Water Quality Portal

JIM KREFT,
USGS CENTER FOR INTEGRATED DATA ANALYTICS
NATIONAL WATER QUALITY MONITORING COUNCIL
WEBINAR
FEBRUARY 18, 2015
Purpose

• Describe the Water Quality Portal
• Explain the benefits of the Water Quality Portal
• Conduct a Demonstration of the Portal
• Demonstrate a series of different uses of the portal
Portal Background

The “New Jersey Problem”

- New Jersey data was collected in a cooperative agreement with USGS
- Data Collected by USGS was in NWIS
- Data Collected by NJ DEP was submitted to STORET
- Combining data was laborious and error-prone

National Water Quality Monitoring Council took on this challenge
Portal Background- 2003
Memo of Understanding

USGS and EPA will deliver data from USGS-NWIS and EPA-STORET in a common format to:

1. Analyze and report on the state of the nation's water environment
2. Provide a common basis for integrated water-quality analysis and protection
3. Provide an information base for scientific inquiry about water quality

An underlying goal is to ensure that the data from these important government databases are documented to describe their quality so that users can establish the utility and comparability of the data.
Benefits of the Portal

Reduces the effort to use other data sources
- Collecting data from multiple sources
- Combining into common format
- Delivering in a single file

Leverages and protects investments in monitoring data
- Common data elements emerging from monitoring community
- Marketplace of what, when and where for monitoring

Supports water quality based decision making
- Comparing to water quality standards
- Identifying hotspots
- Developing protection and restoration plans
- Modeling expected changes
Data are valuable, plan for re-use

Electronic data are more valuable than data in file cabinets

The more data are re-used, the more valuable they become
  ◦ Collect once – use multiple times

Shared data are of even higher value
  ◦ Provide for better planning and management decisions
  ◦ Incentivize collaborative efforts
  ◦ Make the most use of the data collection resources being invested
 Increases Access to Data

With the National Water Quality Monitoring Council (NWQMC), the Water Quality Portal (WQP) integrates publicly available water-quality data, through use of the Water Quality eXchange (WQX), from the USGS NWIS, EPA STORET, and USDA ARS STEWARDS.
Access to Water Quality Data

Over 250 million discrete water data records and 2.2 million stations
Access to Multiple Data Types

Portal Data Records (252m total)
- USGS NWIS – 90m records
- USDA STEWARDS – 1m records
- USEPA STORET – 162m records

Portal Data Contributors
- Federal – EPA, USGS, USACOE, NPS, USBR
- States and territories – 50 with 5 more in progress
- Tribes – 130 agencies
- Other organizations – county, watershed groups, academic

Portal Data types
- Physical/Chemical
- Biological (Coming soon)
- Habitat, Metrics, Indexes (Coming Soon)
Portal Usage continues to rise

Weekly Visits

12/30/13  2/28/14  4/30/14  6/30/14  8/31/14  10/31/14  12/31/14  2/28/15  4/30/15  6/30/15  8/31/15  10/31/15  12/31/15
Result Records Downloaded

6/1/14  7/1/14  8/1/14  9/1/14  10/1/14  11/1/14  12/1/14

Direct Call  WQP Site
Built using Responsive Design

Water Quality Portal works on any device with a modern web browser
Water Quality Portal Demo

Portal data holdings
- [http://waterqualitydata.us](http://waterqualitydata.us)
- Characteristics and query capabilities
- Display stations on a map
- Web Service usage example
Demo- Query Options

- Geospatial
  - Countries
  - States
  - Counties
  - Bounding box
  - Radius around location

- Site Metadata
  - Organization
  - Site type
  - Site ID

- Sampling information
  - Sample type
  - Characteristic Type
  - Characteristic Name

different file types

what does the “RESTlike Query” button do for you?
Demo- Data display and download options

Map Display
- Coming soon-
  - More informative styling
  - Web Map Service
  - Web Feature Service

Download options
- Tabular
- KML for sites
RESTLike Query

The key to advanced usage of the portal.

Every combination of the

```
Sites
http://www.waterqualitydata.us/Station/search?statecode=US%3A55&siteType=Stream&characteristicType=Nutrient&mimeType=csv&zip=yes

Sample results
http://www.waterqualitydata.us/Result/search?statecode=US%3A55&siteType=Stream&characteristicType=Nutrient&mimeType=csv&zip=yes
```
What is a client?

Bridge between the an API and an external tool.

Allows people familiar with a given software tool to easily get started without having to first figure out the WQP API

Water Quality Portal has two clients:

- DataRetrieval for R
- pywqp for Python
Water Quality Portal R Client

https://github.com/USGS-R/dataRetrieval

https://github.com/USGS-R/USGSHydroTools
Using R for data exploration

DataRetrieval

- Officially supported R client for WQP and NWIS data. On CRAN and Github

WQP_Data_Exploration

- suite of R-based data exploration examples for the WQP, available to all on Github
- [https://github.com/USGS-CIDA/WQP_Data_Exploration](https://github.com/USGS-CIDA/WQP_Data_Exploration)
What agencies have sampled for nutrients in the western Lake Erie basin?

<table>
<thead>
<tr>
<th>OrganizationFormalName</th>
<th>count</th>
</tr>
</thead>
<tbody>
<tr>
<td>USGS Ohio Water Science Center</td>
<td>427</td>
</tr>
<tr>
<td>USGS Michigan Water Science Center</td>
<td>351</td>
</tr>
<tr>
<td>Division of Surface water (Ohio)</td>
<td>221</td>
</tr>
<tr>
<td>Indiana STORET</td>
<td>201</td>
</tr>
<tr>
<td>USGS Indiana Water Science Center</td>
<td>107</td>
</tr>
<tr>
<td>Michigan Department of Environmental Quality</td>
<td>74</td>
</tr>
<tr>
<td>Division of Drinking and Ground Water (Ohio)</td>
<td>63</td>
</tr>
<tr>
<td>EPA National Aquatic Resources Survey</td>
<td>16</td>
</tr>
<tr>
<td>USDA Agricultural Research Service</td>
<td>16</td>
</tr>
<tr>
<td>IDEM</td>
<td>12</td>
</tr>
<tr>
<td>EPA National Aquatic Resource Survey Data</td>
<td>7</td>
</tr>
</tbody>
</table>
How many sites in the Western Lake Erie basin have been sampled for phosphorus?
What sites have more than 50 phosphorus samples, with at least 10 of those samples after 2000?
Change Two lines of Code:

data <- readWQPdata(huc="0410*", characteristicName="Phosphorus")

is changed to

data <- readWQPdata(huc="0207*", characteristicName="Phosphorus")

#sites <- whatWQPsites(huc="0410*", characteristicName="Phosphorus")

is changed to

sites <- whatWQPsites(huc="0207*", characteristicName="Phosphorus")
And we have changed to the Potomac
Potomac Organizations and sites that collect nutrient data
Water Quality Portal Python Client

https://github.com/USGS-CIDA/pywqp

\[ y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it} \]
Portal FY15 Activities

Development
- Identify additional data partners and sources
- Examine opportunities to improve data quality
- Serve biological data out of the Portal
- Improve mapping interface and possibly NHD based searches
- Discuss continuous data solution
- Increase visibility of data in the portal

Outreach
- Highlight new community tools
- Promote data sharing
- Attend meetings
- Conduct webinars
Questions

Charles Kovatch
Kovatch.Charles@epa.gov
202-566-0399

Jim Kreft
jkreft@usgs.gov
608-821-3919