

STORET/WQX Conference Call – February 25, 2010, 12:00-1:00 PM Eastern Time

Introduction (Kristen Gunthardt)

- The minutes from all previous conference calls are available over the web: <http://www.epa.gov/storet/confcalls.html>
- The next scheduled call will be March 2010. The exact date will be emailed via the list server when the call gets closer.
- Please email storet@epa.gov and let EPA know you have attended the call so that meeting rosters may be kept.
- If you have a special topic you'd like to lead for an upcoming call, please email Gunthardt.Kristen@epamail.epa.gov
- EPA would like to hear comments you have on the quality of these conference calls. Please send them to storet@epa.gov
- Please subscribe to the STORET automated listserver for announcements regarding conference calls: <http://www.epa.gov/storet/listserv.html>

Latest with STORET Warehouse Webservices (publically accessible) – Kristen Gunthardt

- Original STORET webservices off the STORET Home page are cataloging services and stations and result services - separate from USGS Webservices.
- Tried to address issues where detection limit measure wasn't coming out depending on detection condition type, issues with certain data elements not coming out of the webservices – configured so result qualifier codes, activity time come out. We are trying to make sure that folks know this is a services version and we have additional items in backend processes that make sure data is being populated correctly.
- Kevin stated spaces in station names (pointed out by R10 that WQX supports station names with spaces) – received an incremental release of STORET result webservices, which is up today. This addresses a lot of issues brought up by R10 and would like feedback on those webservices, on whether issues captured and addressed. Issues were surrounding STORET Result webservice. Cataloguing services are unchanged.
- List serve message of availability later today letting everyone know that this service is up, just for get result service. Also posting changed schema file.
- URL for STORET webservices is on STORET Home page, see Features box. These webservices are only for production.
- How do we see submissions to WQX that are not in the Warehouse, yet?
- Everything so far described is public facing services off STORET data warehouse, but also have Exchange network services off of transactional database (query and solicit) can be used to look at test and production transactional databases. Full suite of exchange network services in Flow Configuration Document, that you can verify xml submissions with. Can pull exactly what was submitted.
- User - If there are still issues outstanding after looking at webservices, what is the turn around time? Turn around probably in next 2 weeks. 3 STORET System components - Web Application Download (text file), Web Services(xml), and Watershed summary. Expecting updates to all STORET system components.
- User pointed out - Using Solicit vs. Query, there are differences in the taxon domain list. There is a different number of taxa between these two lists. Should send to STORET@epa.gov . Kevin also recommended using the WQX Web application to verify domain lists using lookup function, as well as projects and monitoring locations that you uploaded to production. Requires registration to use. Contact STORET@epa.gov .
- Kevin mentioned the difference between publicly accessible and proprietary webservices.

STORET-USGS based Web services - Kevin Christian

- Have received production deliverable of the USGS Webservices. Beta in production now. It is publicly accessible.
- Miscue with RTP computer center on evaluating 2nd deliverable from USGS in evaluating for security scans. Working on negotiating these hurdles. Probably by mid-March will have deployment hurdles negotiated. Stay in contact with list serve for messages concerning release. Website will also have a what's new.
- STORET-USGS webservice is not replacing our Webservices, it is just an additional section to our webservice home page. It is a joint format, which is the WQX XML schema format, which we have negotiated the domain values with USGS and other partners.
- Not a lot of change to Beta of 1st release, thus may be released earlier than mid-march to get it in use.
- Services will be featured at some upcoming conferences.
- User – when working with Beta not all data current. When will all data be included and what is the refresh? Will it be possible to get a JSONDE download? We need to talk with USGS about JSONDE. They may be using JSONDE for gauge data. The refresh issue needs to be investigated. Warehouse queries are strictly the data tables, services join between the inventory (cataloguing) tables. Trying to rewrite that logic. USGS services join station_fac table with result_fac table. The counts should come out the same, but Kevin will work with USGS on logic. Our June release of webservices will only query catalog tables if it is a cataloging services. Data services like result and station services only join catalog service and acts like a filter. Only will use the inventory tables when we call the inventory service and when we need stations services that go directly to fac tables or to data tables much like the warehouse application. Tables have been joined between services, thus the discrepancy in what you are seeing. This is not a refresh issue, but something that is based on joins on the backend.

WQX and WQX Web flow not being operational – Kristen Gunthardt

- You may have seen list serve messages about WQX and WQX Web in regards to the flow was not operational.
- We had a recover issue with WQX backend transactional database causing several issues about the last few weeks. Needed to make sure that people weren't submitting to WQX flow until data base was restored. Other errors occurred referring to unique constraint errors were occurring.
- The consequence doesn't affect the majority of users, but there may be a few that we will contact to ensure that we have all of the data.
- This occurred over the President's data weekend (12th). It was a low activity weekend. This affects the production instance of WQX. Both WQX Node and WQX Web work against the production database. File corruption recognized. Recovery instituted and were able to recover back to February 8th, then needed assistance recovering to February 12th. By the time we returned on the 16th, then the best they had was current as of the 8th.
- All major issues are resolved, but there are a few issues.
- 12 organizations had activity between the 8th and the 12th. We will contact those organization asking them to resubmit to ensure that their data is correct.
- User – I have data files to import, is it best to do 1 at a time or do all at once? Each file 30,000 + records. Kristen answered in WQX Web that you might run into hanging files. 30,000 records falls within a spreadsheet worth of records, so the system should be able to handle. In terms of

performance, it should handle it, but will take time. Make sure that Activity IDs are unique across organization so that later submissions do not replace the formerly submitted. No warning generated if Activity IDs will replace old ones, but the processing report will show counts of activities updated.

Q and A

- Q:** Wendy Gayle – We have 80 characteristic name translations set up, will this affect performance? What is the upper limit of translations?
- A:** Kevin said that it will affect performance. Using find and replace engine will give better performance. Let each tool do what they do best. The fewer the translations the better.
- Q:** User - Would it be a problem to submit combined files that yield well over 100,000 records?
- A:** Beyond spreadsheet limit there will be a performance impact. Kevin recommends that you aggregate into smaller chunks. Kristen recommended that different activities referencing the same projects and monitoring locations do not have to be in the same files. Processing software can handle a lot but the XML generation is limited.
- Q:** Jim in AK – We need to have a unique Activity ID every time you upload data?
- A:** Yes, you can think of it as analogous to a Lab sample ID.
- Q:** Wendy Gayle – How often are domain values updated and how are notifications made?
- A:** Kristen - Not on a regular schedule, but as users request. 99% of changes are mostly additions. Changes in taxon names result in retirement of old name. Have not done a good job of notifying people of changes. Tend to rely on folks using exchange network services. Long term plan is to make this self reporting service that you can query. Number of systems that we reference Nemi.gov for analytical method and epa.gov\srs and ITIS or Catalogue of life are adopted. The rest of community knows by querying table.
- Q:** Jim in AK – Can you cover verifying data in WQX?
- A:** For WQX Web, most complete way is check in STORET Data warehouse after a week, or use webservices that applications can use to call data. Node/Node client users can use services that are not in the WQX Web tool. WQX Web has real time data reports under the lookup area for Projects and Monitoring Locations, but will expand to activity list (list of activity ids).
- Q:** User - What is the best way to notify you about hung up files in WQX Web to clear them out? What if dataset has dropped off the list?
- A:** Send email to STORET@epa.gov. If you need to try to load immediately, change filename to try to reload. The drop off problem may be a bug, but it should be there at least a week. One of the releases you can also control the number of things listed on a page.

Participants on call (based on who emailed storet@epa.gov)

Rebecca Ross	Texas CEQ
James Hileman	Region 10, USEPA
R. Bruce Tuttle	Idaho Dept of Water Resources
Dean Tucker	National Park Service
Linda Cartwright	Tennessee Department of Environmental Control
John Tooley	Washington Department of Ecology
Jim Porter	Minnesota Pollution Control Agency

Larry Merrill	Region 3, US EPA
Donald Kean	NC DENR
Jon Becker	Region 4, US EPA
Ian Ogilvie	GULF OF Maine Research Institute
Michael J Whitman	West Virginia Department of Environmental Protection
Paul Koska	Region 6, US EPA
Deb Soule	NH Department of Environmental Services
Arne Sjodin	Public Health & Environment, CO
Shawn Henderson	Region 7, US EPA
Dallaire Thomas	MA Department of Environmental Protection
Johnnie Purify	Region 4, US EPA
Elizabeth A. Florom	Bureau of Water, south Carolina
Erik W. Leppo	Tetra Tech
Paul Andrew	RTI International
Deepti Puri	US EPA
Kevin Christian	US EPA
Kristen Gunthardt	US EPA