

# Water Quality Portal

JIM KREFT,

USGS OFFICE OF WATER INFORMATION

NATIONAL WATER QUALITY MONITORING CONFERENCE

MAY 4, 2016

---



# Purpose

---

- Describe the Water Quality Portal
- Explain the benefits of the Water Quality Portal
- Show what's new on 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.

# Data are valuable, don't just use them once!

---

Electronic data are more valuable than data in file cabinets

The more data are re-used, the more valuable they become

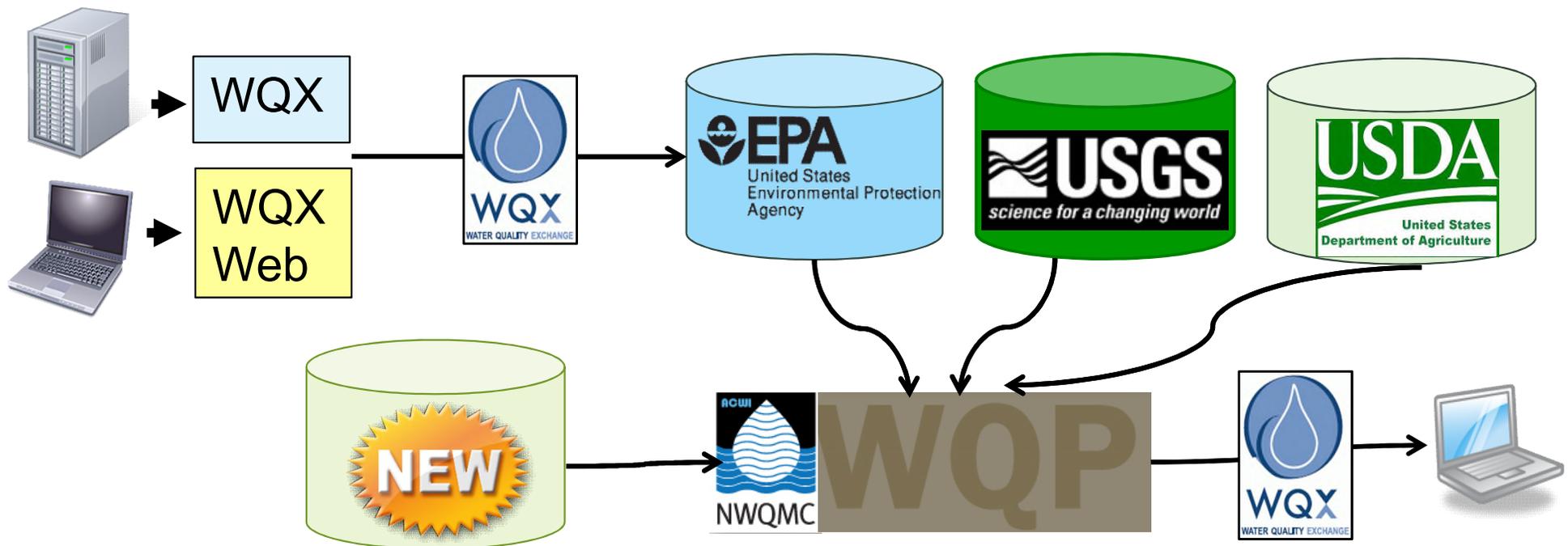
- Collect once – use multiple times

Shared data are of even higher value

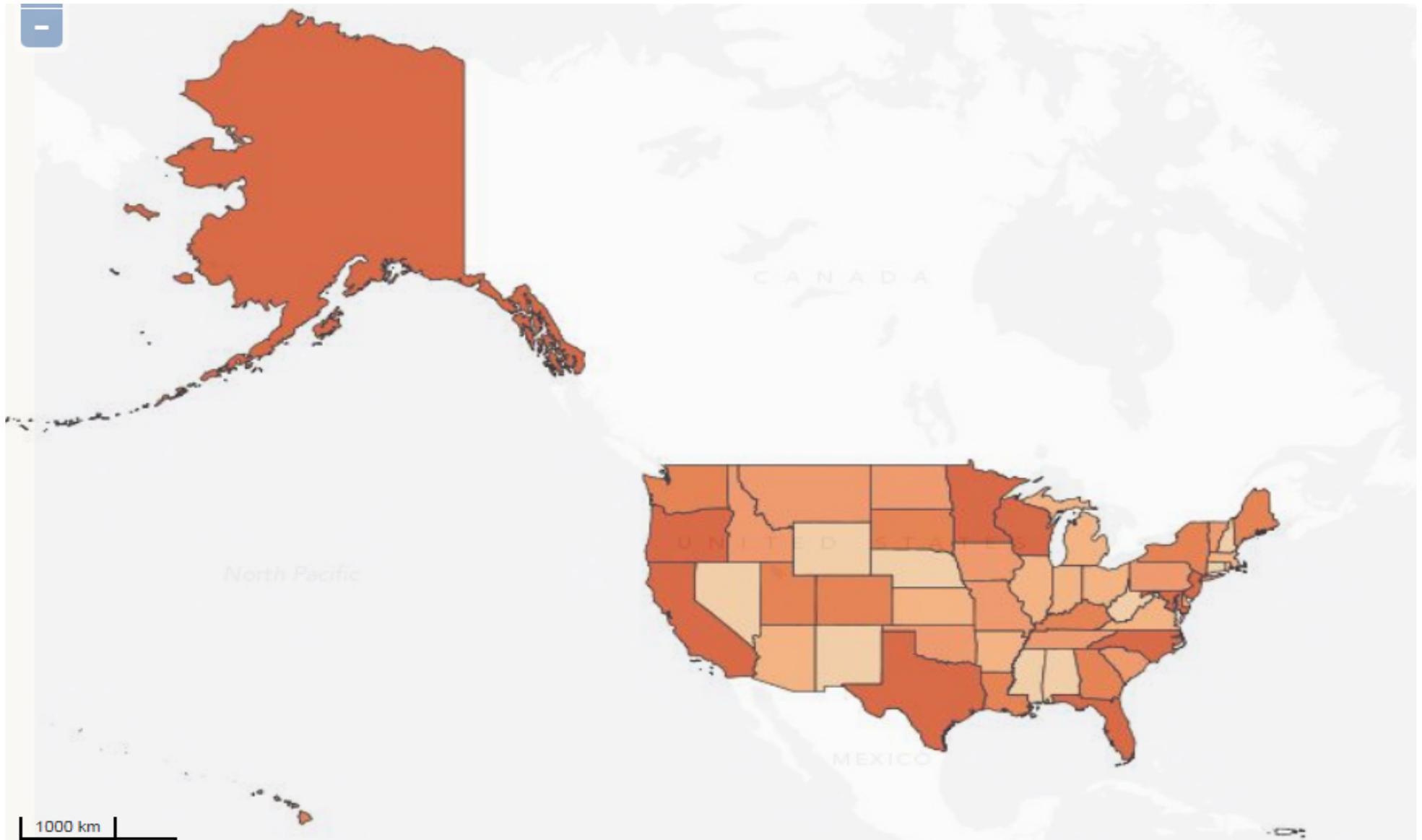
- Provide for better planning and management decisions
- Incentivize collaborative efforts
- Sharing data makes the most use of the data collection resources being invested

# 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 275 million discrete water data records and 2.5 million stations

# Access to Multiple Data Types

---

## Portal Data Records (272m total)

- USGS NWIS – 90m records
- USDA STEWARDS – 1m records
- USEPA STORET – 162m records

## Portal Data Contributors

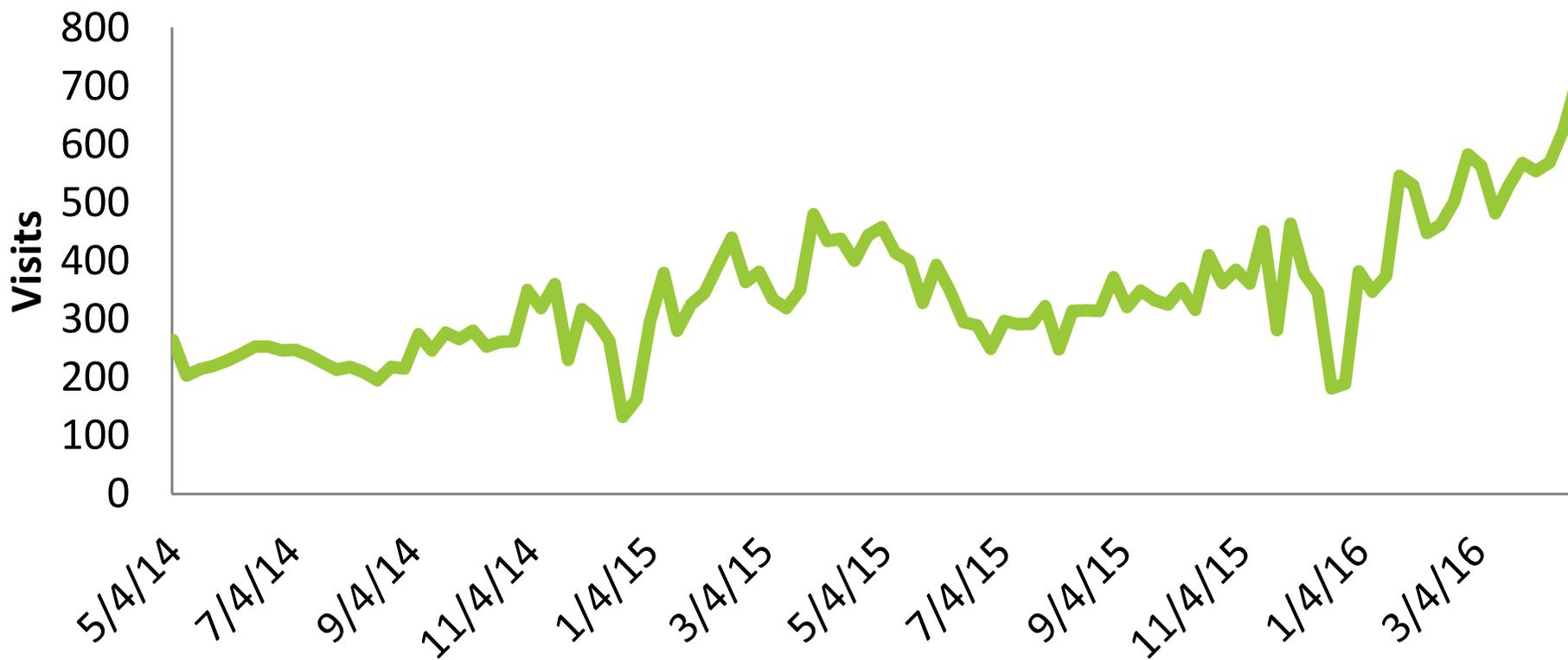
- Federal – EPA, USGS, USACOE, NPS, USBR
- States and territories – 50 with 5 more in progress
- Tribes – 130 agencies
- Other organizations – county, watershed groups, academic

## Portal Data types

- Physical/Chemical
- Biological Collections
- Habitat, Metrics, Indexes (Coming Soon)

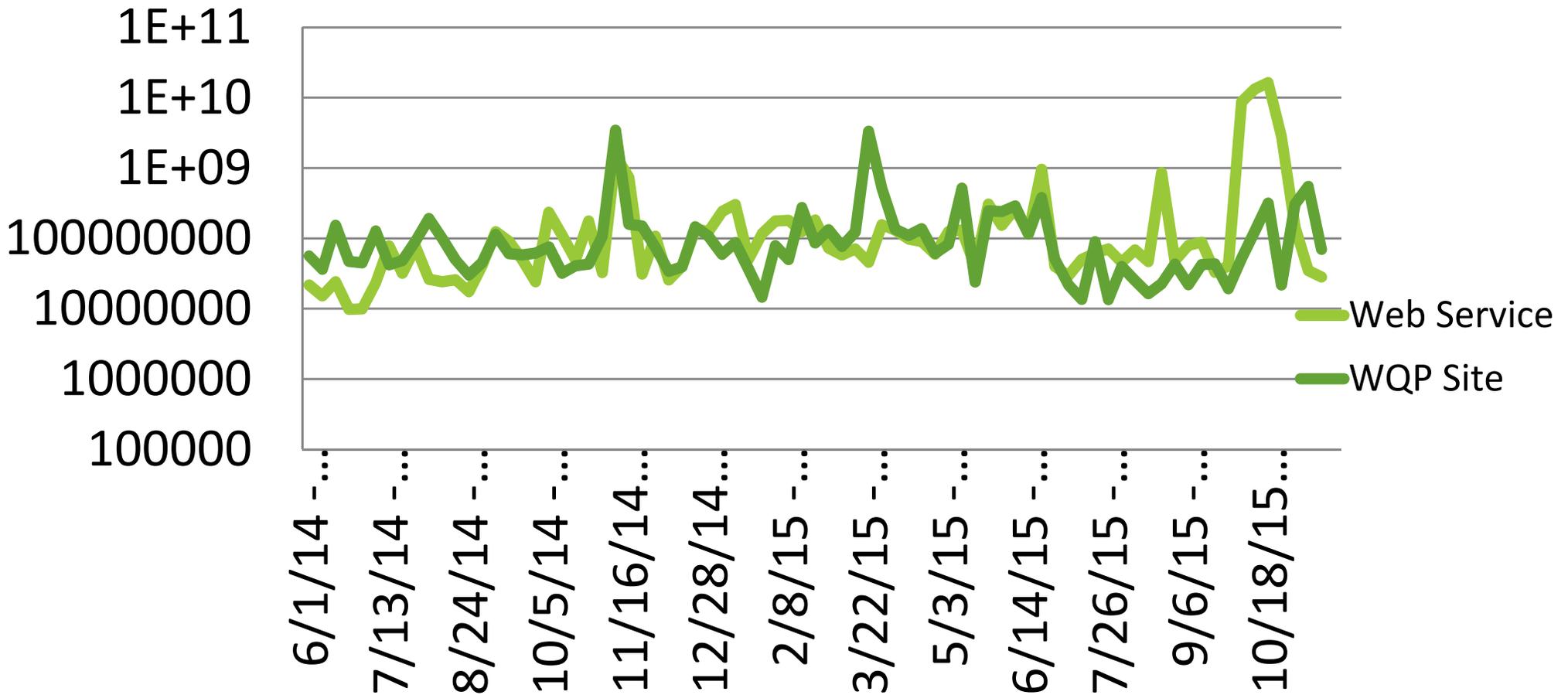
# Portal Usage continues to rise

**Weekly Visits to the WQP**  
May 2014 - May 2016

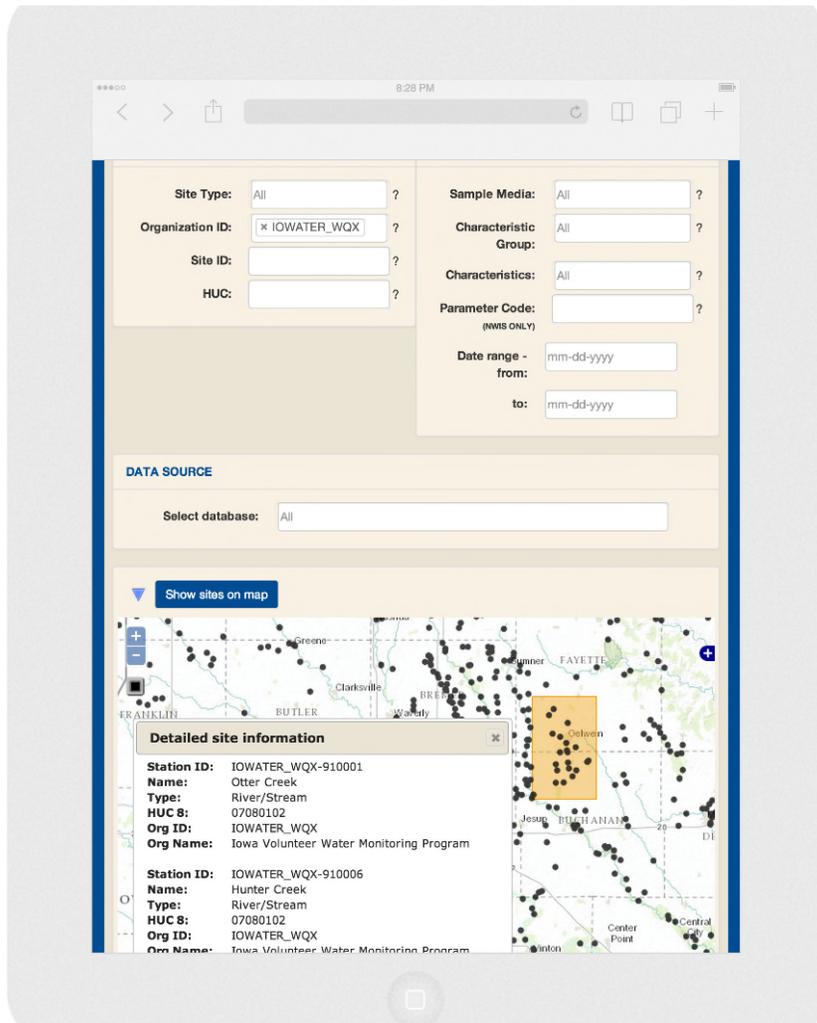


# Web services are key

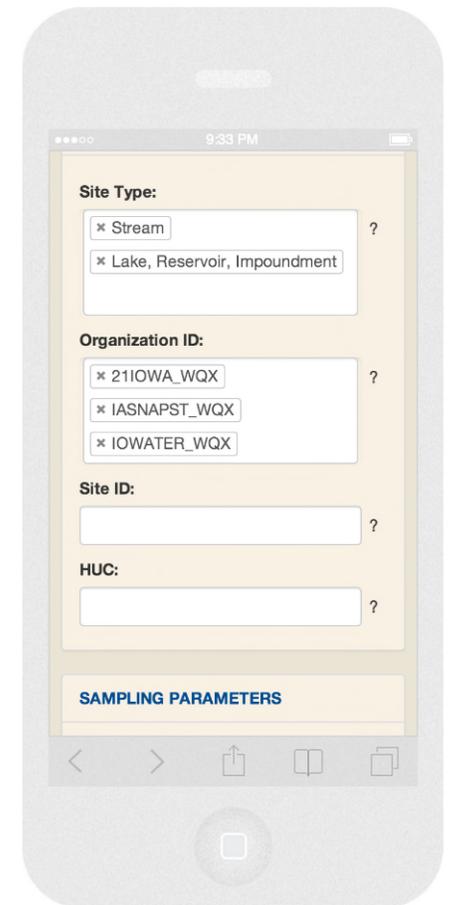
## Weekly records downloaded from WQP



# Built using Responsive Design



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



# WQP Portal Page

**LOCATION**

**Place:**

Country:  ?

State:  x ?

County:  x ?

**Point Location: ?**

Within  miles of

Lat:

Long:

**Bounding Box: ?**

North:

South:

East:

West:

---

**SITE PARAMETERS**

Site Type:  ?

Organization ID:  ?

Site ID:  ?

HUC:  ?

**SAMPLING PARAMETERS**

Sample Media:  ?

Characteristic Group:  x ?

Characteristics:  ?

Project ID:  ?

Parameter Code: (NWIS ONLY)  ?

Minimum results per site:  ?

Date range - from:  to:

**Biological sampling parameters: ?**

Assemblage:  ?

Taxonomic Name:  ?

---

**DATA SOURCE**

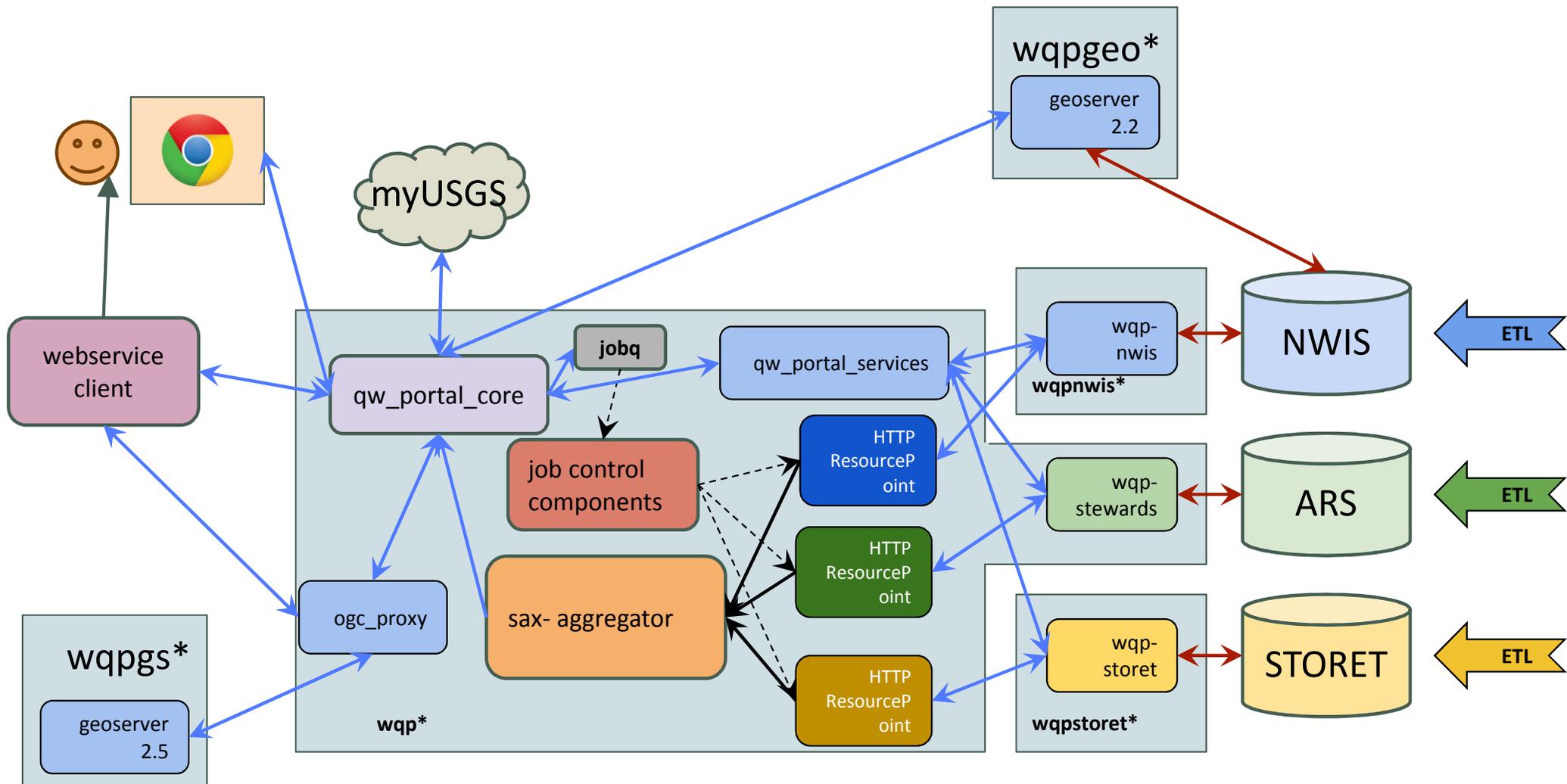
Select database:

# What have we done recently?

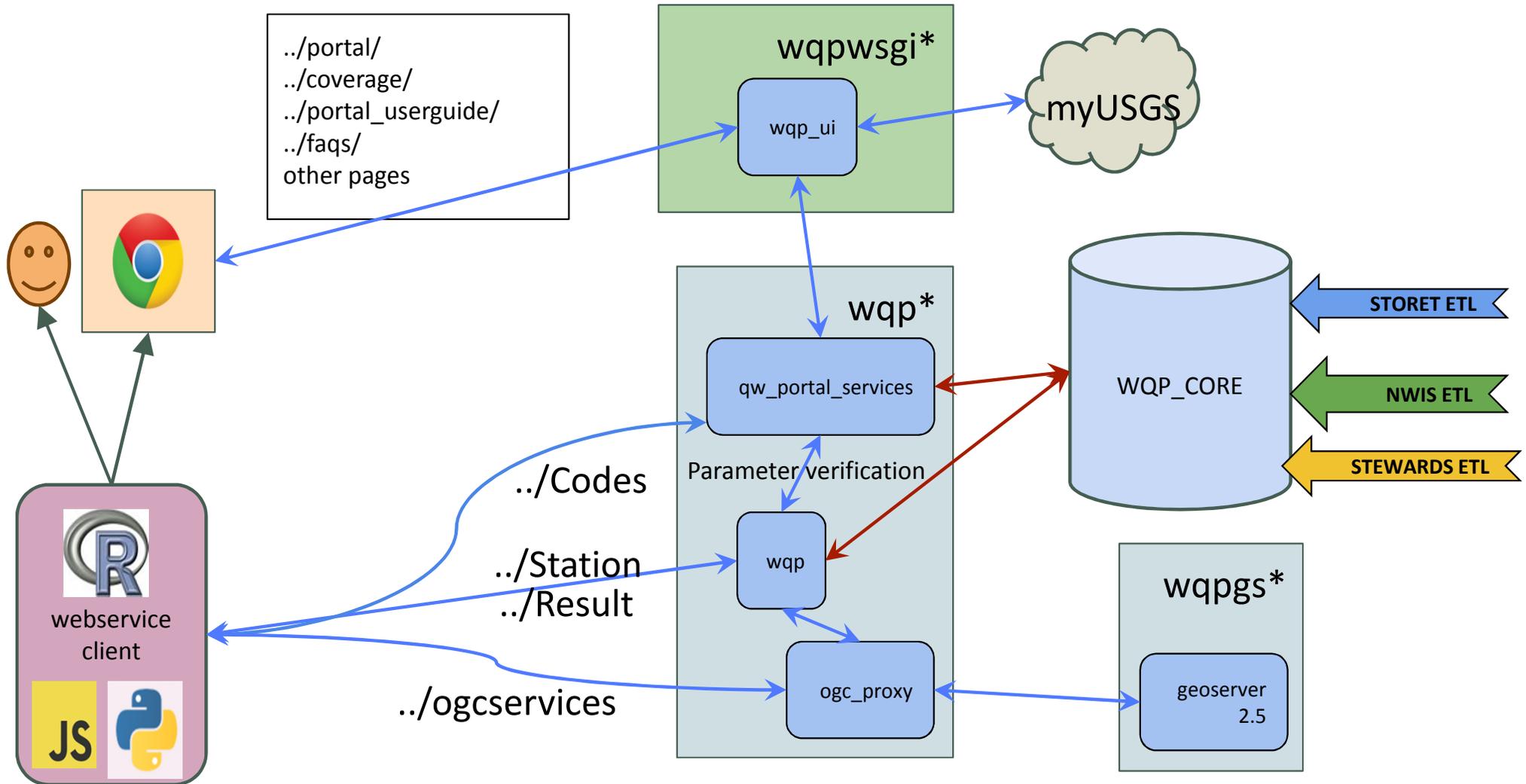
---

- Completely revamped underlying services
- Added Biological data
- Added new query parameters
- Searchable site pages
- Upstream-downstream- coming soon!

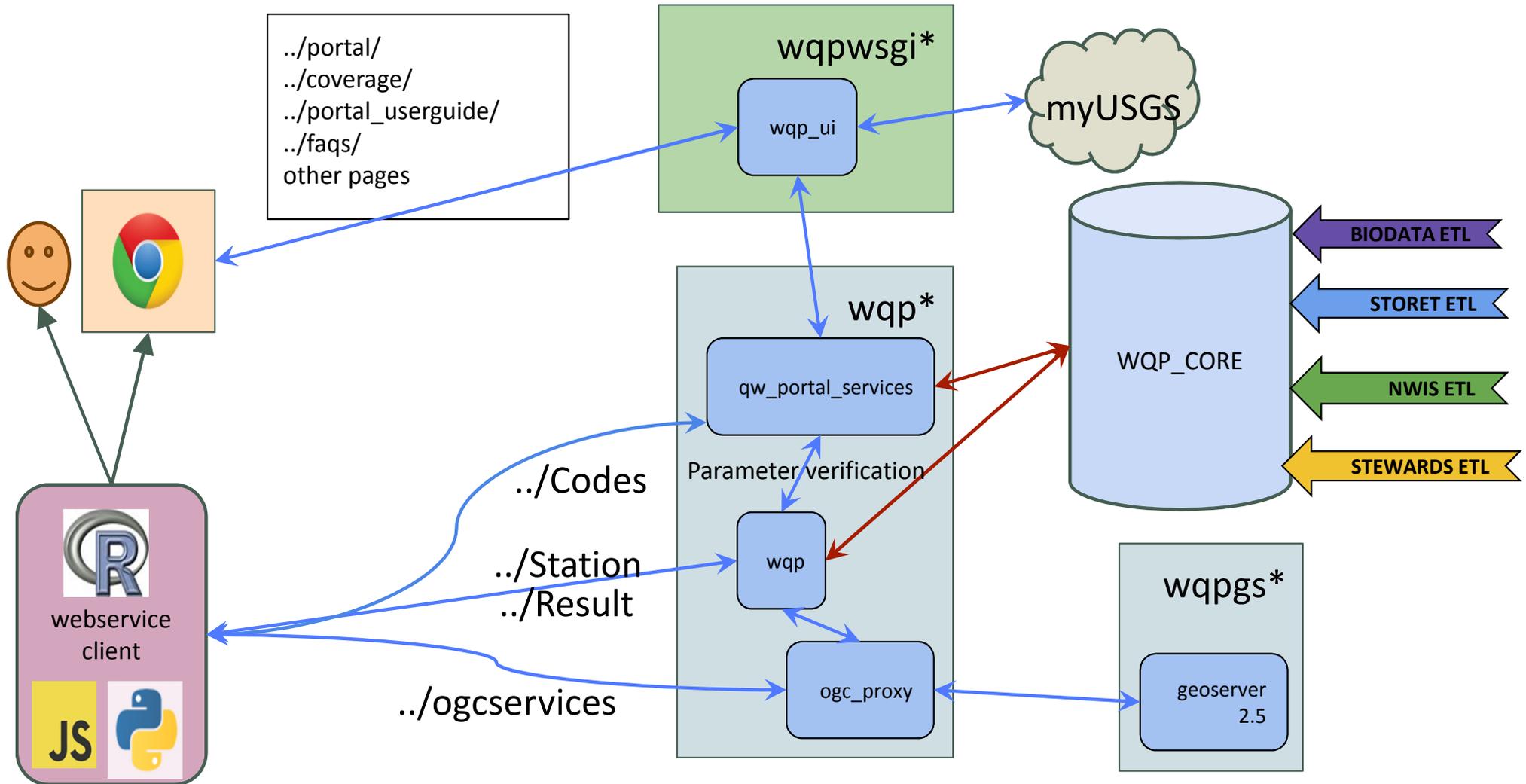
# Architecture Before



# Architecture After



# Adding Biodata was trivial

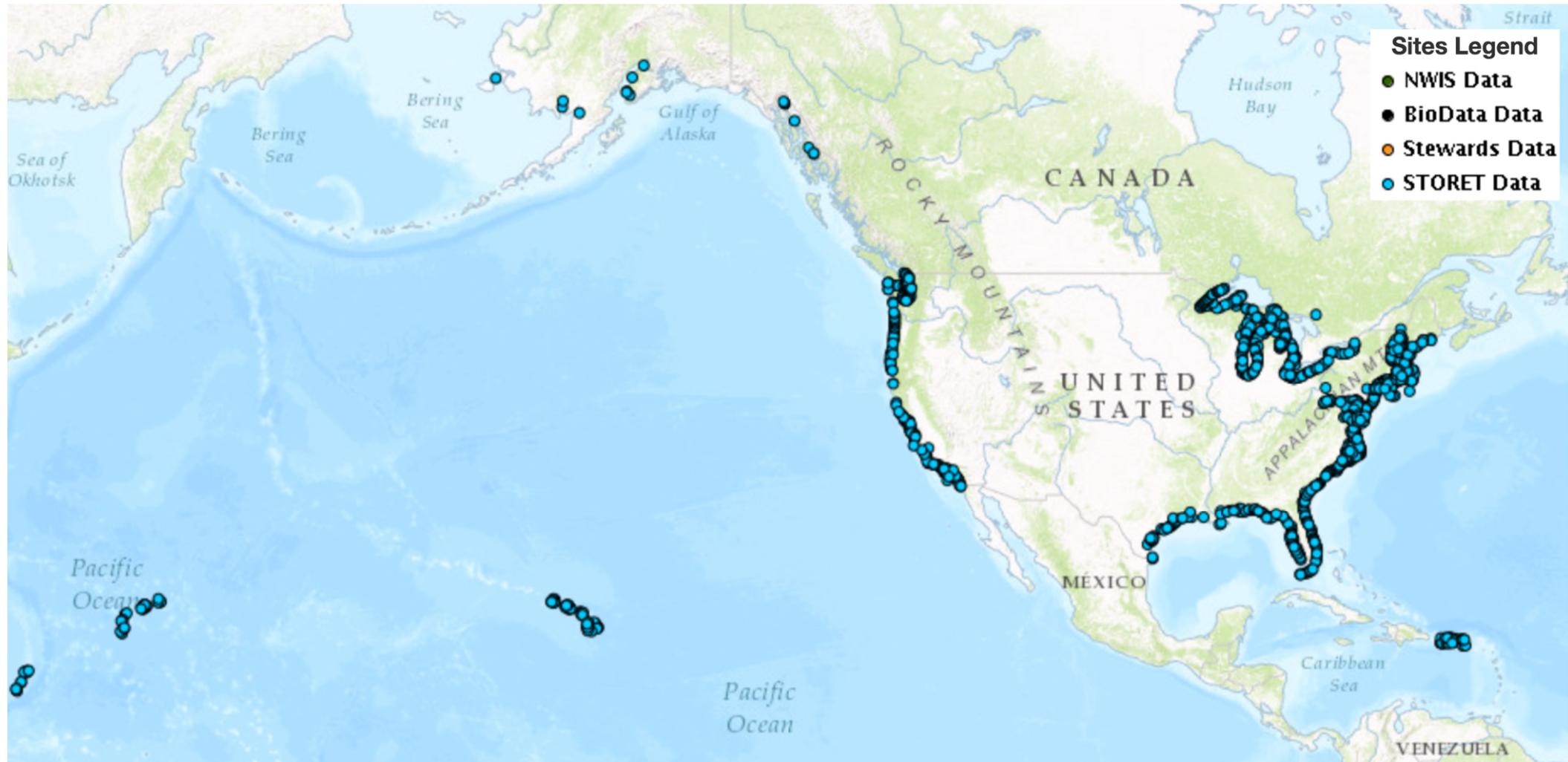


# Additional Query Parameters

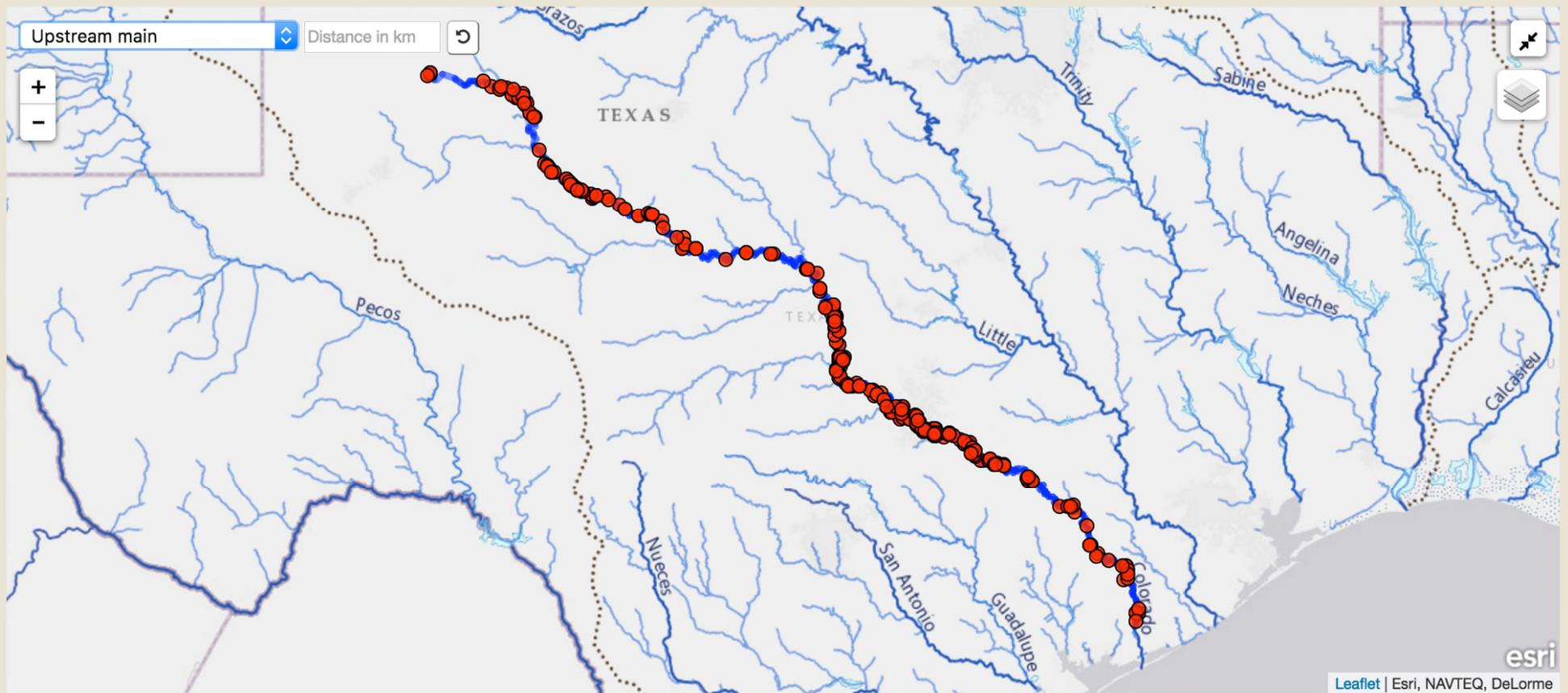
---

- Sampling Parameters
  - Project ID
  - Biological Parameters
    - Assemblage
    - Taxonomic Name
- Site Parameters
  - Coming soon!
    - Upstream-downstream queries

# ProjectID: EPABEACH



# Upstream-downstream queries



DATA SOURCE

# Bringing WQP Content into the light

---

- Water Quality Portal is part of the “Deep Web”
- Only can get WQP data if you know it is there
- Goal:
  - Expose Water Quality Portal content to search engines
  - Make data findable to both humans and robots

<http://www.waterqualitydata.us/provider/>

# Bringing WQP Content into the light

---

[WQP Home](#) > [Providers](#)

## Water Quality Portal Data Providers

Learn more about Water Quality Portal Data Providers [here](#).

[BIODATA](#)

[NWIS](#)

[STEWARDS](#)

[STORET](#)

<http://www.waterqualitydata.us/provider/>

# Bringing WQP Content into the light

---

[WQP Home](#) > [Providers](#) > [STORET](#)

## Water Quality Portal Data Organizations for STORET

[Clear Creek Superfund \(Colorado\) - 0800257](#)

[Clear Creek Superfund - 0800257\\_WQX](#)

[Ogden Railyard \(US EPA Region 8\) - 0800597](#)

[International Smelter \(US EPA Region 8\) - 0800650](#)

[Mystery Bridge Road - US Highway 20 - 0800852](#)

[Summitville Superfund site \(US EPA Region 8\) - 0801194](#)

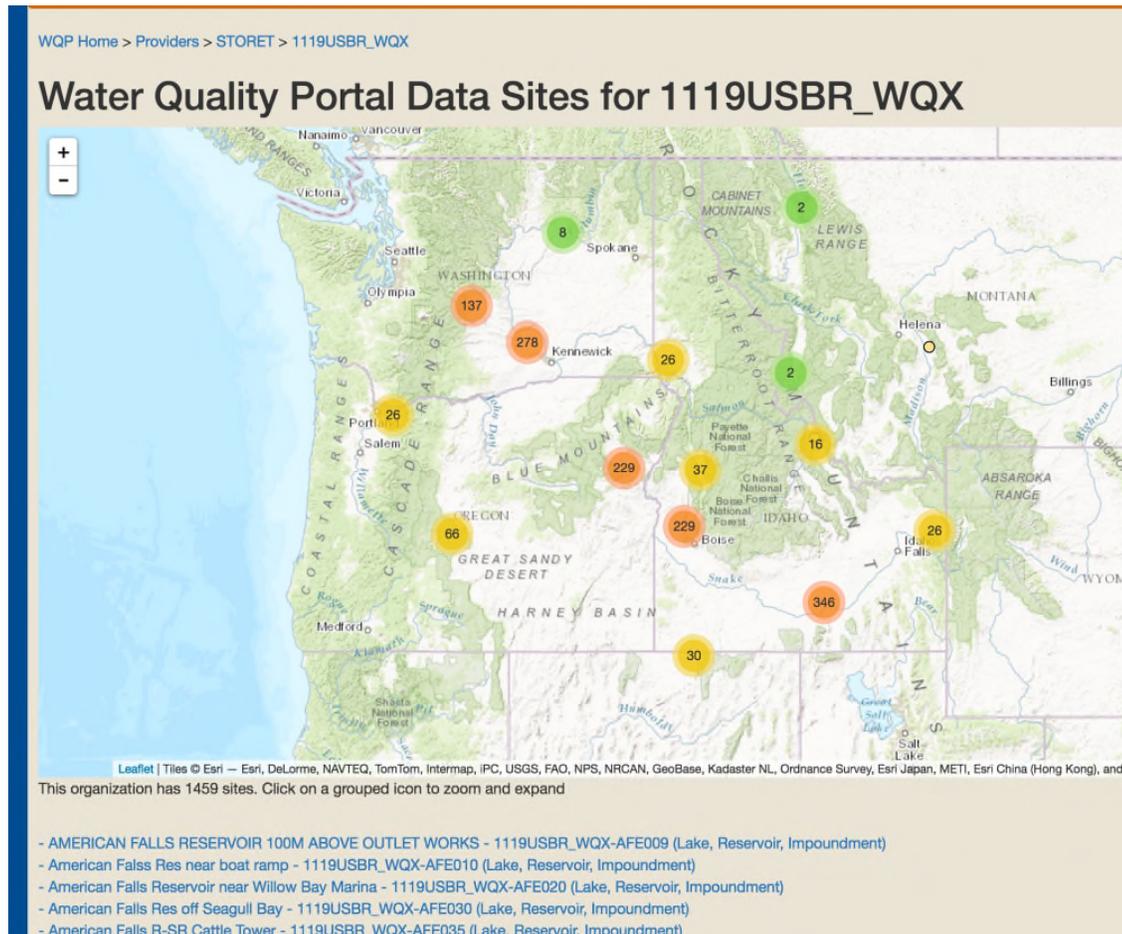
[Red Mountain Pass Zinc \(US EPA Region 8\) - 0801417](#)

[California Gulch \(US EPA Region 8\) - 0801478](#)

[French Gulch Superfund site \(US EPA Region 8\) - 0801505](#)

<http://www.waterqualitydata.us/provider/STORET>

# Bringing WQP Content into the light



[http://www.waterqualitydata.us/provider/STORET/1119USBR\\_WQX/](http://www.waterqualitydata.us/provider/STORET/1119USBR_WQX/)

# Bringing WQP Content into the light

The screenshot shows the Water Quality Portal interface. At the top, there is a navigation bar with links: WQP Home, Download Data, How to use the WQP, National Results Coverage, and About the WQP. Below the navigation bar, the breadcrumb trail reads: WQP Home > Providers > STORET > 1119USBR\_WQX > 1119USBR\_WQX-AFE020. The main heading is "American Falls Reservoir near Willow Bay Marina (1119USBR\_WQX-AFE020) site data in the Water Quality Portal". Below the heading, it states "Data Provider: STORET" with a link to "Learn more about Water Quality Portal Data Providers". The text describes the site's location, watershed, and provides links to EPA and USGS resources. It also includes coordinates and a note about metadata availability. A section titled "Detailed Metadata" includes a link to the "user guide" and a table of metadata elements.

MonitoringLocationTypeName	Reservoir
ProviderName	STORET
HorizontalCoordinateReferenceSystemDatumName	UNKWN
StateCode	16
MonitoringLocationName	American Falls Reservoir near Willow Bay Marina

[http://www.waterqualitydata.us/provider/STORET/1119USBR\\_WQX/1119USBR\\_WQX-AFE020/](http://www.waterqualitydata.us/provider/STORET/1119USBR_WQX/1119USBR_WQX-AFE020/)

# Why expose data like this?



amnicon river water quality



All

Maps

Images

News

Shopping

More ▾

Search tools

About 2,580 results (0.51 seconds)

**AMNICON RIVER NEAR PETZAU, WI (USGS-04024570 ...**

[www.waterqualitydata.us/provider/NWIS/USGS-WI/USGS-04024570/](https://www.waterqualitydata.us/provider/NWIS/USGS-WI/USGS-04024570/) ▾

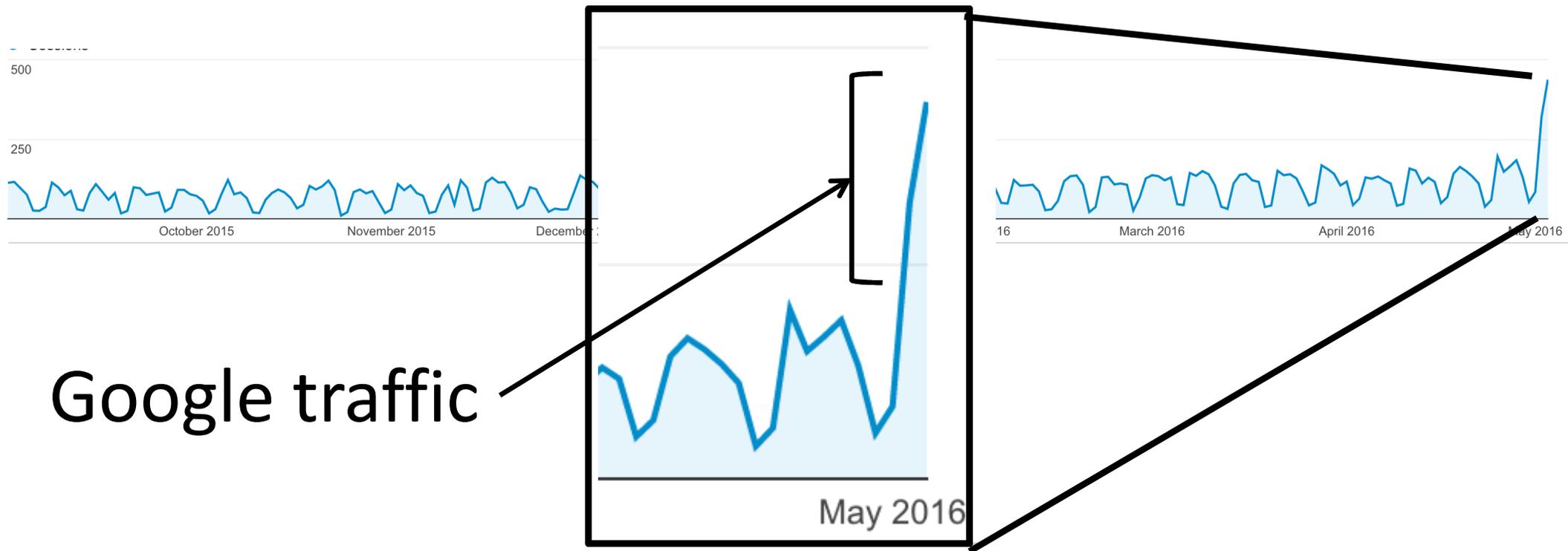
AMNICON RIVER NEAR PETZAU, WI (USGS-04024570) site data in the **Water Quality** Portal. Data Provider: NWIS (Learn more about **Water Quality** Portal Data ...

**AMNICON RIVER NEAR MOUTH NEAR POPLAR, WI ...**

[www.waterqualitydata.us/provider/NWIS/USGS-WI/USGS-04025033/](https://www.waterqualitydata.us/provider/NWIS/USGS-WI/USGS-04025033/) ▾

AMNICON RIVER NEAR MOUTH NEAR POPLAR, WI (USGS-04025033) site data in the **Water Quality** Portal. Data Provider: NWIS (Learn more about Water ...

# Indexing works!



# Web Service Calls

Show Web Service Calls

?

## Sites

```
http://www.waterqualitydata.us/Station/search?statecode=US%3A55&countycode=US%3A55%3A025&characteristicType=Nutrient&mimeType=csv&zip=yes&sorted=no
```

## Results

```
http://www.waterqualitydata.us/Result/search?statecode=US%3A55&countycode=US%3A55%3A025&characteristicType=Nutrient&mimeType=csv&zip=yes&sorted=no
```

## WFS GetFeature

```
http://www.waterqualitydata.us/ogcservices/wfs/?  
request=GetFeature&service=wfs&version=1.1.0&typeName=wqp_sites&searchParams=statecode%3AUS%3A55%3Bcountycode%3AUS%3A55%3A025%3BcharacteristicType%3ANutrient%3Bsorted%3Ano&outputFormat=application%2Fjson
```

# 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>

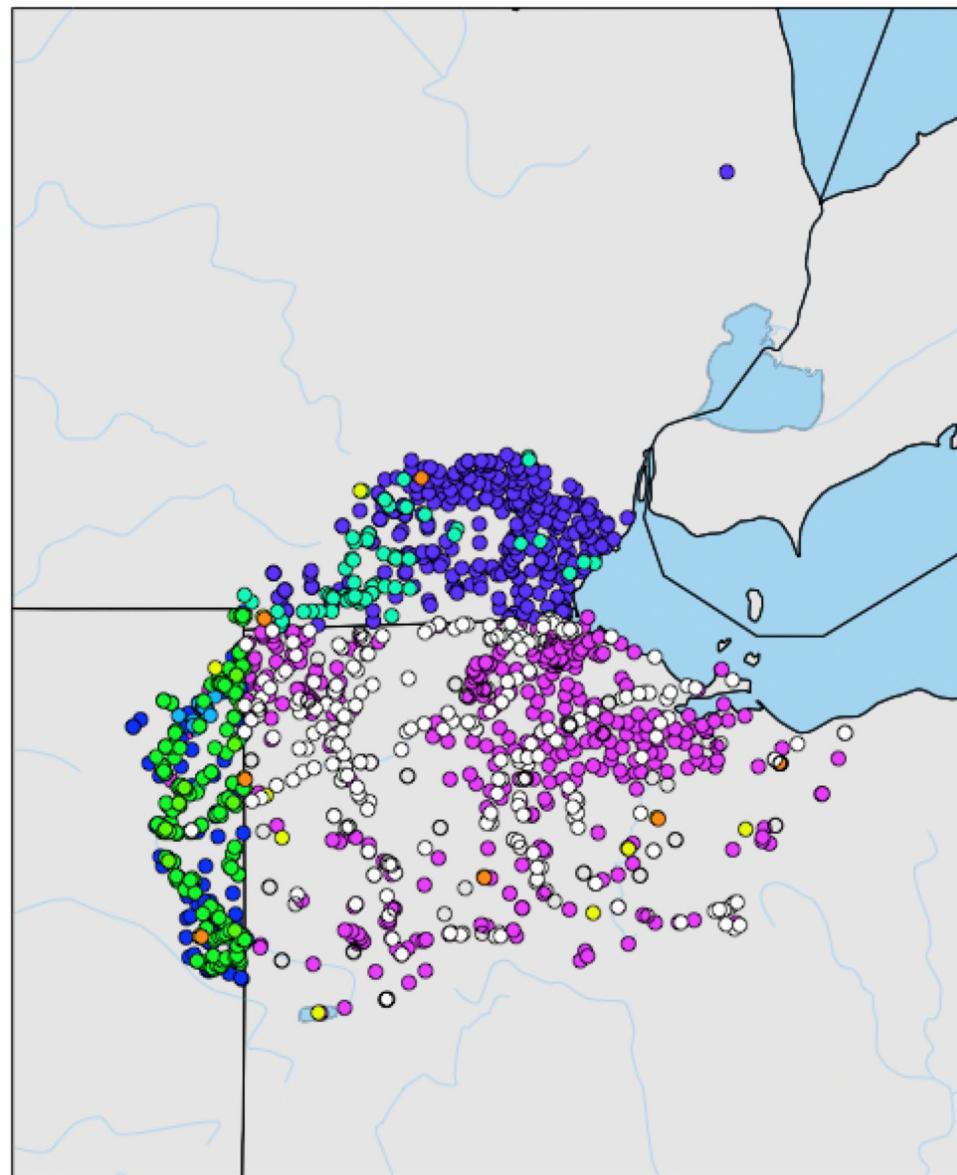


The screenshot shows the GitHub repository page for USGS-R/dataRetrieval. The page includes navigation links for Personal, Open source, Business, Explore, Pricing, Blog, and Support. It also features a search bar and buttons for Sign in and Sign up. The repository name is USGS-R / dataRetrieval, with 19 watches, 37 stars, and 21 forks. The page shows 17 issues, 0 pull requests, and 0 wiki pages. The description states: "This R package is designed to obtain its water quality sample data, streamflow data, and metadata directly from either the USGS or EPA, as well as user-supplied text files." The repository has 1,171 commits, 2 branches, 34 releases, and 6 contributors. The current branch is master. A merge pull request #213 from Idecicco-USGS/master is shown, with the latest commit 485c1b1 from 13 days ago. The file list includes:

File	Description	Last Commit
R	More url stuff	13 days ago
inst	Rebuild	5 months ago
man	Committing the man files!	a month ago
tests	Fix tests	3 months ago
vignettes	Rebuilt vignette.	6 months ago
.Rbuildignore	Update NEWS and version	2 months ago
.Rinstignore	Switching back to non-demo.	2 years ago
.travis.yml	Going back to c.	a year ago
DESCRIPTION	Version update	13 days ago
LICENSE.note	Re doing license with CC0.	2 years ago
NAMESPACE	More attempts to speed things up.	2 months ago
NEWS	Update NEWS and version	2 months ago
README.md	Keep that information in NEWS	14 days ago
appveyor.yml	Adding environment variable to appveyor.	a year ago

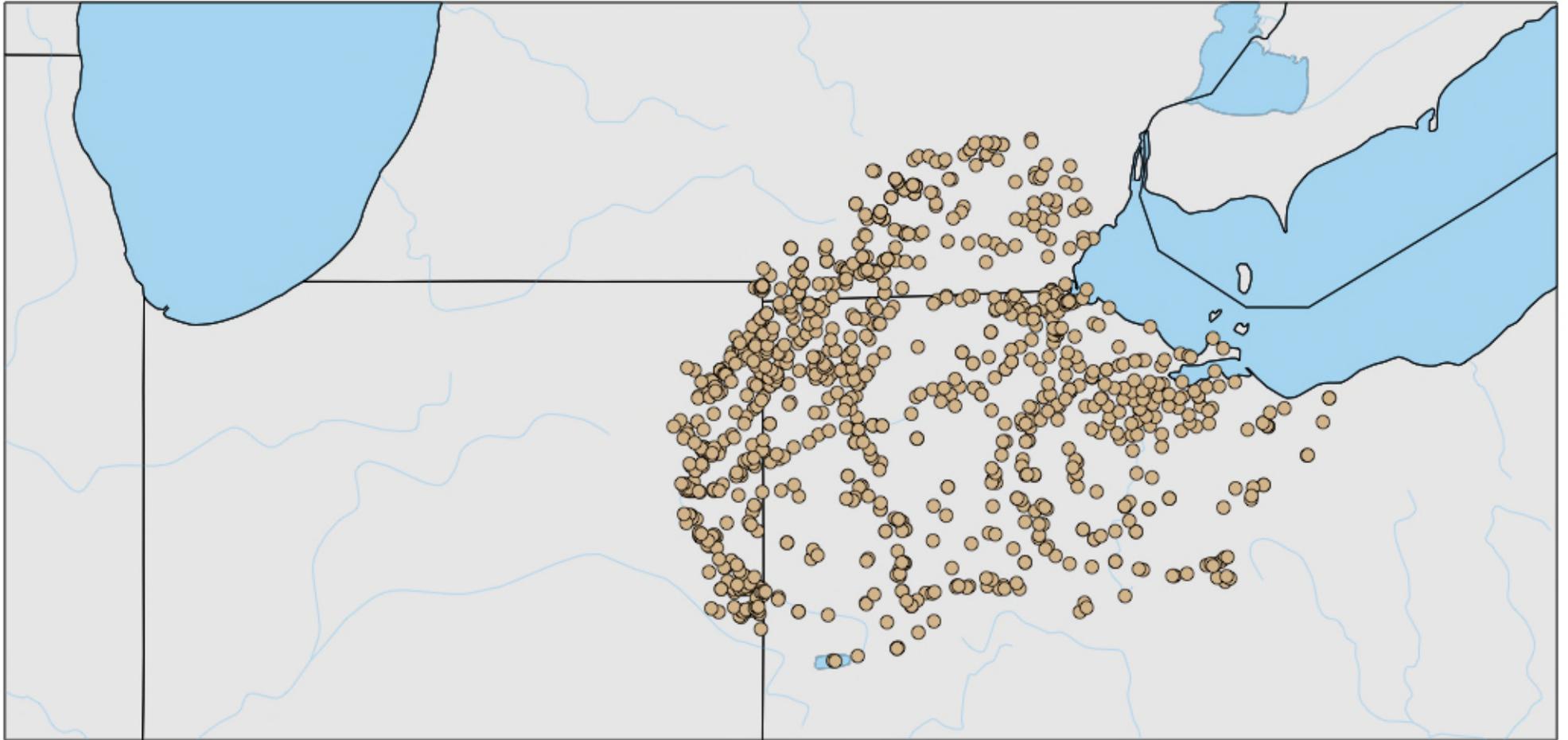
# What agencies have sampled for nutrients in the western Lake Erie basin?

	OrganizationFormalName	count
1	USGS Ohio Water Science Center	427
2	USGS Michigan Water Science Center	351
3	Division of Surface water (Ohio)	221
4	Indiana STORET	201
5	USGS Indiana Water Science Center	107
6	Michigan Department of Environmental Quality	74
7	Division of Drinking and Ground Water (Ohio)	63
8	EPA National Aquatic Resources Survey	16
9	USDA Agricultural Research Service	16
10	IDEM	12
11	EPA National Aquatic Resource Survey Data	7



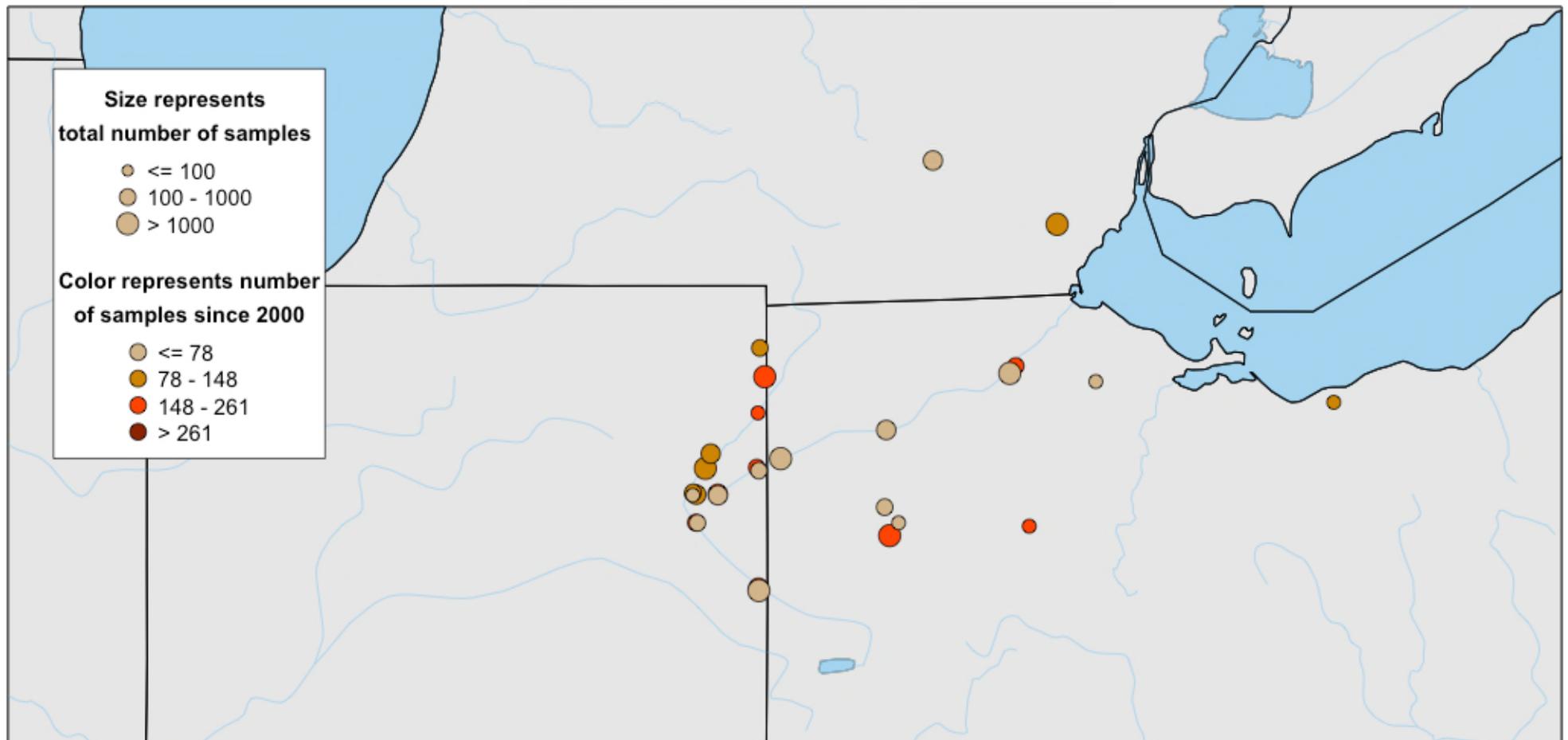
# How many sites in the Western Lake Erie basin have been sampled for phosphorus?

---



# 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

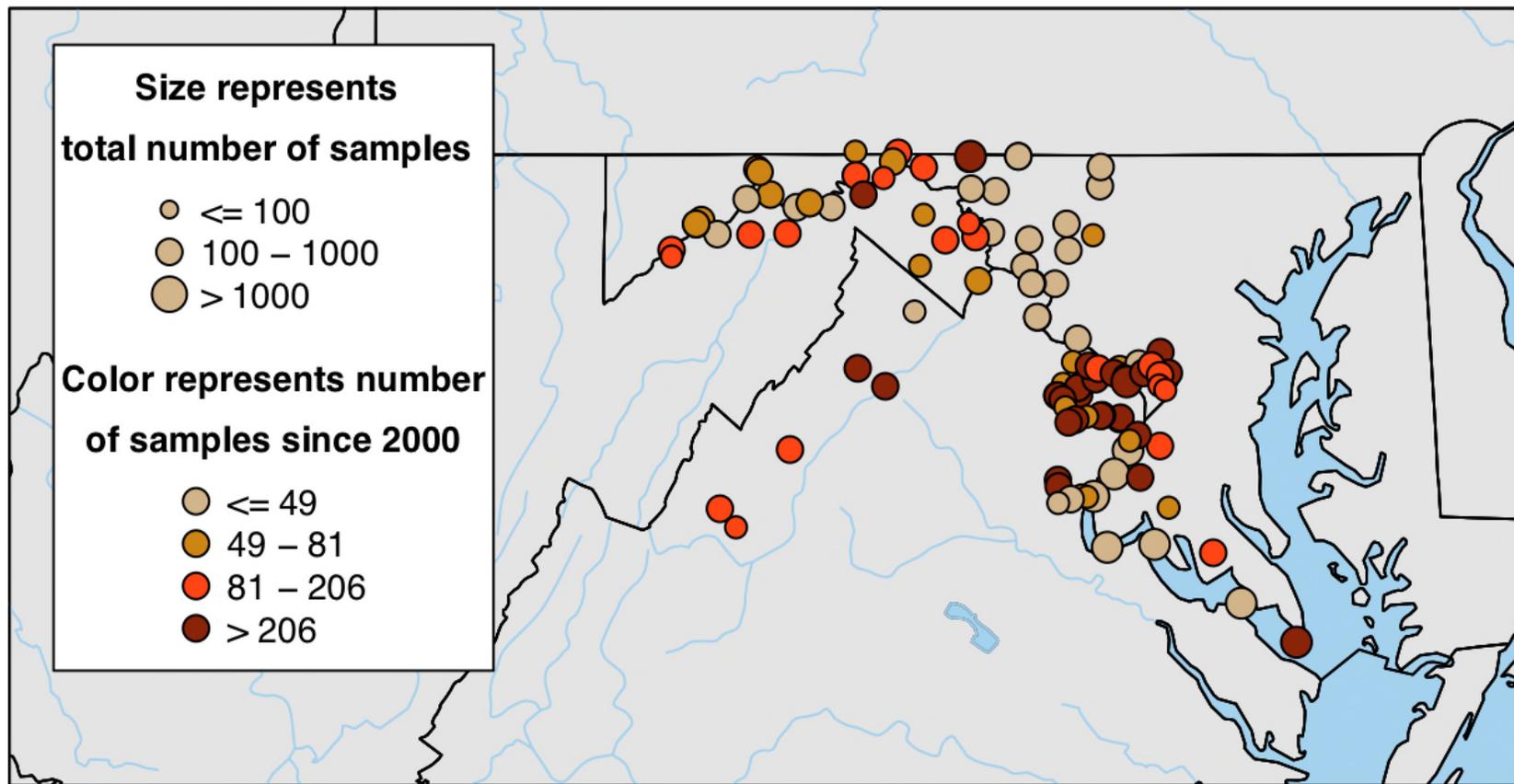
```
data <- readWQPdata(huc="0207*",  
characteristicName="Phosphorus")
```

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

is changed to

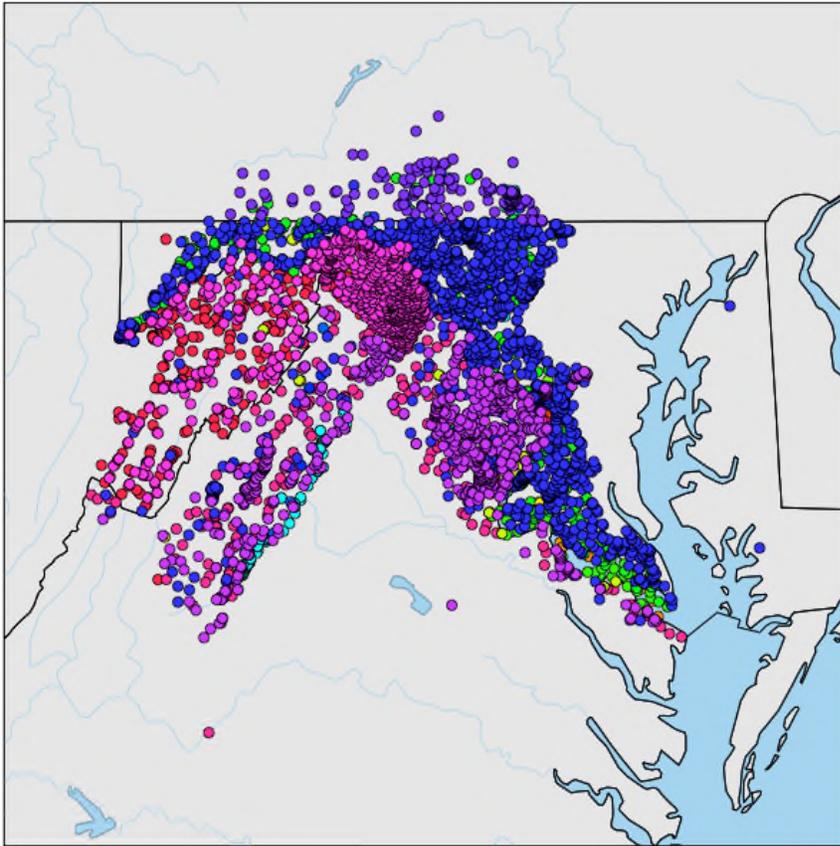
```
sites <- whatWQPsites(huc="0207*",  
characteristicName="Phosphorus")
```

# And we have changed to the Potomac



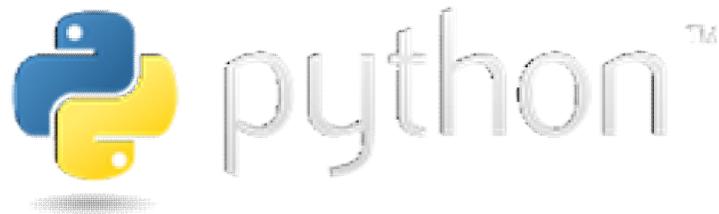
# Potomac Organizations and sites that collect nutrient data

---

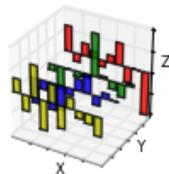
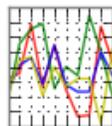
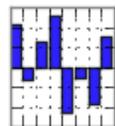


# Water Quality Portal Python Client

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



pandas  
 $y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$



This repository Search Explore Gist Blog Help jkreft-

USGS-CIDA / pywqp Unwatch 13

A generic scriptable Python client for downloading datasets from the USGS/EPA Water Quality Portal: an alternative to manual use of the website at <http://www.waterqualitydata.us>.

83 commits 1 branch 0 releases 5 contributors

branch: master pywqp / +

add documentation for wxq\_mappings

whblondeau authored 16 days ago latest commit: e53b7c4cfc

File	Commit Message	Time Ago
.idea	fixed references in the fileformat section, cleaned up formatting to ...	2 months ago
documentation	mass restructuring following lettuce BDD implementation	2 months ago
pywqp	manual merge with origin master	26 days ago
tests	steps to support dataframe.features	22 days ago
.gitignore	Add LICENSE, edit README and .gitignore	22 days ago
LICENSE	add hard line breaks	22 days ago
README.md	add documentation for wxq_mappings	16 days ago
dev_setup.sh	Added virtualenv, pip install setup script.	2 months ago
requirements.txt	Added virtualenv, pip install setup script.	2 months ago
setup.py	manual merge with origin master	26 days ago

README.md

## pywqp

A generic scriptable Python client for downloading datasets from the Web Services offered by the USGS/EPA Water Quality Portal: an alternative to manual use of the [WQP website](#).

# Portal FY16 Activities

---

## Development

- Identify additional data partners and sources
- Examine opportunities to improve data quality
- Discuss continuous data solution

## Outreach and planning

- 5 year Strategic Planning exercise
- Highlight new community tools
- Promote data sharing
- Attend meetings
- Conduct webinars

# Mission and Vision

---

## Vision

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

- Vision: what is our desired end state? What are the major issues or problems? What would success look like?

## Mission

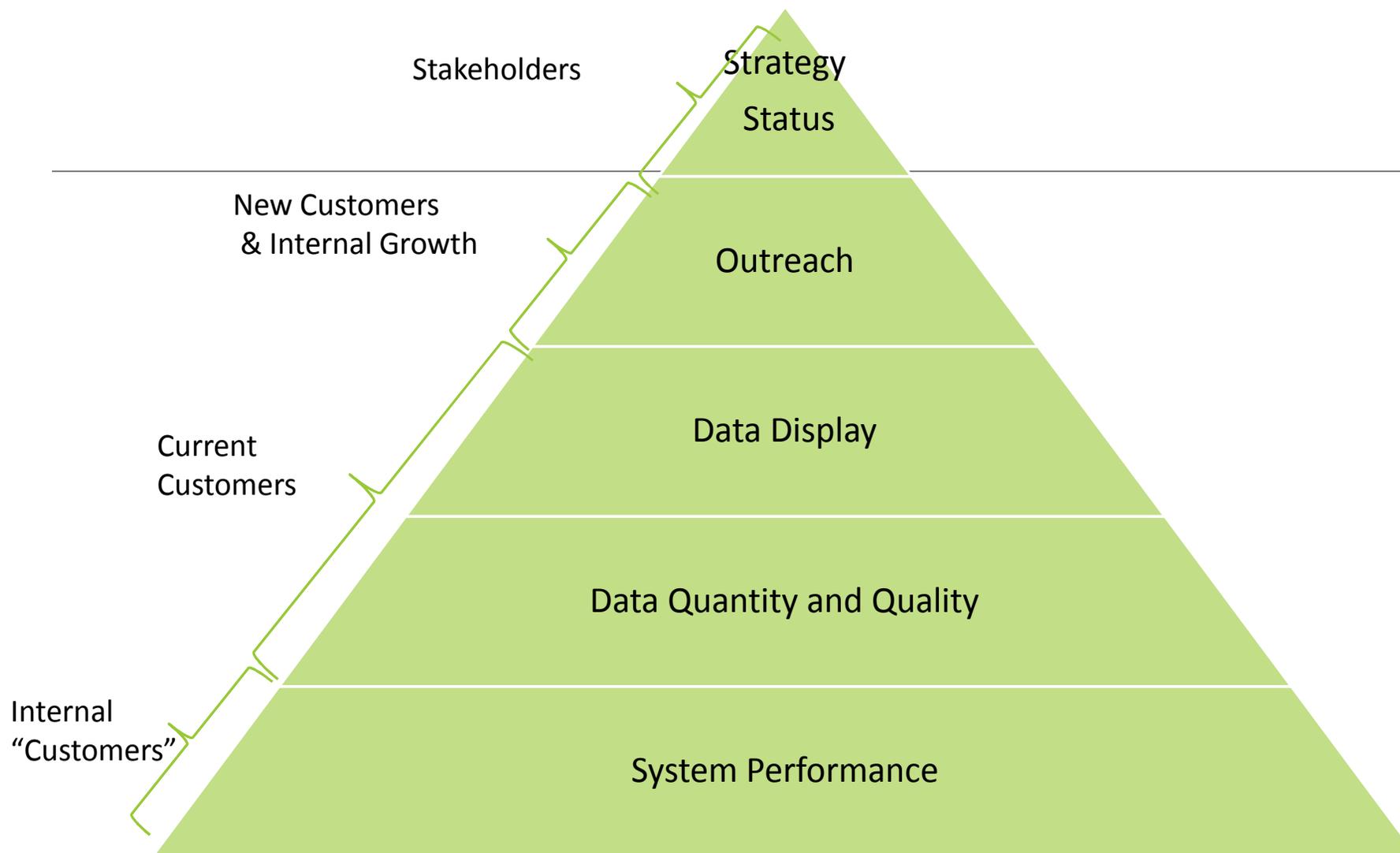
Provide easy access to all water quality data, facilitate improvements in data quality, and enhance data discovery and data summaries to inform sound water-quality decision making at local, state, regional, and national scales.

- Mission: Why we exist, Who are we, what do we do, why do we do this work and for whom?

## Scope

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

- Scope: What types of data, and where, will we focus on?



Relationship of 5 Portal Strategy Sections to Customer Groups

# Questions

---

Jim Kreft

[jkreft@usgs.gov](mailto:jkreft@usgs.gov)

608-821-3919

Dwane Young

202-566-1214

[young.dwane@epa.gov](mailto:young.dwane@epa.gov)

