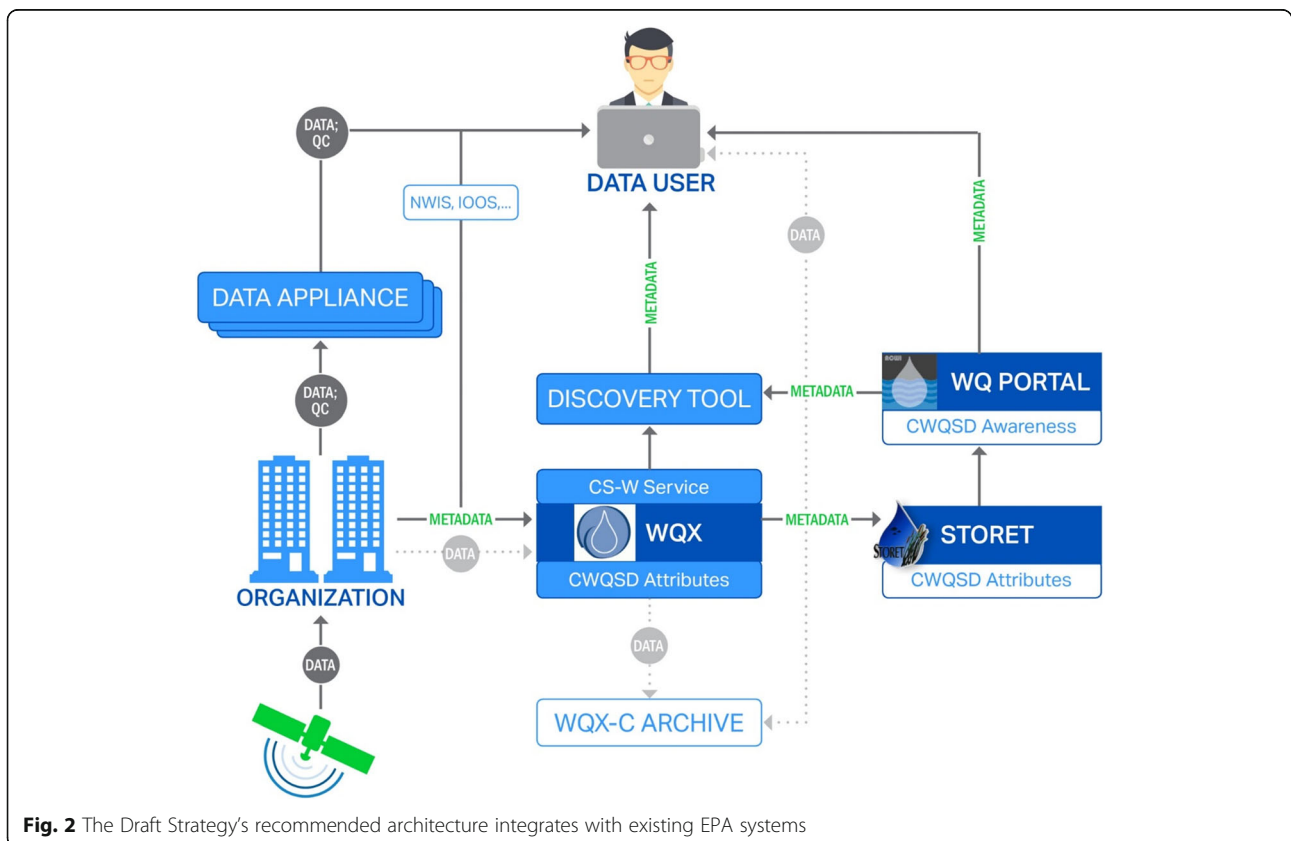
- ③ EPA’s AirNow centralized system handles continuous data well, but is currently focused on a highly controlled, homogenous set of parameters;
- ④ The US Geological Survey’s National Water Information System delivers centralized water data using OGC services; and
- ⑤ The Integrated Ocean Observing System is built around OGC standards such as SOS and combines a centralized catalog with distributed data;

A wealth of additional detail on these designs as well as descriptions and results of the review process is available in [10] and [3]. The recommended system architecture for an Interoperable Watersheds Network (IWN) shown in Fig. 2 called for the standards-based implementation of a centralized catalog, data appliances and archive, where (1) Site and deployment metadata are submitted to an extended WQX serving as a catalog; (2) Data are available through service endpoints exposed by data appliances tied to the organization; (3) Data users discover data of interest with a discovery tool by querying metadata in the catalog, and then retrieve the data from the data appliances; and (4) Data are archived for backup, redundancy and/or regulatory reasons in a modified WQX. The design process led to the identification of the Open Geospatial Consortium (OGC) Sensor

Observation Service (SOS) [8] and WaterML [9] standards as sound bases for an initial implementation. Use of these standards is consistent with OGC Best Practices as identified by the OGC Hydrology Domain Working Group [7]. The development of the draft strategy was coordinated with the Open Water Data Initiative, a complementary US Federal activity exploring the integration of water information into a connected national water data framework [2].

EPA subsequently funded the 2015–2016 pilot implementation of the IWN to validate the recommended architecture. The design laid out for the sensor data sharing network took advantage of the OGC standards to allow sharing of continuously monitored data using a common format. The recommended network therefore facilitates both discovery and dissemination of data, and contains the following key features:

- Organizations publish their data using SOS services through a variety of means;
- Data services and organizations are registered in a centralized catalog;
- Discovery and analysis are supported through a portal complementary to WQP for human use and through an application programming interface (API) for machine-to-machine use cases.



**Fig. 2** The Draft Strategy’s recommended architecture integrates with existing EPA systems

The pilot implementation of the IWN focused on two watersheds where continuous sensors were being deployed by organizations willing to partner with EPA in sharing their data:

- Hackensack-Passaic, New Jersey. The New Jersey Department of Environmental Protection (NJDEP) and the Meadowlands Environmental Research Institute (MERI) operate sensors in and around the Passaic River.
- Little Miami, Ohio. EPA’s Office of Research and Development (ORD) and Clermont County each operate sensors on reservoirs, tributaries and the main stem of the Little Miami River.

Stakeholders in the watersheds were engaged in site visits and on monthly calls to develop use cases, to define data workflows and attendant technology stacks, and to provide feedback throughout.

**Methods**

A straightforward software development approach was used that first elicited requirements for the major projected components and then iteratively implemented the components with many opportunities for stakeholder input.

Stakeholder workgroups were engaged to:

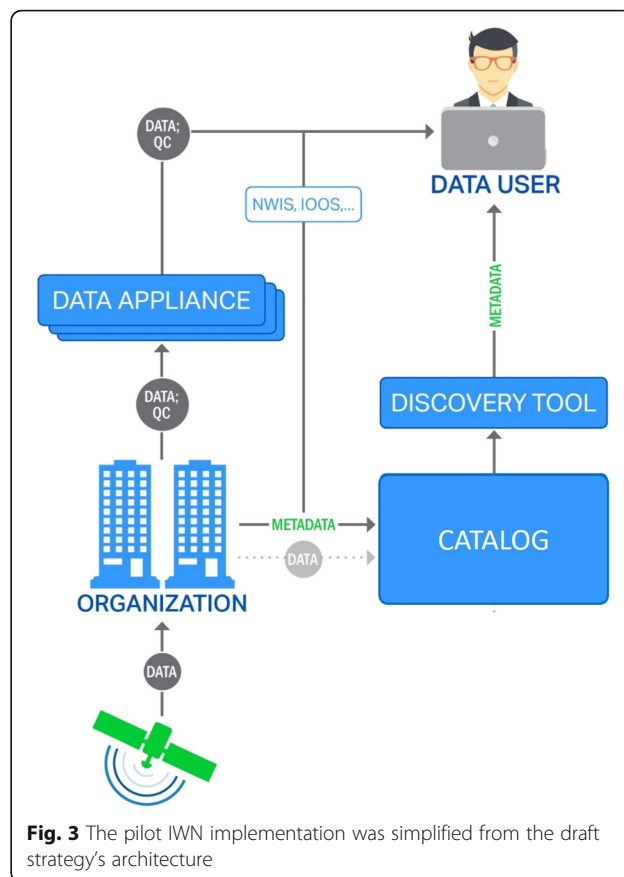
- Identify required attributes and query capability for a centralized metadata catalog,
- Specify technological and data requirements for data providers, and
- Define desired functionality for a web-based discovery tool that provides access to the catalog services and provider data.

Short, simple descriptions were solicited from representatives of the pilot watersheds to define user stories. These descriptions of desirable features presented from the stakeholder perspective were used as the launching point for an agile development process. Regular interactions with stakeholders served to inform the implementation towards its responsive endpoint.

**Results and discussion**

The system architecture identified in the draft strategy research was simplified for the purposes of the pilot IWN implementation (Fig. 3). This simplification was dictated by anticipated logistical issues with integration of the IWN into existing EPA systems. The simplified architecture resulted in three concurrent efforts:

- **Catalog Development.** A metadata catalog was developed to contain and serve necessary information to meet user expectations.



**Fig. 3** The pilot IWN implementation was simplified from the draft strategy’s architecture

- **SOS Services.** Partner organizations identified how best to make their data available through SOS services and then implemented—or helped to implement—the services.
- **Discovery Tool.** The Currents discovery tool consumes metadata and data services to enable data discovery and access.

As implemented, IWN data is currently made available using WaterML 2 and SOS 2.0 through either 52 N or Kisters servers with the SOS 2.0 Hydrology Profile enabled, so data services are compliant with the requirements in the Best Practice document for the OGC SOS 2.0 hydrology profile for SOS 2.0 implementations serving OGC WaterML 2.0 [7]. In addition, the related catalog and Currents discovery tool fulfill the common cases requirements for data discovery and download established in the Scope section of the Best Practice document.

The agile development process used in these efforts was heavily informed by regular interactions with the pilot watershed stakeholders. The development of user stories with the stakeholders was one of the key elements that determined the success of the pilot from their point

of view. These short, simple descriptions of desired new features, written from the stakeholder perspective, are presented in Table 1.

The Source Water Protection (Hackensack-Passaic A) and Water Safety (Little Miami A) user stories share a need for discovery and visualization, while the Water Quality Assessment (Hackensack-Passic B) and TMDL Implementation (Little Miami B) user stories call for large multiple-site, multiple-parameter downloads.

Consideration of these user stories together with other stakeholder input gave rise to the conceptual model shown for the metadata catalog in Fig. 4.

Key features of the catalog include:

- **Standardized vocabulary for parameter names.** Parameter names supplied by data providers are mapped to the appropriate name in the Substance Registry Service, which is EPA's "authoritative resource for basic information about chemicals, biological organisms, and other substances of interest to EPA and its state and tribal partners." [6]
- **Quality Assurance/Quality control (QA/QC) field.** The "Sensor QAQC" field provides a simplified mechanism for linking to appropriate QA/QC data such as sensor maintenance reports. The expectation is that data providers will populate this field with a hyperlink that points to the providers' collection of relevant QA/QC data and metadata.
- **QA/QC status.** Although some providers (e.g. the US Geological Survey ()) are able to provide observation-specific data qualifiers, QC status is generally not consistently available, and is not directly represented in the catalog data model. QC status is instead encoded as part of the SOS procedures.

**Table 1** User Stories From Pilot Watersheds

Hackensack-Passaic River Watershed	
A. Drinking Water/Source Water Protection Early Warning	<i>"As a water manager, I want to view trends in selected parameters so that I can predict and remediate a water quality issue before it occurs."</i>
B. Water Quality Assessment (for Clean Water Act Integrated Reporting)	<i>"As a water quality manager, I want to download continuous monitoring parameter data so that I can compare it to numeric criteria and evaluate if the water is meeting state standards."</i>
Little Miami River Watershed	
A. Water Safety (Drinking Water and Recreation) optimization: Maximizing output while minimizing cost with a Harmful Algal Bloom focus	<i>"As a water manager, I need to detect water quality issues, such as a harmful algal bloom, so that I can alert the public."</i>
B. TMDL Implementation	<i>"As a water quality manager, I need to be able to download parameter data for use in running TMDL models."</i>

An initial effort to build the catalog around the 52° North implementation of SOS [1] was set aside after recognition that new API queries would need to be coded. The metadata catalog was instead realized in PostgreSQL with accompanying REST services implemented using the Java Spring framework to deliver JSON or GeoJSON responses for different API queries:

- **GetOrganizations** retrieves the list of organizations that are currently registered as data providers along with service end point, the date of the most recent data harvest, when the server was last pinged, and an indication of whether the endpoint is available. The service accepts an optional organization id (org\_id) parameter which limits the results to the requested organization.
- **AvailableParameters** returns the list of parameters that are available for query via the metadata catalog.
- **GetSensors** (multiple) returns a feature collection which specifies the siteId, siteName, orgId and geometry (type and coordinates) of a sensor. There are separate services for spatial filtering by county, hydrologic unit, circular buffer, bounding box, and upstream/downstream relationship. All of these services accept an organization id (derivable from the getOrganizations service) and parameter id (from availableParameters), as well as a minimum and maximum observation date to constrain results.
- **GetSensorParameters** returns the list of parameters that are registered in the catalog for the specified input sensor ID.
- **GetOrganizationParameters** returns the list of parameters that are registered in the catalog for the input organization ID.

GetSensorParameters and GetOrganizationParameters results both include the organization's parameter IDs for use in querying data by parameter directly from the organization's service endpoint. The catalog harvests metadata from registered organizations' service endpoints daily.

### SOS services

The draft strategy and the design for the pilot IWN implementation both relied on the use of WaterML2 as the common format for data access and of SOS for management of the data. Version 4.3.6 of the 52° North implementation of SOS [1] was identified as a suitable platform around which to build "data appliances" for data providers as it supported WaterML 2 reporting of observations. Four different configurations were implemented for the pilot watershed partners (Table 2).

Details of setup and configuration are provided in a supporting GitHub repository at <https://github.com/IWN-Currents/OGD-materials>.



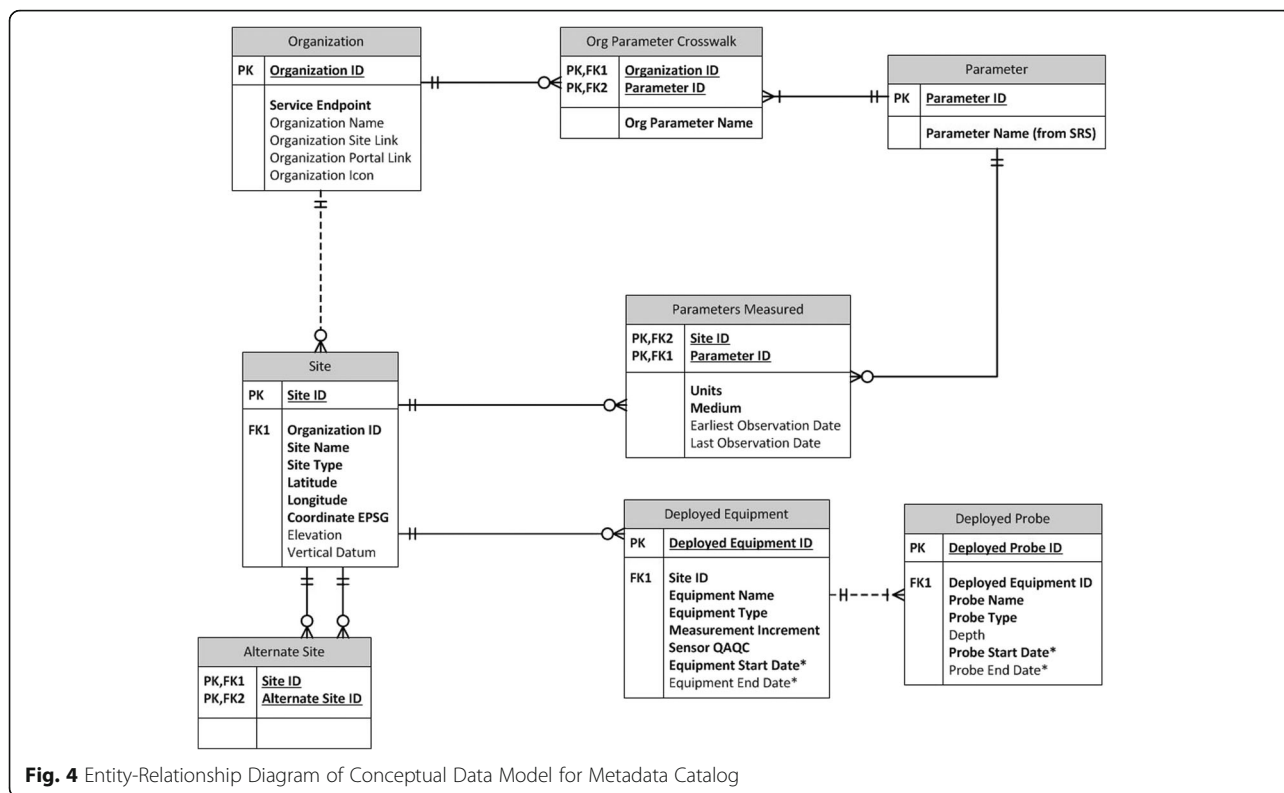


Fig. 4 Entity-Relationship Diagram of Conceptual Data Model for Metadata Catalog

Additional “partners of opportunity” were identified in the course of the project and incorporated into the IWN (Table 3). The ready incorporation of Region 1 and Region 10 data into the EPA server demonstrated flexibility of the IWN-configured data appliance approach, while integration of data from WaterML 2/SOS-aware commercial software validated the assumed interoperability of heterogeneous data server components.

**SOS ingestion**

Data ingestion procedures were designed to emphasize the use of simple text files. A Python script was written to parse comma-separated value (CSV) files containing observational measurements and combine the measurements with parameter and station metadata to form appropriate SOS 2.0 InsertSensor, InsertResultTemplate, and InsertResult service calls. The sensor networks

implemented on IWN data appliances share many key attributes:

- There is a common naming scheme for procedures, offerings, features and templates that reflects the IWN project, object type, data provider organization and sub-organization, location, data status, and parameter (e.g.urn:x-epaiwpp:template:epa:ord:esf-weather:raw:light-3).
- Observed data from each sensor in the system are presented to the user as a SOS Observation Offering.
- Offerings are each linked to an SOS Procedure describing the sensor that produced the data in the offering.
- Sensor procedures for all of the sensors at a station are grouped together as children of a station procedure. Each station procedure has an offering that is “undefined”.

**Table 2** Partner Configurations Implemented for Pilot IWN

Partner	Configuration	Comments
NJDEP	Linux server hosted at Rutgers University	Manual batch updates only
MERI	Linux server hosted on Amazon Web Services t2.micro instance	Near-realtime updates pulled from third-party site
EPA ORD	Linux server hosted in EPA private cloud	Manual batch updates only
Clermont County	Windows server on premises	Near-realtime updates pulled from Flowlink (vendor) database on same server

**Table 3** Pilot IWN “Partners of Opportunity”

Partner	Configuration	Comments
EPA Region 1, Region 10	Linux server hosted in EPA private cloud (same as EPA ORD)	Near-realtime updates pulled from EPA Region 1 and third-party sites
EPA Region 7	Kisters WISKI/KiWIS installation	SOS 2.0/WaterML 2 services provided by KiWIS
USGS	National Water Information System	Provides WaterML2-formatted data retrieval; required custom code for metadata harvesting

- Station and sensor procedures contain sub-organizational contact information, while the Provider section for the SOS installation contains organizational contact information.

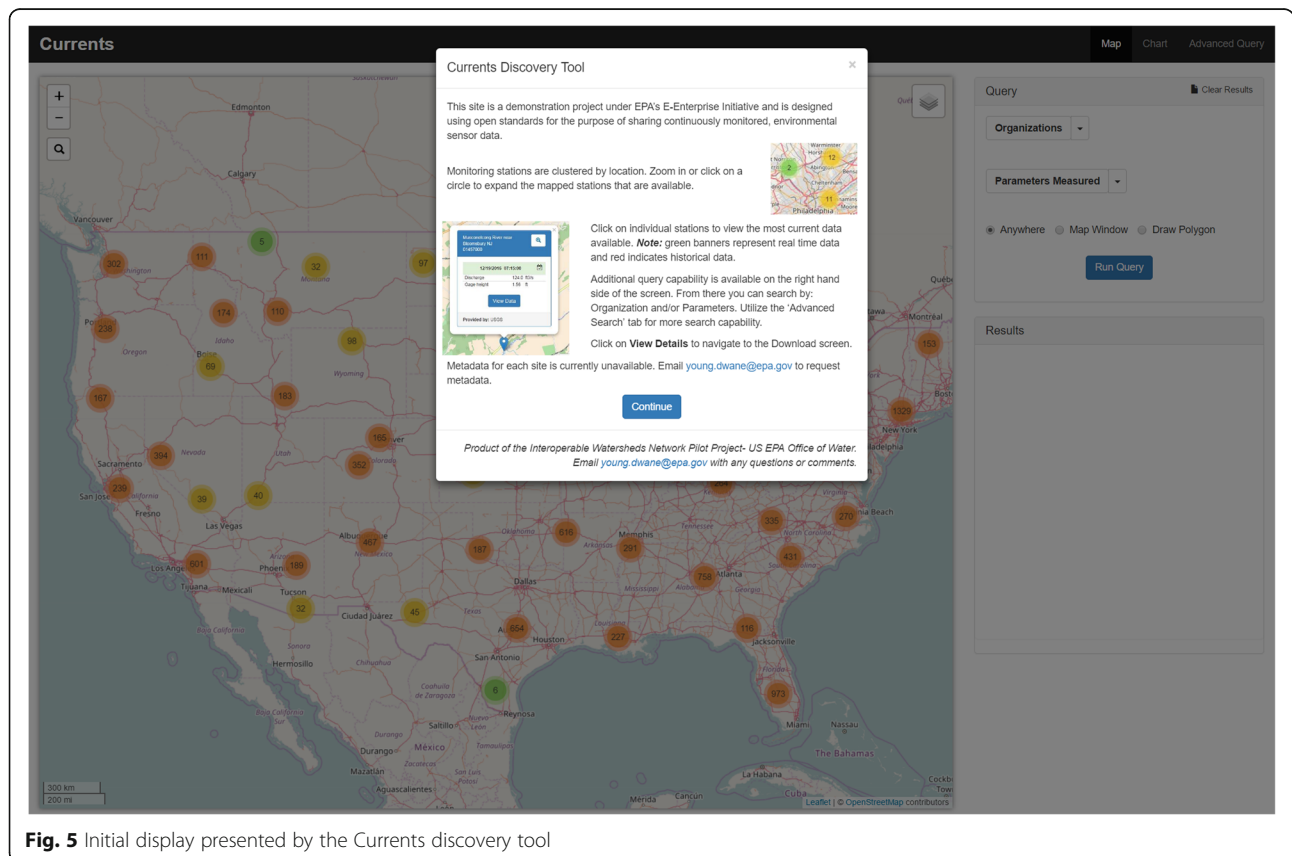
The ingestion code inserts stations, sensors, and observations using the SOS API, which allows it to be run locally or remotely, though local operation is recommended to simplify security settings for the SOS client. The code checks the SOS database to identify the most recent available observation for a given parameter and station, and only uploads observations that are more recent. Two typical use cases have been identified in the IWN pilot project: direct manual (batch) use for the occasional injection of long-term, typically historical and lengthy records, and scheduled invocation of `a.sh` (Linux) or `.bat/.vbs` (Windows) script for continuous near-realtime updates.

The ingestion script, example supporting files, and data appliance setup instructions are available from the GitHub repository at <https://github.com/IWN-Currents/OGD-materials>.

**Discovery tool**

The pilot IWN project also resulted in the development of the Currents Discovery Tool; a link to the tool is maintained at <https://github.com/IWN-Currents/OGD-materials>. Currents leverages the metadata catalog and SOS APIs to support discovery and visualization of IWN data. The initial Currents architecture is implemented entirely as client-based JavaScript.

Currents initially displays an interactive map that displays the sites registered in the metadata catalog (Fig. 5). Users can select sites directly in the map and view data, or narrow their selections using the simple query tools



**Fig. 5** Initial display presented by the Currents discovery tool

available on the mapping page or by using the more detailed tools available through the Advanced Query.

Features selected for query functionality were identified using input from the partner workgroups. The Currents tool allows users to filter data by organization, parameters monitored, and by identifying a date range for the observation results. Users can additionally use spatial parameters, such as the current map window, a user defined polygon and HUC-8 watershed or county boundaries to refine their selections. Partners also expressed a desire to select sites using a point and specified buffer distance and using stream network navigation; these features are included in the metadata catalog services, but are not yet available in the Currents tool.

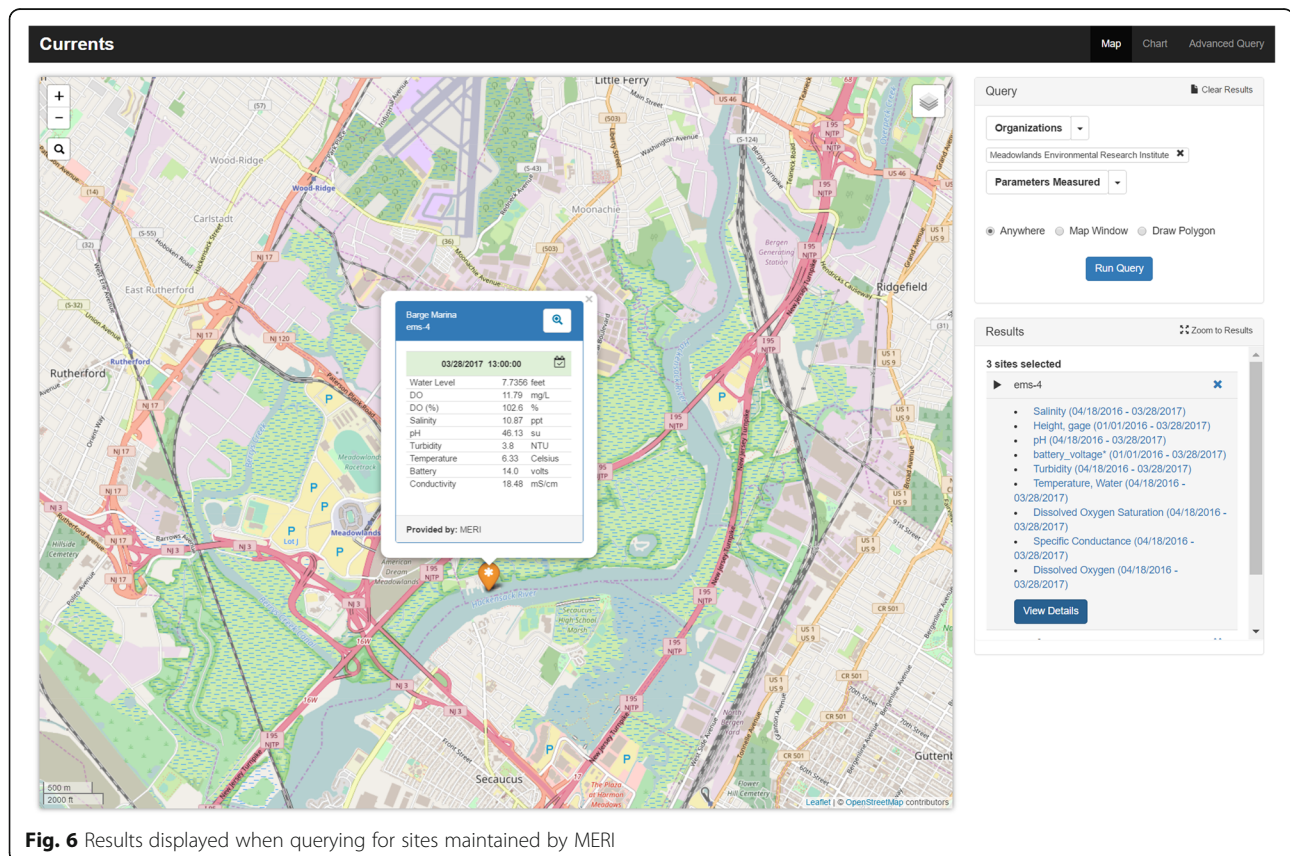
Query results are presented as an interactive list. Users can expand a listing for a site, view the data availability for the available parameters, and view the most recently measured observation data in the map pop-up window (Fig. 6). Users can select a parameter from the list and view a time series chart of the data (Fig. 7). The time series parameter data can also be downloaded as a comma-delimited text file from the site detail page (Fig. 8), which also provides access to data request URLs.

## Conclusions

The successful implementation of the pilot IWN demonstrates the feasibility of the original strategy for sharing continuous data, although scalability of the approach will be a concern. In particular, bandwidth, storage, and CPU requirements for the catalog server will likely increase as data providers engage with the IWN and register more data appliances. Data providers are deemed unlikely to run into scalability issues as data appliances configured for this pilot ran successfully on with minimal resources (e.g. Amazon Web Services' most-lightweight hardware configuration – t2.micro).

The IWN Project's overall successes include:

- Deployment of data appliances with varying configurations matching providers' data output formats.
- Implementation of an automatically updating metadata catalog and attendant API for web-based queries.
- Standards-based integration into the catalog of metadata both from IWN data appliances and from other interoperable data sources, demonstrating that a standards-based approach can address data source heterogeneity.
- Design and development of the web-based discovery and access Currents tool to fully leverage the catalog



**Fig. 6** Results displayed when querying for sites maintained by MERI



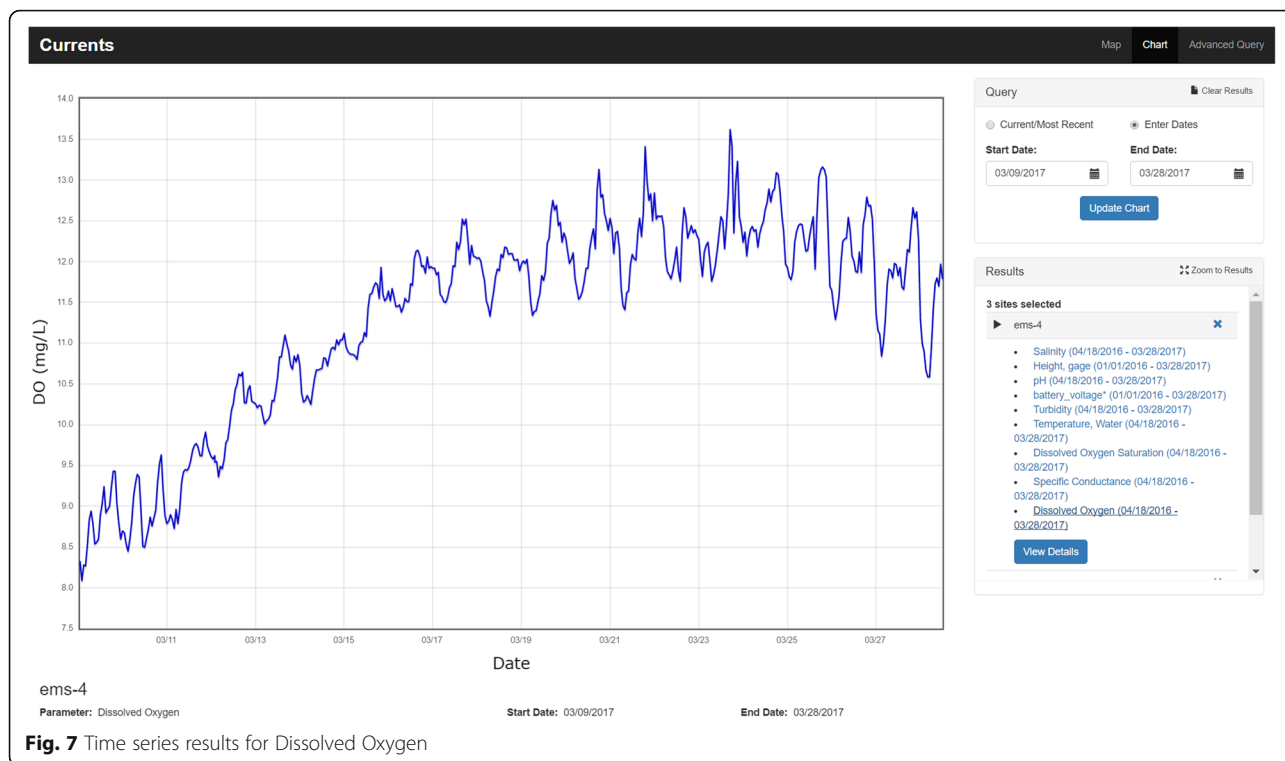


Fig. 7 Time series results for Dissolved Oxygen

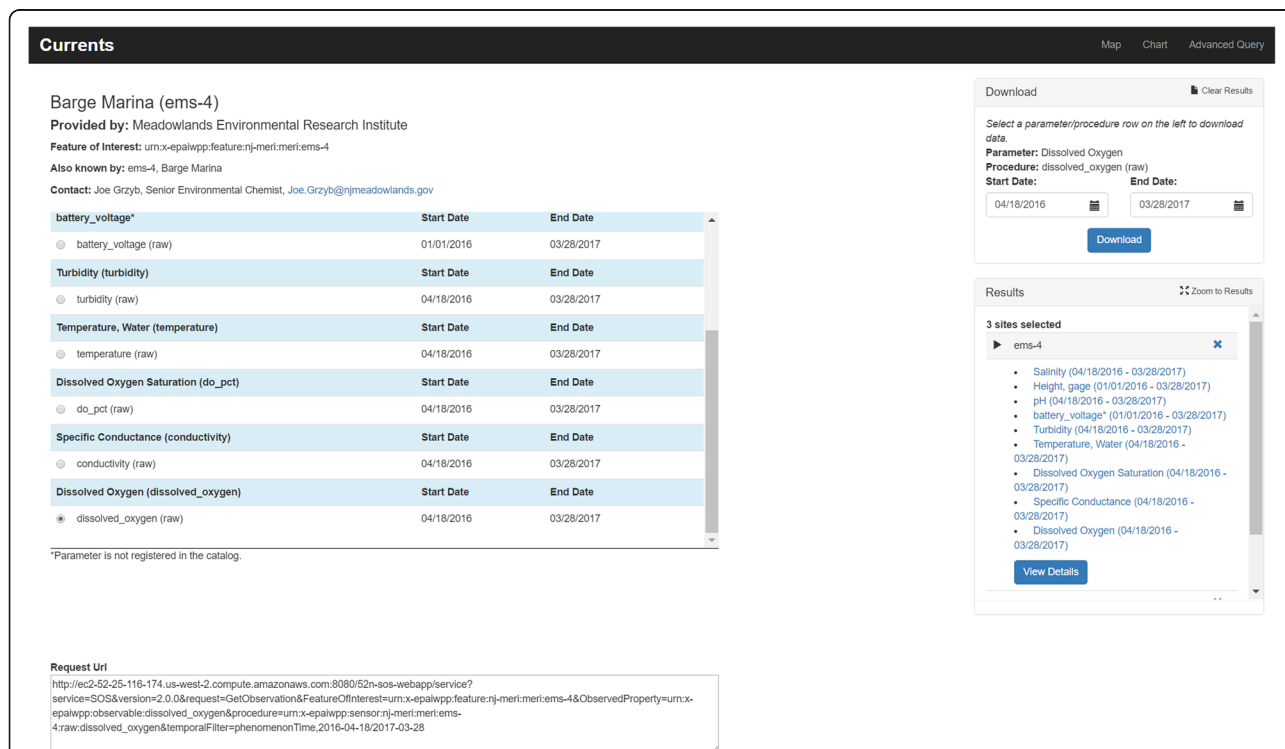


Fig. 8 Site detail for MERI station EMS-4, providing access to downloads and data request URLs

and data source APIs, e.g. by adding upstream/downstream selection and access to all metadata elements.

During the course of the pilot project, consideration of the various user stories and of stakeholder feedback helped identify feature requests for incorporation into SOS:

- The DeleteObservation request added by 52 North as an extension to the SOS standard is of high value and worth adding to the standard. Data partners sometimes identified errors in their data after posting, and DeleteObservation supports the replacement of erroneous data
- Observation-specific data qualifiers would be useful for the IWN to support user quality control information needs, but data qualifiers as defined in the WaterML2 standard (e.g. <wml2:qualifier xlink>) are not yet supported in the 52 N SOS database model. Observation-specific qualifiers can be included with InsertObservation requests using the <om:parameter > tag in the current development branch for 52 N SOS, but cannot be entered in InsertResult requests. Implementation of wml2 qualifiers and/or om parameters is desirable.
- Downloading of results from large GetObservation requests can be time-consuming, and it would be useful to provide the user with feedback on the progress of their request. One way SOS might help is to allow a temporal filter to be placed on the GetDataAvailabilityRequest to allow the querying individual/software to assess roughly how large a given retrieval might be to set expectations and perhaps strategy, such as breaking down the retrieval into smaller subretrievals.

As of this writing (December 2016), EPA is exploring next steps for expansion of the IWN towards full national-scale deployment. During this expansion, improvements and additional features will be implemented to reflect lessons learned in the pilot, including:

- Guidance on harmonizing data appliance deployment with organizational IT policies
- Improved handling of QA/QC.
- Multiple-parameter, multiple-station visualization and download capability in Currents.
- Addition of sub-organizational contacts to metadata catalog and Currents discovery tool.
- Selection of stations in Currents using the API's point-and-buffer and upstream/downstream services.
- A mobile Currents application.

To align complementary efforts and promote interoperability, the next IWN phase will encompass

coordination and cooperation with other Federal agencies (e.g. USGS, NOAA) and academia (e.g. Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI)). Additionally, EPA hopes to engage with the private sector to encourage sensor and data management vendors to provide SOS and WaterML 2 access to data.

### Supplemental information

A description is provided here of the naming scheme for SOS objects on IWM data appliances. Basic recipes and other information for installing and configuring 52°North SOS and the pilot IWN ingestion script are provided at <https://github.com/IWN-Currents/OGD-materials>.

### IWN uniform resource name (URN) scheme

Uniform Resource Names (URNs) are used extensively to provide unique machine-readable identifiers for different entities represented in 52 N SOS-based data appliances deployed on behalf of Pilot partners. URNs were chosen instead of Uniform Resource Locators (URLs) to simplify data provider requirements by removing the need to provide resolvable endpoints on data appliances.

In general, URNs consist of the term “urn:” followed by a namespace ID and a namespace-specific string. The namespace ID is currently “x-epaiwpp”, so all URNs will begin with the text “urn:x-epaiwpp”. The organization, suborganization, station, and parameter IDs are specified in metadata files for the data appliance.

### Organizations, suborganization and station IDs

Every organization in the Pilot must have a unique name identifier. The name identifier will begin with either a two-letter state postal abbreviation or “US” for national-scale organizations:

- usepa – United States Environmental Protection Agency
- njdep – New Jersey Department of Environmental Protection
- njmeri – Meadowlands Environmental Research Institute (located in New Jersey)
- ohclec – Clermont County, Ohio – OR OH39025 (FIPS-BASED)

Organizations are assumed to have suborganizations such as:

- usepa:ord – Office of Research and Development (EPA)
- ohclec:wrp – Water Resources Division of Clermont County, Ohio
- njmeri:meri – no suborganization, organization acronym repeated

Station IDs are assigned by the organization, and consist of alphanumeric characters.

### Parameter IDs

Parameter IDs are used to uniquely identify observable properties, sensors, features, offerings and templates within 52 N SOS. The parameter IDs must be consistent across all organizations.

### Observable property URNs

Observable property URNs are consistent across the entire network, and consist of the namespace ID, followed by the classifier “parameter” and the parameter ID:

*urn:x-epaiwpp:parameter:temperature*

### Station, offering, sensor, feature, and template URNs

URNs for stations, sensor, features, and templates are created by concatenation:

- Station URNs identify platforms deployed for sensors, and consist of the namespace ID followed by the classifier “station”, and the organization, suborganization, and station IDs:

*urn:x-epaiwpp:station:ohcleccty:wrdefr34.8*

- Sensor URNs identify sensors deployed at a platform, and consist of the namespace ID followed by the classifier “sensor”, the organization, suborganization, and station IDs, a data quality status indicator (“raw”, “provisional” or “final”), and the sensor parameter:

*urn:x-epaiwpp:sensor:ohcleccty:wrdefr34.8:raw:temperature*

Offering, feature, and template URNs are structured similarly to the sensor URNs but use a different classifier:

*urn:x-epaiwpp:offering:ohcleccty:wrdefr34.8:raw:temperature*

*urn:x-epaiwpp:feature:ohcleccty:wrdefr34.8:raw:temperature*

*urn:x-epaiwpp:template:ohcleccty:wrdefr34.8:raw:temperature*

### Funding

The work reported on in this paper was funded by the U.S. Environmental Protection Agency (EPA). EPA staff participated in the work, providing substantial input on design specifications and reporting requirements, and also identified and coordinated with stakeholders in the pilot watersheds for this project. Additionally, EPA staff (D. Young and B. Dean) are co-authors.

### Authors’ contributions

All authors participated directly in the work reported on in this paper and contributed materially to the manuscript. Specifically: TS led the implementation of “data appliances.” He drafted the outline, abstract, results (data appliance) and methods, and coordinated synthesis, editing and submittal of the manuscript. KS was responsible for development of the Currents discovery tool and also led the elicitation of specifications and user stories from pilot watershed partners. She wrote the results (discovery tool) section of the manuscript and provided editorial review. BB was responsible for design and implementation of the metadata catalog and deployment of the catalog on EPA-supplied hardware. He was responsible for the writeup of the catalog in the manuscript. DY was the EPA project lead and also served as liaison to EPA and other federal agencies and offices with similar needs. He was responsible for the paper’s introduction and conclusions. BD supported D. Young in liaison roles, and also worked on the use of the EPA SRS as a basis for a standardized parameter set. She provided editorial review on the paper. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

### Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>LimnoTech, 501 Avis Drive, Ann Arbor, MI 48108, USA. <sup>2</sup>USEPA Headquarters, William Jefferson Clinton Building, 1200 Pennsylvania Avenue NW, 4503T, Washington, DC 20460, USA. <sup>3</sup>RTI International, 3040 East Cornwallis Road, Research Triangle Park, P.O. Box 12194, Durham, NC 27709-2194, USA.

Received: 23 December 2016 Accepted: 20 April 2017

Published online: 24 May 2017

### References

1. 52° North. “Sensor Observation Service”. 2016. <http://52north.org/communities/sensorweb/sos/>. Accessed 20 Dec 2016.
2. Advisory Committee on Water Information. “Open Water Data Initiative Overview”. 2014. <https://acwi.gov/spatial/owdi/>. Accessed 28 Mar 2017.
3. EPA. “Continuous Monitoring Data Sharing Strategy.” Prepared by Michael Baker International. Washington: LimnoTech and MapTech, Inc; 2015. under EPA Contract EP-C-12-052 Order No. 0005.
4. EPA. “What is STORET and how does it relate to WQX”. 2016a. <https://www.epa.gov/waterdata/frequent-questions-about-storage-and-retrieval-storet#101>. Accessed 20 Dec 2016.
5. EPA. “STORET/WQX: What is WQX?”. 2016b. <https://www.epa.gov/waterdata/frequent-questions-about-storage-and-retrieval-storet#103>. Accessed 20 Dec 2016.
6. EPA. “About Substance Registry Services”. 2016c. [https://iaspub.epa.gov/sor\\_internet/registry/substreg/home/overview/home.do](https://iaspub.epa.gov/sor_internet/registry/substreg/home/overview/home.do). Accessed 20 Dec 2016.
7. Open Geospatial Consortium. “OGC® Sensor Observation Service 2.0 Hydrology Profile”. 2014. [http://docs.opengeospatial.org/bp/14-004r1/14-004r1.html#requirement\\_1](http://docs.opengeospatial.org/bp/14-004r1/14-004r1.html#requirement_1). Accessed 28 March 2017.
8. Open Geospatial Consortium. “Sensor Observation Service”. 2016a. <http://www.opengeospatial.org/standards/sos>. Accessed 20 Dec 2016.
9. Open Geospatial Consortium. “OGC® WaterML”. 2016b. <http://www.opengeospatial.org/standards/waterml>. Accessed 20 Dec 2016.
10. Slawecki TAD, Young D, Perez B, McLellan P. “A Draft EPA Strategy for Sharing Continuous Monitoring Data”. In: Proceedings of the Water Environment Federation, WEFTEC 2015: Session 610 through Session 611. 2015. p. 5291–5303(13). doi:10.2175/193864715819522919.
11. USGS, EPA, USDA. “What is the WQP”. 2016. [http://www.waterqualitydata.us/wqp\\_description/](http://www.waterqualitydata.us/wqp_description/). Accessed 20 Dec.