



# Using Web Services to Publish Data

---

2010 OEI Symposium  
May 11, 2010

Lee Kyle; EPA/OEI/OIC  
Pravin Rana; EPA/OW  
Kristen Gunthardt; EPA/OW/OWOW  
Dwane Young; EPA/OSWER/ORCR



# Overview

---

- EPA Program Offices have enhanced data access by implementing public-facing web services to allow for better access to programmatic data
- This approach allows for services to be re-used across many different applications, and also allows for the data to be integrated in ways that it hasn't been before



# History of data exchange at EPA

---

- Historically, web services have been used at EPA to publish:
  - Geospatial data
  - Exchange Network (EN) data flows
- EN utilizes XML schema for EPA-trusted partner data exchange of regulatory flows
- EPA also publishes data from Envirofacts and program-specific databases via queries, reports, text files, and downloads



# The future: using web services

---

- Goal: reusable, interoperable, standards-based tools, services, and components implemented '*a la carte*' to promote discovery/access
- Web services promote public-facing data access, as well as trusted-partner access
- Support re-emitting of data in a variety of types and services, including OGC and spatial web services (WFS/WMS, KML, GeoRSS)
- Integrate existing metadata capabilities, via web services, to allow enhancement of metadata and integration with a variety of systems

Application Layer

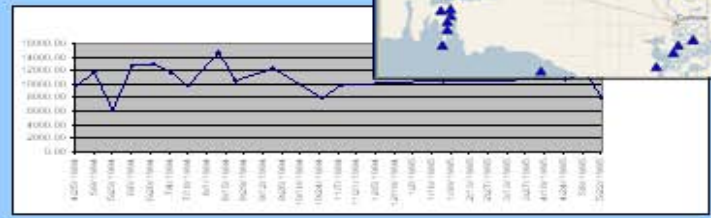
**EPA** United States Environmental Protection Agency

**MyEnvironment**  
environmental information for my area  
How to use this page? Help Us to Improve

Sites Reporting to EPA near 84307, UT

Search Enter fact

Sites and facilities by:  
Industry Release info



**Cleanups in My Community**

Search: All EPA @ This Area

Initial Searched Geographic Area: zip 84307

Area of Interest: Choose One

Water Monitoring Stations (To see all, visit these data.)

**STORET Watershed Station Summary: 16010204**

This page provides summary information for water quality data that are available within the watershed. You can access the water quality monitoring data listing on any of the list details links. These data are generated from EPA's STORET Database.

The data provided on this page were generated using STORET Web Services. For more information on how to incorporate these data into your applications, see the links.

For this watershed, the following organizations have reported monitoring data:

- USFWS** - National Park Service Water Resources Division
- FRD** - EPA National Aquatic Resource Survey Data
- UDNR** - Utah Department of Environmental Quality

Organization ID: 11971692 - National Park Service Water Resources Division

Station Name/Summary Information	Period of Record
<a href="#">830108 (See Details)</a>	10/11/1978-11/1979
<a href="#">830108 (See Details)</a>	10/11/1978-11/1979
<a href="#">830108 (See Details)</a>	10/11/1978-11/1979

Organization ID: 11971692 - EPA National Aquatic Resource Survey Data

Station Name/Summary Information	Period of Record

View on Map (Requires Google Earth)

Station	Monitoring organization	Activity/Start Date	Analyte	Result Value	Units	Comment
830-OK-BL-MANTUA-RES	7/11/1979	Dissolved oxygen (DO)	6000	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Temperature, water	19.70	deg C		
830-OK-BL-MANTUA-RES	7/11/1979	Specific conductance	325.00	uS/cm		
830-OK-BL-MANTUA-RES	7/11/1979	pH	7.80	None		
830-OK-BL-MANTUA-RES	7/11/1979	Nickel	0.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Selenium	0.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Silver	0.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Strontium	0.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Zinc	2.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Chromium	200.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Lead	0.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Nitrogen, Nitrite (NO2) as NO2	50.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Phosphorus, orthophosphate as P	0.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Mercury	0.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Total Coliform	83.00	#/100ml		
830-OK-BL-MANTUA-RES	7/11/1979	Nitrogen, Nitrite (NO2) + Nitrite (NO2) as N	250.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	CO2, Chemical Oxygen Demand	210000.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Nitrogen, Kjeldahl	1200.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Nitrogen, ammonia as N	200.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Nitrogen, Nitrate (NO3) as NO3	900.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Phosphorus as P	80.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Sulfate, Dissolved	194000.00	ug/l		
830-OK-BL-MANTUA-RES	7/11/1979	Specific conductance	315.00	uS/cm		

Services Layer

Catalog Services

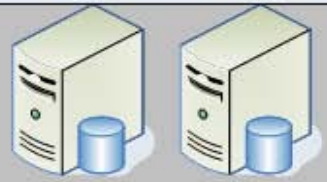
Data Services

Geo Services

EPA Program Data Sources

EPA OEI Shared Services

Shared Geographic Services



Data Layer

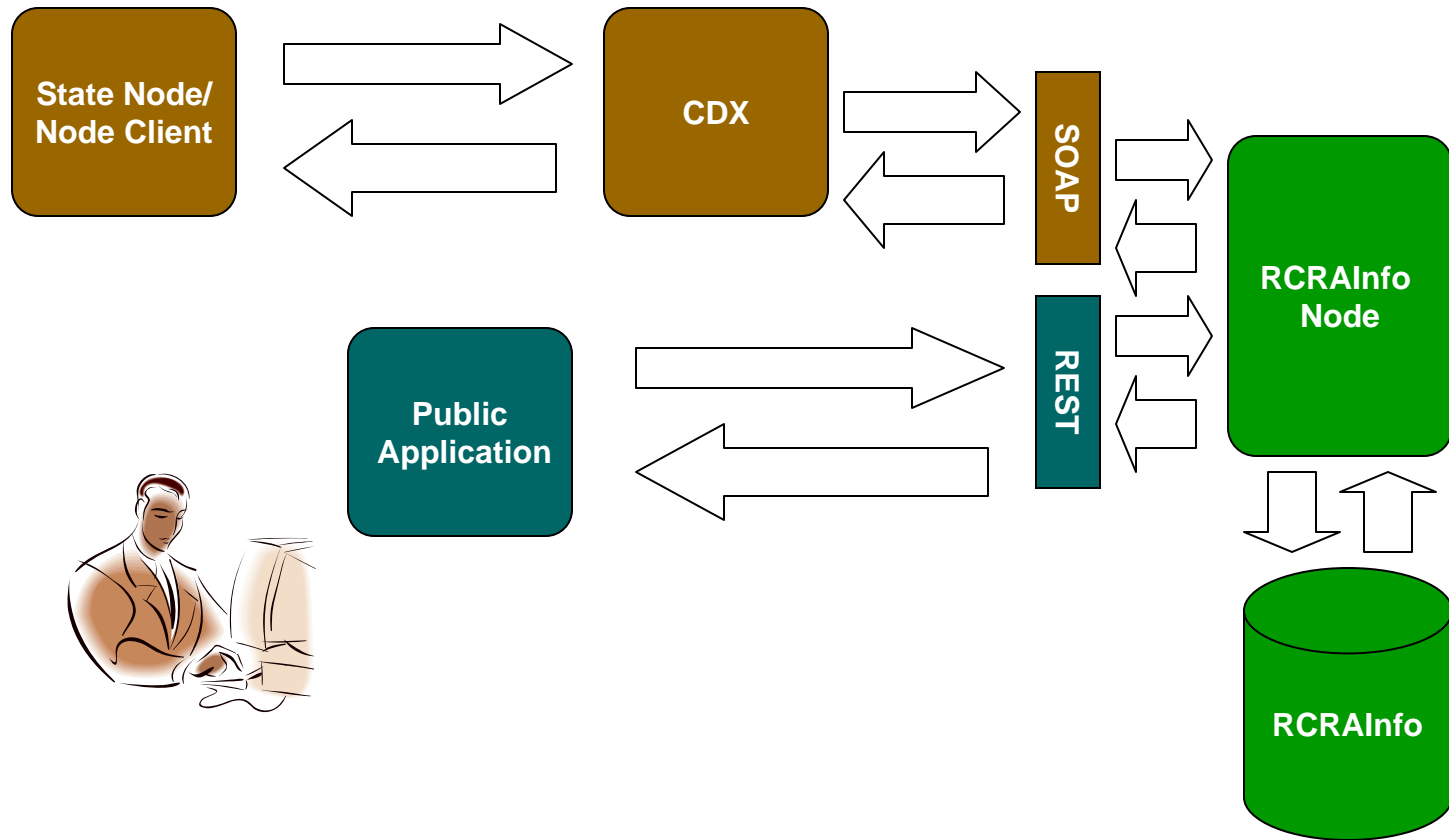


# The OSWER Approach

---

- RCRAInfo will have both Exchange Network (SOAP) and Public (REST) web services
- We'll use the same application to provide both
- Both types of services will use the same schema

# Query and Public REST Service



# Why we did it this way?

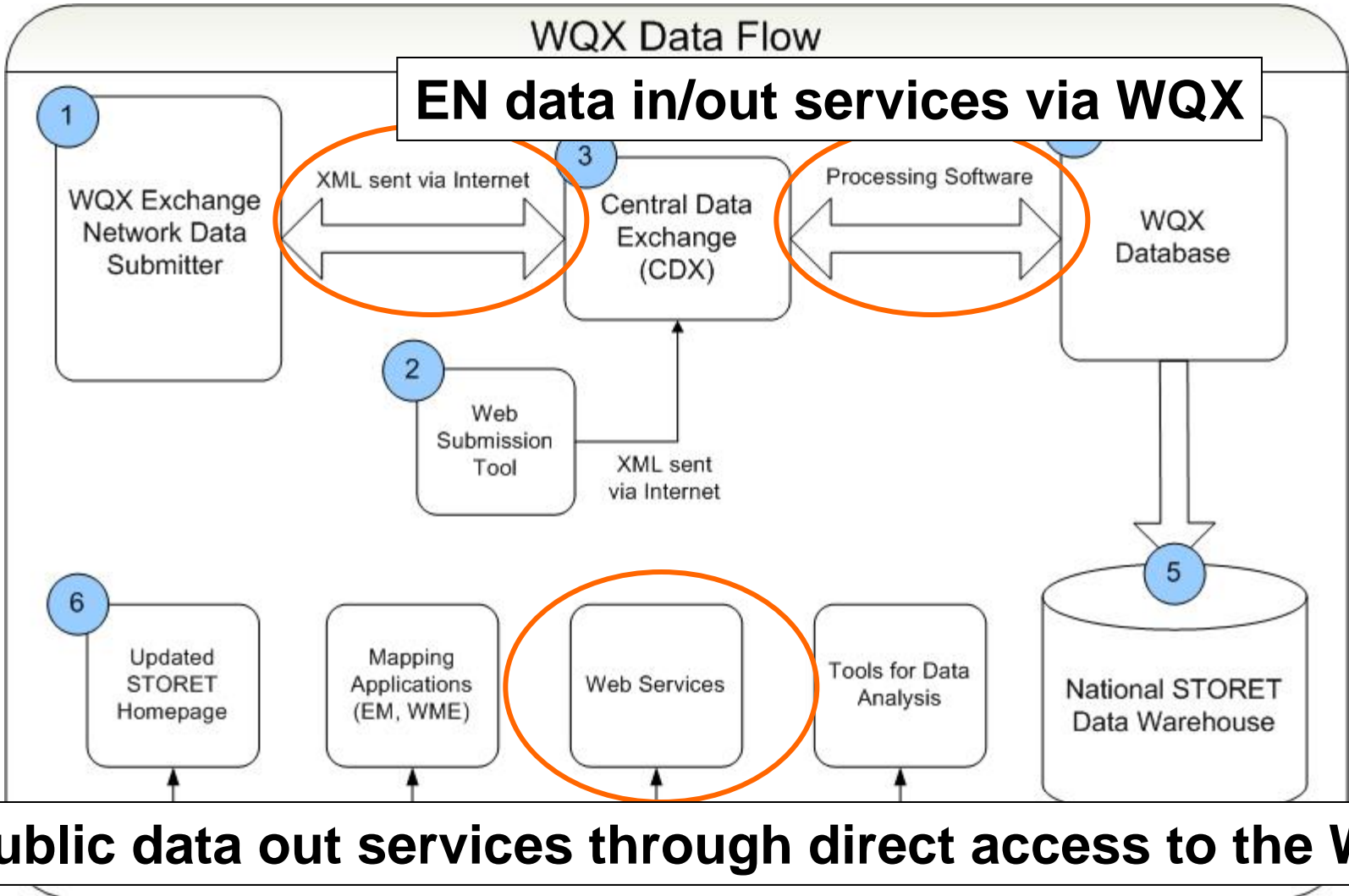
---

- Have both Public and Exchange Network data sharing needs:
  - The REST services provided a better public interface
  - The public data needed to be outside a username/password protected area
  - EN use case for States to sync their databases with data contained in RCRAInfo

# OW's approach with WQX

---

- Have taken an EN approach for supporting data submitters
  - WQX data services allow users to confirm data presence and accuracy in the WQX database
  - All Organization specific services return data in WQX v2.0 schema format (same format as input)
- Have taken a public approach for supporting data consumers
  - STORET Warehouse web services in SOAP ; REST like formats



**Public data out services through direct access to the WH**

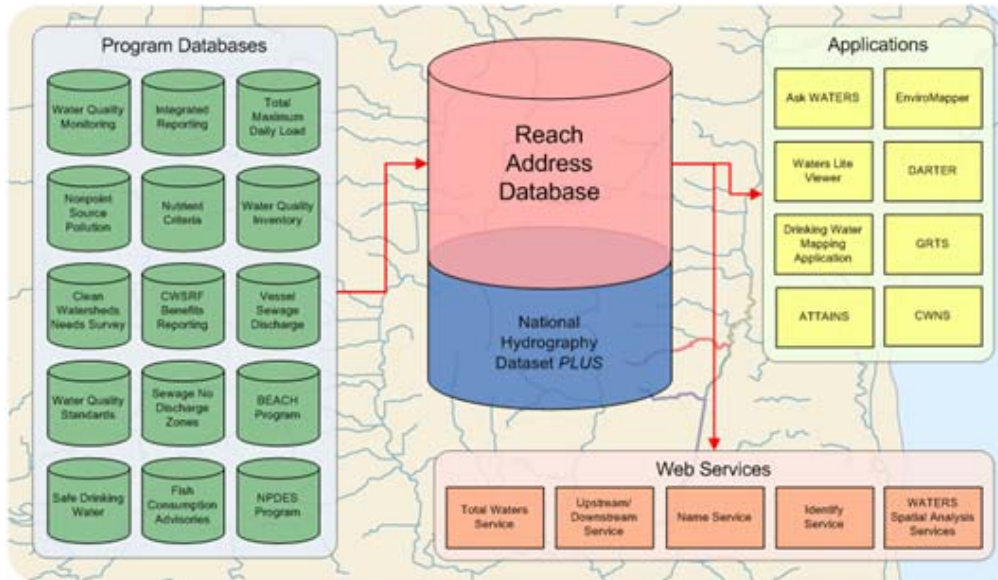
# Why did OW approach data access in this way?

---

- Data sharing needs have been met from serving data out of the STORET Warehouse structure
  - Data sharing needs are public in nature
- Data synchronization is key, but not a program requirement
- Data cataloguing for enhanced data discovery

# WATERS Overview

- **Watershed Assessment, Tracking & Environmental Results (WATERS)** - Unifies federal and state data from many different sources to provide easily accessible water quality information. The primary goal of the Office of Water is to achieve clean and safe water. Many different OW programs work together toward achieving this goal and WATERS helps them share important water quality information to meet this goal.



- Unites water quality information that was previously available only from several independent and unconnected databases.
- Program databases are connected via the digital network of surface water features known as the National Hydrography Dataset (NHD).
- Powers a variety of applications, web services, database services and Geographic Information Systems mapping services for both public and internal EPA consumption.

Additional Information: <http://www.epa.gov/waters>

# Available WATERS Data

---

- National Hydrography Dataset (NHD) Plus
- NHDPlus Smoothed Catchments
- Watershed Boundary Dataset (WBD)
- Legacy Watershed Boundary Dataset
- Water Program Features indexed to the NHD
  - 303(d) Listed Impaired Waters
  - 305(b) Assessed Waters
  - Beaches
  - Combined Sewer Overflows
  - Clean Water State Revolving Fund Benefits Reporting
  - Clean Watersheds Needs Survey
  - Facilities that Discharge to Water
  - Fish Consumption Advisories
  - Fish Tissue Data
  - Impaired Waters with TMDLs
  - Nonpoint Source Projects
  - No Discharge Zones
  - Water Quality Standards

\* Note: The above list does not include program system databases contained within WATERS.

Additional Information: <http://www.epa.gov/waters/data/prog.html>

<http://www.epa.gov/waters/doc/auxiliary/index.html>

# Why WATERS Enhanced Web Services

---

WATERS provided web services but:

- Not easy to understand or consume
- Lack of a visual interface
- Little or no outreach/implementation assistance available

## Goals

- Broad usage of services
- “Easy” to understand and consume
- Provide implementation assistance through documentation and hand-holding

# Web Services

---

- HTTP, SOAP, and/or Database Services
  - Validate and Convert Lat/Long
  - Get Entities by Lat/Long, HUC, and Radius
  - Get Program Waters Entities
  - NHD Event Indexing
  - NHD Point Indexing (Raindrop and Distance Snap)
  - Identify
  - Total Waters by state, HUC, FCODE, bounding box, and custom geometry
  - NHD/GNIS Name Lookup
  - Upstream/Downstream
  - Navigation
  - Navigation Delineation

Additional Information: <http://www.epa.gov/waters/geoservices/index.html>

# How we achieved our Goals

---

- Evaluated and deployed ArcGIS Server (AGS) 9.3.1 services
- Developed and expose REST-like endpoints for SOAP services
  - JavaScript Library
  - Create a Navigation Delineation Service that allows users to create user-defined drainage areas
  - Create SOAP-based versions of existing database Event Indexing, Point Indexing, and Navigation Services
- Exposed code samples so developers on how to use the services
- Developed a user interface for demonstrating services
- Created comprehensive documentation
- Made others aware by holding focused training/outreach with follow-up

# Created easy to understand documentation

- To promote service usage and to follow best practices, the WATERS team updated the WATERS Services documentation easier to find and interpret. The previous version of the web pages contained sparsely documented input and output service parameters, limited service description, and hard to find service endpoints. Based on the new content and layout, consumers at all skill levels have access to in-depth service information which has led to an increase in service usage.

The screenshot shows the 'Web and Database Services' section of the EPA WATERS documentation. It includes introductory text about the services and a table listing specific services with their supported protocols.

Service Name	JavaScript Library Function	HTTP Service	SOAP Service	Database Service
<a href="#">Validate Lat/Long Service</a> Utility function for validating a point's geographic latitude and longitude given in decimal degrees for use against the Reach Address Database (RAD).			✓	✓
<a href="#">Convert Lat/Long Service</a> Utility function for converting a point's geographic location latitude and longitude from degrees, minutes, and seconds to decimal degrees.			✓	✓
<a href="#">Get Entities by Lat/Long Service</a> Service for identifying details of all water entities within a given radius of a specified geographic location for specific WATERS programs.			✓	✓

The screenshot shows the 'Mapping Services' section of the EPA WATERS documentation. It features a diagram of 'Example Consumers' and their connections to servers and databases, followed by a section on 'ArcMap Based Map Document' and a 'Download' link.

**Example Consumers**

- EnviroMapper for Water: MapServer, OGC WMS
- WATERS Lite Viewer: OGC WFS, REST
- RAD Download: KML
- Bing Maps
- Flex/Flash
- Google Maps
- Google Earth: ArcIMS Image, ArcIMS Feature, OGC WMS, OGC WFS

**ArcGIS Server (9.3.1) Based**

**ArcMap Based Map Document**

For convenience, WATERS provides a preconfigured ArcMAP based map document (MXD) that contains pointers to the available ArcGIS based WATERS datasets. The MXD utilizes the NAD83 MapServer services and requires the use of ESRI ArcMap 9.2 (or greater). WATERS also offers the [WATERKML](#) files that are KML (Google Earth) based documents.

**Download**

[WATERSGEO\\_SERVICES.mxd](#)

**OW**

Service Name	Formats						Description
	Map Server	OGC WFS	OGC WMS	KML Server	REST	Cache	
<a href="#">NHD_Med_Basic_Reached_WMERC</a>	✓		✓	✓	✓	✓	Provides access to the reachable features of the...



# Outreach/Training

---

- Since the release of the enhancements last September, the WATERS team has conducted over 18 sessions with application owners, program offices, regions, organizations, and developers regarding WATERS Services.
- We followed-up and ensured that users had the support necessary to integrate the service



# Demo

---

- Using existing web services, we'll show how a simple spreadsheet can be used to pull RCRA data and water monitoring data together, without having to know anything about the underlying data structure