Anne Weinberg

Okay, good afternoon and welcome to today's webcast titled: Using the New Water Quality Portal. This webcast is sponsored by EPA's Watershed Academy and EPA's Office of Wetlands, Oceans, and Watersheds also known as OWOW. I am Anne Weinberg with EPA's Watershed Academy. And I will be moderating the webcast along with Michael Brennan who is an Oak Ridge Institute for Science and Education, we often say ORISE, fellow who works with the WQX/STORET team. Thank you all for joining us today.

We will start by going over a few housekeeping items. The materials in this webcast have been reviewed by EPA staff for technical accuracy. However, the views of the speakers and the speakers organizations are their own and do not necessarily reflect those of the EPA. Mention of any commercial enterprise, product, or publication does not mean the EPA endorses them.

First, I would like to briefly summarize some of the features of today's webcast. We encourage you to submit questions to our speakers during the webcast. To ask a question simply type in the "Questions" box on your control panel and send click. Or click "Send." If your control panel is not showing, simply click on the small orange box with the white arrow to expand it. If you have any technical issues, can let us know by entering them in the questions box as well. That is to the right side of your screen. And then clicking on the "Send" button. We will do our best to respond to your issue by posting an answer in that same questions box. The webcast will be recorded and archived so you can access it in a few weeks after today's live presentation. The archived webcast will be posted on EPA's Watershed Academy webcast page at www.epa.gov/watershed webcasts.

So now that we have completed that discussion of housekeeping items, let's kick off today's webcast. Today we will be viewing the new Water Quality Portal for water quality data which was developed as a team effort by the US Geological Survey, US Environmental Protection Agency, and the National Water Quality Monitoring Council. The portal brings together chemical, physical, and microbiological data from USGS's National Water Information System or what we call NWIS and EPA's Storage and Retrieval Data Warehouse that we refer to as STORET. The portal provides this information to scientists, policymakers, and the public in a single user-friendly web interface. The portal reduces the burden to data usage of searching, compiling, and formatting of water quality monitoring data for analysis. This webcast will include both informational presentations, some PowerPoints, and a demonstration of the portal in action. While we encourage you to look at the Water Quality Portal in the future at www.waterqualitydata.us. Please do not access the system during the webcast as we don't want to overload the server. The webcast will also highlight and provide a demonstration of EPA's water quality exchange or what we refer to as

the WQX Web.   This is a web based data entry tool that enables data owners to upload their data so that the public has access to that data through our new Water Quality Portal.

Without further ado, let me introduce our speakers.   Our first speaker, Susan Holdsworth, is Chief of the US EPA's monitoring branch and is co-chair of the National Water Quality Monitoring Council.   Susan has served as Chief of EPA's monitoring branch for number of years.   Monitoring Branch has responsibility for managing a number of water quality data systems including the STORET Water Quality Exchange System and the new Water Quality Portal which is a team effort by EPA, USGS, and the National Water Quality Monitoring Council.   The Monitoring Branch also leads EPA's National Water Quality Resource Survey for lakes, streams, wetlands, and coastal areas.   Recently the Monitoring Branch has also been involved in mitigation efforts to address the Gulf oil spill.   Susan will kick off the webcast and provide a brief introduction to the new Water Quality Portal and the Water Quality Exchange.

Our second speaker, Nate Booth, is participating remotely in Madison, Wisconsin.   He is the lead architect at the Center for Integrated Data Analysis, CIDA.   I am sorry for all of the acronyms.   CIDA, again Center for Integrated Data Analysis.   CIDA partners with USGS programs to provide innovative computing solutions for conducting interdisciplinary science across a growing array of massive environmental data collections.   Nate also serves as the Data Synthesis leader for the USGS National Water Quality Assessment Program.

And our third speaker is Charles Kovatch who is the Water Quality Exchange and STORET team leader in the US EPA's Monitoring Branch.   The Water Quality Exchange and STORET team assists federal, state, local, and tribal agencies, universities, and watershed groups, and others with their WQX water data transmissions to the STORET data warehouse.   And helps make the data available for analysis.

One final note before we get started with our first speaker, we will try to answer as many questions as possible throughout the webcast.   However, due to the large number of participants not all questions will be answered.   In the event that your question is not answered, please feel free to contact the speakers after the webcast.   You can access the speakers contact information on one of our final slides for today's webcast and on the additional resources document posted at www.epa.gov/watershed webcast.   And one final note, we will post a copy of the slides after this webcast, but a large portion of the webcast is going to be a demonstration of course, we can't capture that with the slides.

So with that, we are about to begin our webcast.   Our first speaker is Susan

Holdsworth who is Chief of EPA's Monitoring Branch and co-chair of the National Water Quality Monitoring Council.   Susan, the floor is yours.   Take it away.

Susan Holdsworth
Thank you, Anne.   Hello, good day to everyone.   This is Susan Holdsworth from the monitoring Branch in EPA's Office of Wetlands, Oceans, and Watersheds.   I am co-chair along with Mike Yurowitz from the USGS of the National Water Quality Monitoring Council and we are thrilled about this product and the opportunity to share it with you today.

The National Water Quality Monitoring Council is chartered under the Advisory Committee for Water Information and the website for the Council is acwi.gov/monitoring.   I encourage you to come and visit the website and learn about our workgroups which include the water information strategies workgroup which is hosting the Water Quality Data Portal.

Again, I am thrilled for the opportunity to bring to you all today the Water Quality Data Portal.   It was first released for public access just this past May at our national monitoring conference.   A meeting that we hold every two years that draws about 1,000 participants from states and tribes, federal agencies, volunteer organizations, municipalities, and other public and private sector organizations.

The Water Data Portal that you will learn more about today is serving over 150 million records contained in the USGS National Water Information System and the EPA's Water Quality Exchange STORET data warehouse.   This portal serving data not just from EPA and USGS but from a number of other federal agencies, state organizations, tribal governments, volunteer monitoring organizations, academic communities.   Over 200 organizations archive and store data and are able to serve it out through the Water Quality Portal.

This project has been a long time coming.   The 2003 Memorandum of Understanding was signed by the Director for the Office of Wetlands, Oceans, and Watersheds and the Associate Director for Water under USGS.   So EPA and USGS signed the MOU reiterating their commitment to developing an integrated interoperable data sharing tool working with the National Water Quality Monitoring Council.   As you can imagine, given how long we have been working on this, many hours have been spent developing an inventory of the data elements in the system and working with the broad community of data collectors to identify what are the core data elements necessary to support the documentation of water quality monitoring results ranging from where were they collected, why were they collected, how were they collected.   And you will learn more about that as we go through.

Another activity that took a lot of time is working through the development of a common library for chemical name, analytical method, so that again, these data are comparable and shareable.   And we were fortunate that during this time there were big advances in information exchange technology that allowed us to do this much more efficiently than we probably envisioned back in 2003 when signing this MOU.

All of this activity, getting us to the point today where these two large information systems are merged virtually through this web portal which USGS's Nate Booth will show you soon.   You can retrieve data from both systems and as I mentioned earlier, about 200 organizations.   This data comes out organized in a common format and delivered to you in a single file.

So with that I am going to reiterate some of what I see as the benefits of the portal and I would certainly encourage you all to share any thoughts on how you see the portal being of benefit to you.   I think one of the first things that this portal does is reduce the level of effort that it takes for you to compile data from multiple sources to be able to use it.   I am sure most of the folks on the phone have had the experience of needing to collect data not just data that you have available but data from other organizations in order to support your analytical needs and I'm sure many of you have toiled with the efforts of trying to consolidate that data into a single format, to get the appropriate documentation about how the samples were collected so that you could use it for your analysis.   This is a huge benefit of being able to go to the Water Quality Data Portal.   It reduces the effort to use data from other sources, reduces the effort for collecting data, combining it into that single format, and delivering it through a single file that is now ready for your analysis.

Another benefit of the portal and the Water Quality Exchange is the service that it provides to leverage and protect all of our investments in monitoring data.   The common data elements emerging from the monitoring community and adopted by the Advisory Committee on Water Information are one way that we can share data and ensure that data that is collected is available for future use, collect it once, use it multiple times.   The other thing the portal provides is somewhat of a marketplace of what data are available, when and where monitoring is taking place that you all can bring into your decision needs.   And of course, the overarching benefit, the really "why did we do this" is that having this data allows us to support water quality base decision-making bringing information and science to the decision-making process.   Enabling us to pair ambient data to applicable water quality standards.   Enabling us to identify hotspots around the country.   Enabling us to develop protection and restoration plans that are based on sound science.   And enabling us to pull these data into building better models

to predict water quality and to show how water quality may improve under different control strategies.

So what is next for us today? We are going to have Nate Booth from the USGS lead you all through a demonstration of the portal and show some applications of the data now that it is available through the portal. And then Charles Kovatch from EPA will show you how to add your data to the portal through the Water Quality Exchange.

And so I would like to invite you to help to provide a couple of examples of how you can help us maximize and leverage the investments that you all are making in monitoring data and that we have made in the portal and the Water Quality Exchange. So I would encourage you to use the portal after the call. To use the data. I want to let you know as Anne said the webinar is being recorded so that you can refer to it later and you can refer other folks to looking at it to get an overview of how to use the portal and get data into it. I also want to mention that the portal has a user guide which you can see from the portal website.

I would also like to invite you to add more data through the Water Quality Exchange. Make more data available through the portal. And to invite you to visit the tutorials that are available through the Water Quality Exchange for detailed instructions beyond the overview provided in this webinar.

And finally, I would like to ask you to provide us feedback on likes and suggested improvements that you think about, identify as you are going through the portal. And with that I will open it up for a few questions or clarifications before we get to the meat of this webinar. Thanks.

Anne Weinberg
Okay, thank you, Susan. We are now going to have, that was an excellent presentation. We now have time for a few questions from the audience. Michael Brennan will pose a few questions for Susan.

Michael Brennan
Sure. Susan, we have one question that you can answer right now before we move on to the next part of the presentation. Can you comment on the scope of the portal? Do all the states include their data in the Water Quality Data Portal?

Susan Holdsworth
We have about 10 or 12 states that do not have current data in the Water Data Portal. But as you'll see later from one of the maps, there is some data, whether it is USGS data from NWIS or other sources for every state. And as you will see it later in the map of the data distribution there is a lot of room for getting more data and we are actively

working to help states and tribes flow their data through the Water Quality Exchange and invite other organizations whether it is local government or volunteer organizations, academic groups to come on board.   Thanks.   That's a great question.

Michael Brennan
I think that's it for questions for now.

Anne Weinberg
Okay.   Thanks, Susan.   We will likely have some more questions for you later on. People will have other questions that come to your mind for EPA, Susan will stay with us.   So our next presenter is going to be Nate Booth from USGS.   And we will make the screen really wide.

Okay, Nate, take it away.

Nate Booth
Okay, you see my screen, okay?

Anne Weinberg
Yes.

Nate Booth
Okay, very good.   Thank you.   Alright, hello, everybody.   My name is Nate Booth; I'm from the US Geological Survey.   It is a privilege to present today on behalf of all of the folks that have contributed to this project over the years as Susan mentioned.   This has been a long time in the making.   It's had many facets and just here within a few months ago we finally have been able to produce this portal that provides easy access to this large collection of water quality data and I'm going to walk you through some of the features of that portal.   This slide that I am showing here was the title slide for our launch at the National Water Quality Monitoring Council back in May.   And as Susan suggested, the portal provides access to over 150 million water quality records from states, tribal partners, US EPA, and USGS.

The portal has these 4 general characteristics and they are as follows.   So again, we've mentioned this point several times already but the portal brings together these two large water quality database systems, USGS's National Water Information System and EPA's STORET.   Charles will be talking about how data flows from various organizations into STORET after I finish my slides.   The Water Quality Portal is based on the Water Quality Exchange data standards and convention and that data standard provides several benefits.   One major one is that it provides data sources or providing organizations a common convention for expressing information about our monitoring sites, watershed characteristics, the results of our sampling, as well as important field and laboratory metadata to help ensure the proper use of the data in analysis.

The Water Quality Exchange standard also provides a common target for folks to develop other applications that use data from the portal and I will talk about few examples of that later. The portal is -- a lot of the time was spent in developing the portal has been in optimizing performance. So I'm sure many of you have been able to access data in the past for a few monitoring locations at a time and the Water Quality Portal is organized to be able to support broad regional and national organization assessments so there is very little limits to the portal in terms of how much data you can down load both geographically and temporally. The Water Quality Portal is updated every night from NWIS and every week from STORET.

This next slide is a very simple schematic again that Charles will provide some additional detail to but on the USGS site all data are compiled from our state offices around the USGS in the portal overnight and data is compiled in the EPA's STORET data warehouse on a weekly basis and then that makes its way into the portal.

Okay, so moving along into the demonstration. So when you come to the portal at www.waterqualitydata.us. This is the page that you will come to. The page is organized into four sections starting in the lower right there is information that provides information about background on the portal, information about contributing organizations. There is a map that displays the data coverage across the country that gives an indication of which organizations are providing and how much data is being provided by each, by state, by county, and by watershed. There is quite a bit of material on how to use the portal so there is good user guide material as well as user guide material for the web services that are available from the portal and I will describe those later. And then finally, the part that I will be demonstrating next which is the download data application within the portal which is on the left-hand side of the screen.

So when you click that, when you click that link this is what you will see. This is the query page. We tried to design the query page so that all of the information that you need to define your queries in one screen there is a lot of links to help materials and pick lists across the page. And many of our users have described this webpage as seven clicks away from virtually any water quality data you want to download. So I will be walking through the sections of this query page next.

So the first section on the top indicates the filter criteria that are geographic in nature so you can search by county, country, state. There is actually some water quality data from outside the United States. You can search for data within a radius of a position that you are near or some other arbitrary position that you can type into the form. You can also search by a bounding box so those

two are useful when you want to be able to access portal content through another mapping application as well.   The next section is a series of filters that are based on monitoring location information.   So these filters add the who, what type, and specific monitoring locations you might be looking for as well as hydrologic units.

Next, you can search by sampling information and this includes the filters around what media has been collected, the types of constituents that were analyzed. You can search for particular characteristics or constituents of interest.   And you can also specify a range of dates that you are interested in.

Next, you specify what you would like to download.   So you can download information about the monitoring locations or you can download information about the sampling results.   You can choose to search across both USGS, the National Water Information System and STORET or one of them in particular. And then you specify which download format you are interested in.   You can download tabular data in CSV or tab delimited or directly into Microsoft Excel format.   And then this next slide shows an example of what a water quality result retrieval looks like in Excel.

And this slide gives some specifics of data elements that are typical in a site download.   So this example shows a stream monitoring location, Fanno Creek which has a 23 square mile watershed.   And it is maintained by the USGS Oregon Water Science Center.

You can also download a geographic file format called KML.   KML is now a standard that was developed by Google.   And KML files can be opened in Google Maps, Google Earth, and other compatible GIS software.   So it's an easy way of displaying the monitoring locations that meet the filter criteria that you specify.   And when you click on points within Google Earth or Google Maps you can get similar information about the monitoring locations as I displayed in the last slide.

If you download the water quality results, this slide gives you an example of one typical water quality result.   So this example shows a suspended sediment of nitrogen concentration result that was taken on April 25th of last year at the Fanno Creek site.   It was collected by the USGS.   It was analyzed at the USGS National Water Quality Lab.   The sample was taken during a storm event during the rising stage of the hydrograph.   There is also metadata listed here for the analytical method that was used as well as the long-term method detection limit.

Okay, the portal has been developed so that it works on smartphones and tablet computers.   And if you have one of those devices with a GPS chip on board, you

can come to the form that I showed and you can click a link there that retrieves your current location using the GPS and quickly search for monitoring locations nearby and display those locations on a map on your tablet device.

So I am going to now run through a slightly more specific example retrieval.   In this scenario I am interested in -- I'm involved in a study that is looking at nutrient conditions in the early 2000's within the Big Thompson River Basin in Colorado. So what I am going to do is come to the portal.   And I will first choose a stream site location.   I'm interested in surface water so I will choose "streams."   Within the portal on the portal query page, if you click on any of these links to the right of any of the form elements you will get either a list of values like what is shown here or you will get some description about what sort of information can be entered into that form element.

So once I chose "stream" I will also specify the hydrologic unit that corresponds with the Big Thompson Watershed.   And I will have found out which hydrologic unit that is by clicking the question mark here and finding the hydrologic unit for Big Thompson Colorado.

Next, I'm going to choose the sampling media of water.   And next, I am going to choose a group of constituents from the characteristic group choice of nutrients. And so once I made those two selections they will appear in the form.   And then next, I will specify the date range that I'm interested in.   So 2000-2004.

Next I will select the sites that -- I will choose sites to download next.   And then I will press "Download" to download those results.   The point I want to make with this form is that you can fill the form out once and you can download sites that meet those criteria and then the results that meet those criteria.   So any of you that have been involved in water quality data analysis I am sure have looked for data in sort of a two phase approach.   First, looking for sites that meet some criteria.   And then secondly, downloading the water quality results for those sites.   And so this form is set up to make it easy to do just that.   So you specify the filters around a set of water quality data you are looking for, download the sites that meet those criteria and download the water quality data after that.   So once we choose sites and press "Download" the portal will prompt you with how many sites will be downloaded.   And if you think that number is satisfactory, you can press "Continue" and download that file.   And then if you change the download results, you will be prompted again with how many results that will include as well as from how many sites that is.   And again, you can press "Continue" to download that file.

So this map here shows the results of that query.   And the point being that while that is perhaps a nice amount of data from either the USGS NWIS system or the

STORET system, by pulling these data together collectively we, of course, have a much more robust monitoring network to base our studies and analysis.

Okay, so moving on beyond the demo I'm going to next show a few example applications of how folks have been using the portal to support other more specific needs.   So when we started developing the portal we very deliberately designed it to not only have a simple search interface that I just demonstrated but also provide access to those same data programmatically so through software using a technology called Web Services.   So these next applications use that capability to use the portal and the data behind the portal that is updated regularly in those applications.   And we invite all of you in the community in general to continue to develop these sorts of applications and we would very much like to know about those applications.   We would like to share those applications with others around the community.

So this is the first example.   The Texas office of the USGS developed what is called the NWIS Snapshot tool.   This is a tool that works within ESRI's ARC GIS platform that lets you pull together stream gauging locations with water quality sites and groundwater levels altogether in a GIS and the water quality information to access the portal for that content.

Secondly, a sister committee to the National Water Quality Monitoring Council within the Advisory Committee on Water Information Subcommittee on Groundwater is actively developing a national groundwater monitoring network that is available at the URL that you will see in the slides.   And for that monitoring network you can retrieve information about groundwater wells that are part of the network, time series that describe the water levels within each well, and also water quality information from those wells.   And for certain states, the water quality component of the national groundwater monitoring network comes from the Water Quality Portal.

Next, the National Water Quality Monitoring Council has for a long time had the goal of monitoring from summit to sea and this next example is supporting modeling of the near shore environment to be able to characterize the effects of tributaries in the Great Lakes.   And this tool is based on data standards for both WQX and other standards that have been endorsed by NOAA's Integrated Ocean Observing System.   So this tool that is also available on the Internet and you can get the URL from the slides lets you integrate water quality information from the portal with real-time rainfall information across the watershed of interest, integrated with other time series information including water levels and stream flow as well as real-time currents from hydrogen models in the Great Lakes.   So this is another good example of how Web Services from the portal provide water quality monitoring information to other applications.

And finally, a slightly more advanced example is supporting what appears to be a new paradigm in analyzing water quality data and that is accessing distributed data across the Internet using Web Services directly within statistical tools. So this screenshot shows a product or an application called Our Studio which is an open source statistical software package. And what we are doing within this environment is accessing nitrate data from a particular site. It uses the Web Services from the portal to populate a data frame within R. and within a few lines of code we are able to quickly develop these box and whisker plots that shows the distribution of nitrate concentrations by month throughout the period of record. So it's a very powerful set of water quality statistical tools within this platform. This is being developed by Bob Hirsh and Laura DeCicco within USGS. You can download this package again from another URL that will be part of the slide deck.

Okay, so wrapping up my part of the presentation. We have tried to assemble a good bit of user documentation and resources within the portal and we welcome you to access user guides both for the portal and for the Web Services. We have assembled a good list of frequently asked questions as well as materials for data providers that are not currently part of the portal that would like to be. And we also can make some training materials available on request from a training course that we gave at the monitoring conference.

And closing for my section, I will just quickly mention a few feature enhancements that we have heard from user feedback across the community. So these are not things that we will necessarily get to right away butt they are things we've heard about and that we hope to get to given adequate resources. So I will just go through them quickly. Folks have asked many times for a richer mapping display within the portal. That is something we are actively developing. So you will be able to do something similar to filling the form out on the portal like I showed and be able to display that information directly on a map without having to download a KML file and going through the steps that I showed. It's a more direct mapping of information. That is one.

Secondly, going back to the upper left, we would like to use the National Hydrotherm Feed data set as a common river network to link monitoring locations together. Again, many of you that worked with water quality data sets in the past know that different agencies may be monitoring the same location but may have that location coded slightly differently. And so linking things to a common river network would give us a way of integrating that data spatially and let us analyze river reaches of interest.

Next, we will continue to work with the National Environmental Methods Index or

NEMI, which is another project within the National Water Quality Monitoring Council. It provides standardized references to analytical methods.

Besides bringing on additional data partner that Charles will speak to next, there has been also quite a bit of interest in displaying and providing access to real-time monitoring data within the portal as well as accessing biological and habitat information.

And then finally, we would like to continue to highlight tools that are developed across the community like the last few that I showed. And we will try to do that in some way through the portal and that National Water Quality Monitoring Council. And so with that, I will close and if we have a few minutes for questions, I can try and take them.

Anne Weinberg
Okay. Thank you, Nate, for that information. Our next – thanks, Nate, again. Our next presentation will be by Charles Kovatch but before we do that we will have a couple of questions for Nate. And the question session will be led by Michael Brennan.

Michael Brennan
Thanks a lot, Nate. I would just like to note now that we have had a lot of questions so far and we will get to all of your questions after the meeting. But at this point, we had quite a few questions on the topic of data quality. Nate, could you comment on if all fields are required, need to be completed when submitting a data request or can you leave some of the fields blank when using the Water Quality Data Portal?

Nate Booth
Sure. So I'm guessing that the question is related to data being supplied to the portal rather than form fields on the form for retrieving data. So if it's the first one, the answer is no, not all fields are required and Charles will probably provide some more information about that during his presentation. But QA and QA/QC quality assuring the data of course, is quite important. There are QA/QC checks that are run both on the NWIS system and on the STORET system that provides some quality assurance to the data. In terms of which fields are mandatory, that set of data elements can be provided as part of a spreadsheet. I think it is part of the portal user documentation. If it's not we can follow up with that in the spreadsheet to anyone who's interested.

Michael Brennan
All right, thanks, Nate. Next, if you could comment on the HUC parameter. When users select the HUC parameter will that apply it all data inclusive or is there a possibility to leave out data within a watershed? In other words, is it an optional field to fill in when data suppliers submit data to the Water Quality Portal?

Nate Booth

Yes, so if you were to specify any of the fields within the portal including hydrologic unit code, the data is going to need to be, the data is -- I'm sorry the data needs to be -- needs to satisfy that filter in order for to be retrieved. Now with hydrologic unit code you can specify -- you can use wildcards so you can specify hydrologic unit code 04* for example, to get all of the hydrologic unit code's within that larger area. But yes, if you specify filter the data it needs to meet that filter in order for it to be retrieved.

Michael Brennan

Thanks, Nate. I guess one last question. Is there a limit to the number of stations and results that could be downloaded from the Water Quality Data Portal like a limit of 1 million results?

Nate Booth

Yeah, there is not a limit. There is a limit within an older version of Excel to open the large files but there is not a limit within the portal as far as how much can be downloaded.

Michael Brennan

Great. Thanks a lot, Nate.

Anne Weinberg

Okay. Thanks, Nate. Are those the last questions? Anything else? Okay. Our next -- our third speaker is Charles Kovatch, team leader for the WQX STORET team and take it away, Charles.

Charles Kovatch

Thanks, Anne. As we have talked about through this webinar we had an overview of what monitoring data is and how it serves multiple communities and we showed the portal, how you are able to get data out to utilize within different analyses and modeling tools. And as you've heard, what enables you to do that is standardization of data. And that's what WQX and WQX Web enable data owners are two different ways to apply the water community's data standards. And so that is what we will be talking about. We have a tool that will help you to share your water quality data to participate in a portal and so you can go and grab your information. And so the WQX tools plus the portal increase the value of your data by making it available to such a broad community of users. And really, like we said, it is a data standard that has been vetted in the water community and through that it improves the ability to share data.

So what does the tool do for you? It enables you to share your data in one format which, as we heard from the portal, is really important. You are able to -- thanks. You are able to get your information out of the portal. In one format

that you can incorporate into your analysis.   It improves interoperability of state data systems, enables you to publish your data at a national level where oftentimes if it was just at a local or state level.   It increases your ability to use others data which we also saw from the example that Nate provided because you are pulling data from both USGS and EPA STORET and because of the fact that EPA STORET includes data from federal, state, tribal, local agencies as well as other watershed organizations and academia.   And the really nice thing is that WQX standard enables you to manage data in the format that best serves your program needs.   So if you are still able to use your own system and then have the ability to incorporate whichever tool is best.

So what are our tools?   Our tools are both WQX and WQX Web.   WQX stands for the Water Quality Data Exchange.   It's an XML scheme that provides standard data elements and a solid file format.   It's intended for high-volume data users, meaning you have a lot of data and you submit data frequently. Frequently being maybe monthly or even could be daily for that matter.   WQX Web, of course, is the Water Quality Data Exchange Web template.   It is based in Microsoft Excel.   If you can use a spreadsheet this tool is for you.   We mean that.   It was designed as a simpler tool to use for folks that are not coders or don't have access to coding.

The tools enable you to take -- and following the schematic your data from your field sheets or right from the monitoring laboratory bench and whether you have the high-volume user, you are on your server, you are using WQX.   If you're just using your laptop and your Excel spreadsheet then you are using WQX Web. Both of those will travel through the exchange network which is a system of -- a network of shared users that exchange environmental information.   And then it will reside within STORET and be available on the Water Quality Portal.   And of course, hitting these points again, that you are enabled to share data in one format at the national level through the data standard.   And still again, managing the data in the format that best serves your program needs whether you are on the server or using an Excel spreadsheet.

Nate covered a little bit of this.   This is going into a little bit more detail of what is involved with the database standard.   It is the who, what, when, where, why your sample is collected.   It is not enough to just say we have five somethings of copper.   You want to know who collected it.   In this case this example it is the Friends of the Potomac River, a watershed organization.   They are sampling copper.   Why was it collected?   This is important for us to know too to get extent of what is that monitoring project.   Is a one-time sampling event or is a quarterly or is it monthly or even weekly?   Where was it collected?   Station name as well as that location so that you can go back to the portal and do these, like Nate showed, like latitude/longitude bounding boxes or searching within an

area, you know, 1 mile, 5 miles out from a point. When was it collected of course, is important. Date. What was the method, of course, is also important so you can understand the detection limits of that method and also the sensitivity of that method. And of course, the result value. End result units is critically important. Are you measuring parts per million or parts per billion?

So the WQX schema establishes that data structure for the water monitoring fields. It allows the data owner to use your existing database and it requires you to do a crosswalk between your data fields in your database and the WQX schema, the XML schema. So if you are calling something sample location, we call it sample station. So you have to conduct that crosswalk for the commonality of terms. And we make it a little easier for you because we have a list. It's called domain values or a pick list to help you match your database fields to the schema.

And again, the XML schema is designed for the high-volume user. You are submitting large amounts of data quickly. It requires coding in order to generate that schema. So you have to have the technical capability. It allows for automated machine to machine data submissions. Once this is established, you can have an automated submission to share data on any cycle you want: quarterly, monthly, weekly. And while it does require a high front and investment, you also get a long-term return on your investment because of the fact that you have established the system in order to call your data and you are able to return to it and utilize it multiple times.

If you have never seen an XML schema before, this is what it looks like. I am not going to go through all the details of this, but you can see that it lays out who is the organization, what is the description, so the organization here is a test organization. You know, we talk about the activity that we have. We have a water sample. The date that it occurred was July 19th, 2010 and the other information that follows suit. So this is just to provide you with an image of what the schema looks like.

Now the WQX Web tool, similar fashion. Still has the same structure for water monitoring data fields. It allows you to use your own database or spreadsheet. It still requires that crosswalk between the database and the WQX data standard. And again, you get to use the same domain values or pick lists in order to ease that matching. Here are the differences. Again, this is designed for the lower volume data owner. So you don't have large amounts of data. Maybe you are only submitting once a year. So this is, like I said, lower volume. It does not require coding. You are operating within a Microsoft Excel spreadsheet. So you don't need to do any coding. You don't have to rely upon some of the other technical support. It is a manual user to machine data submission. So while

the other one is automated, this one requires you to have that manual interaction in order to share the file. And it's a lower front end investment which means you have a shorter term return so that the other one you are able to use multiple times while this one you can use multiple times you will have to have a manual interaction in order to complete that transmission.

This is the WQX Web tool. As you can recognize it's a Microsoft Excel spreadsheet and its color-coded. Red for the export function, yellow for the data entry, and green are all the different pick lists, domain values, lookup tables that are embedded within the spreadsheet.

And this slide, this shows just one example of how this works. So we are looking at monitoring location fields. And in this case we have WQX test is the name of this location and you can see that the pick list is embedded for monitoring location type and looking at river or stream, but you have options here to look at spring, wells, and I can assure you there are lakes and other water bodies that are contained within that pick list. And this is the same for the other fields across the spreadsheet. The same thing with the results field. Here we have an activity type. We have sample routine or field observation. We are definitely in a water media. And once you populate when the activity start date -- when was your sample taken. The sample equipment name. We are using a water bottle here, but of course, there other options available. The characteristic name being we have temperature or turbidity, ammonium or nitrate. Again, there is a pick list that enables you to select the characteristic for which you are monitoring or measuring. Result value you have to populate that one on your own. But the result unit, we have milligrams per liter or degrees Fahrenheit. Again, a pick list and this last example here, the result sampled fractions this is dissolved. It could be total or other fractions that are within that list. So again, it lays it out for you logically within a spreadsheet and enables you to utilize the main values or pick lists to help populate the fields in a standard format.

Once you have completed populating the monitoring stations, the results then you get to the export function. This is that red section I told you or the pink. And we have three buttons here that grab from those spreadsheets and create the text files, flat files that you will be submitting over the exchange network. So you can export your projects, export the monitoring location, and export the results so that you will be able to share via the exchange network.

So reviewing this again, you are taking your field data sheets or your monitoring lab results. If you are the high-volume user of the database through WQX you have an XML file and then you are sending through the exchange network to STORET and making available under the Water Quality Portal. Otherwise the other option is running it through the laptop or smaller database using WQX web

tool, the Excel spreadsheet, using the buttons within the spreadsheet to generate the file that you can then send via the exchange network to STORET and then making available on the Water Quality Portal.

Once you submit your data, you are joining over 390 federal state tribal agencies, watershed organizations that are using WQX and WQX Web.   This is the map right off the Water Quality Portal page that shows you this is a combined data for STORET and USGS.   And of course, a darker color would indicate that there is more data available within that state and a lighter color is less data available.   The point here is that we have a lot of participation representing over 150 million records that are formatted to this WQX water quality data standard and making that data available for analysis.

You know, WQX Web improves the interoperability of data systems.   And by using these data fields you are increasing the value of your data making it available to a broad audience.   So multiple users can access your data through the portal.   And you are increasing your ability to use other data.   So if you are looking at a river or a watershed, and you have your organization's data.   You would also like to incorporate data from a federal agency or another state agency or a watershed organization, now you can pull that information into your analysis in the same format.

And another selling point here is that you are able to manage the data in a format that best serves your program needs.   If you are high-volume user or a lower volume user, you can select the tool that best suits your organization and still be able to share using the same common data elements.   The tools, as you can see from both the XML example as well as from the Excel spreadsheet through those domain lists require you to contain those critical pieces of information that increase the value of your data.   So now when you go into the portal, you can extract that information or purchase that information because those fields are populated.   And the tools make it easier for you through that pick list of common names for chemicals and analytical methods.

And with that I want to say that you are not alone when you venture into either the WQX or the WQX Web tool.   We have our help desk that can provide you assistance.   We have monthly user calls where you're able to engage with other data owners so that you can ask questions and to engage on new tools or updates to our system and we have the STORET list serve which again, provides another mechanism that you can receive updates as well as ask questions and share information.   And our website is listed below and that will be made available on the slides.   And with that I will take any questions.

Anne Weinberg

Okay, thank you, Charles, for your presentation. We are going to have time for questions for the audience but first I want to make a few announcements.

So this is Charles Kovatch. You probably know him. And these are pictures and contact information for all of our speakers if you have any questions that was not asked today we have quite a bit of time left so we are going to be asking quite a few of the questions. So I hope we will get too many of them. But if you do not get your question asked please do feel free to contact these folks and ask the questions. Again, Susan Holdsworth, Nate Booth and Charles Kovatch are the three speakers. And I also want to announce our next webcast will take place in November. It will highlight our new application that we recently released on the anniversary -- the 40th anniversary of the Clean Water Act. The app is called: How's my waterway?. And we are going to feature that app that you can find on the website now. And we will also feature some other water quality apps in November. So registration for this webcast will be posted at www.epa.gov/watershedwebcasts. So please check on that website regularly. At the beginning of November we will have an announcement of that date and registration will be open. So please join us for that one.

And finally, I also want to encourage you to download your certificate. The certificate can be downloaded at this URL. You can type in your name. If there are multiple people in the room you can type in multiple people's names and personalize their certificate for everybody that is watching at your location.

I also want to encourage you at the end of the webinar there will be an evaluation form and we encourage you to fill out that survey to let us know what you thought of today's webcast. Give us your feedback. We really do appreciate your feedback.

And now at this time we are going to have a few questions for Charles that Michael Brennan will lead us through and also some of the other speakers. Okay.

Michael Brennan
Thanks for your presentation, Charles. And thank you, everyone, for your questions. I guess first I would like to revisit HUC field and WQX and the Water Quality Data Portal. Charles, does the HUC field pull into the portal that supplied that generated the HUC field in STORET or does it supply the user supplied HUC field in STORET?

Charles Kovatch
Thanks, Michael. That's a good question. The HUC field within STORET is generated from STORET. When a user submits their data via WQX, once it is in

STORET we are able to identify through the location of coordinates from the station which HUC that station resides.   So STORET is able to generate the HUC field.

Michael Brennan
Thanks.   And can you shed some light on the whole process of how data goes from WQX into the Water Quality Data Portal.   Do the users have to push the data to the Water Quality Data Portal or does it just automatically happen?

Charles Kovatch
For all intensive purposes you could say it automatically happens.   It does not reside on any of the -- any other efforts from the data owner.   Once data is submitted then STORET is able to take care of that next step.   I can explain a little bit of what happens, not to get too much into the details, but once the file is submitted, EPA's STORET has Web Services which the Water Quality Portal, and Nate mentioned these earlier, that the Water Quality Portal calls and that's how we are able to help get some of the information out of STORET and make it available through the portal.

Susan Holdsworth
I guess I just want to add one caveat.   I think from the user's perspective if you have flagged your data as provisional or draft, it will not make it all the way to the STORET warehouse and the portal.   It has to be final data.

Charles Kovatch
That's a good point, Susan.   And actually we would like to stress that point again.   STORET offers -- WQX offers the option for users and data owners whenever they are submitting their data fields, their data files, some of the fields can be listed as being provisional data like for some reason they still needed to go through some sort of review or another QA step before they wanted to go as a final data submission.   STORET will not publish or make available any data that is listed as provisional.   Only data that is listed as a final.   It's a nice option to have, any data owner must be aware that field must be checked as final before it will be made available on the Water Quality Portal.

Michael Brennan
Great, thanks.   Now that we are on the topic of data submitters, can you comment on who can submit water quality data?   Does the organization need to meet a certain criteria be it a nonprofit, be academic, be a state?   Do they need any specific qualifications in order to upload their data?

Charles Kovatch
That is a really encompassing question.   I will try and break it down.   First,

anybody can participate to submit data via WQX and WQX Web.   We already have a wide variety of organizations represented within STORET warehouse. We have some federal agencies like the National Park Service.   We have of course, EPA regions and research laboratories.   We have a number of tribes, counties, state environmental organizations.   We have watershed groups and we also have universities that participate within STORET.   So in order for you to become able to submit your data, you would have to establish an organization ID.   So you register with us and we will provide you with that ID that you will be able to submit your data.   Kind of like with your checking account with your bank.   You want to make sure that you are linked with a number that the funding will go to your account and not somebody else's.   The same thing.   You have to have that unique identifier for your organization in order for to go into the STORET.

Susan Holdsworth
So in summary, anyone can submit data.   You just have to request an org ID and be able to align your data with the common data elements.

Charles Kovatch
Exactly.

Michael Brennan
Great, thanks for that.   Let's take a step back and ask Susan a question real quick or if Charles wanted to take this on you can.   How do you see local environmental health and public health departments utilizing the Water Quality Data Portal?

Susan Holdsworth
Well, I think that local government, local state health departments are a source of data coming into the portal through the Water Quality Exchange.   Already a number of public health departments submit data through the beaches program where they are monitoring beaches for particular water quality parameters such as bacteria indicator organisms.   And I certainly hope that they find this portal of value to them in that they can get access not just to their data but to data from other sources of similar or other correlating variables that may be of use.   And then thirdly, I am hoping that local organizations may be able to build some tools like those that Nate showed which might display some of the data for their user community if they are posting beach closures or things like that using polling data from the portal through the Web Services which is a machine to machine communication they may be able to support web notifications for their users. And I would encourage anyone who is interested in that kind of thing to come back next month and see the webinar on the applications, the How is my Waterway app.   And I think we have a beach quality app that we are hoping the

developer of that would demonstrate as well.

Michael Brennan

Thanks. If we could ask another question from Charles. Charles, are the biological data that are now being accepted via WQX are they going to be available via the portal?

Charles Kovatch

That is a good question. We are working on making biological data available through the portal. We have obviously -- we have heard there's great interest in this and I can tell you we are working on it. I'm not able at this time to provide that timeline. We are hoping sometime in 2013, but I don't have an exact date on it.

Susan Holdsworth

Right. You might recall that was on the list of future enhancements that Nate reviewed. As he said, we are working with the water information strategies workgroup under the National Monitoring Council to identify priorities for enhancements to the portal and of course, we have to live within the funding constraints that all of us are facing now. But that is definitely something that we want to add as soon as is feasible. So appreciate your feedback that that must be important to you.

Michael Brennan

Great. And let's get Nate back in the mix. Nate, can you comment on groundwater in the Water Quality Portal? I know you have presented a little bit on USGS's other initiative for groundwater flow, but do you have a sense of the percentage of data that is surface water versus groundwater in the Water Quality Data Portal? And is there a way to indicate information such as aquifers whether there is a well screen, the depth of the boring in the Water Quality Data Portal?

Nate Booth

This is Nate. Those are good question. So there is quite a bit of groundwater information. I would say three quarters or more of the sites from USGS and NWIS are groundwater wells. Maybe half of the water quality results are groundwater. There are not many water quality results for groundwater from STORET within the portal currently. In terms of what information related to groundwater is available through the WQX format, some of that information is available. I would have to check to see which of all of those fields is accessible. I think there is certainly an indication of which aquifers sampled. I think there is -- I think that well depth is a field as well as screened interval, but I don't have that information right in front of me. But we can follow up with you if you are

interested after the call.

Susan Holdsworth

Hey Nate, this is Susan.   Can I ask a follow-up question?   Should we be kind of steering people toward the groundwater portal that you were describing earlier?

Nate Booth

I think that if folks want to access groundwater water quality data comprehensively, come to the Water Quality Portal.   The National Groundwater Monitoring Network is a subset of monitoring locations from various organizations that meet the criteria of the groundwater monitoring network nationally.   So it's not comprehensive.

Michael Brennan

Alright, thanks, Nate.   Also on the topic of data in the portal, can you comment on whether or not there is microbiological data such as pathogen counts available in the portal?

Nate Booth

Yes, microbiological data including pathogen counts are available within the portal.

Michael Brennan

Thanks.   And let's switch back to WQX real quick.   Charles, is there any sort of vetting process included or QA imposed on people submitting water quality data to the Water Quality Exchange?   Are there any QC fields available where you can flag data if there is any QC issues?

Charles Kovatch

We get this question often.   People are interested in the integrity of the data for which they are querying.   First, I would say we work a lot with state data and tribal data as well.   And those organizations have what is called a QAAP, a Quality Assurance Approval Plan.   And those plans layout the quality control protocol that those organizations take and follow whenever they are collecting, analyzing, and I guess you could say reviewing their data.   So that whole step is conducted by the data owner.   What we do have, we have fields within the WQX which one, signify that this was a quality -- this was reviewed quality control confirmation.   There are also comment fields that enable someone to say if there was any sort of error or anomaly or whatever it might be with that file or that result rather that they can populate that comment field with the description.   We also have required fields so if you are submitting a result, we had five copper, you must provide units so you just can't leave copper five standing alone.   It has to be part per billion or part per million or whatever it might be.   So we have

some options within the schema.   We have some required fields.   However, the bulk of QA QC is the responsibility of the organization and that's why it is so important that those organizations provide the contact information for who collected it and which methods, etc., etc. so you can go back and contact them and ask them about those results.

Michael Brennan
Great.   While we are on the topic of QA can you also expand more on non-detects and how they are handled when someone analyzes the water quality sample and are not able to measure the results from it?

Charles Kovatch
What do they want to know again?   Say that again.

Michael Brennan
Can you submit non-detects?   Can you indicate a measured cadmium in the water quality sample and I got a non-detect?   I got a non-value that did not show up in the measurement.

Charles Kovatch
We have, I mean, if you have a sensitivity range on your instrumentation.   You have a sensitivity range on your instrumentation so obviously it's quite possible you are going to have a non-detect.   So we can accept a field, a record that says non-detect.   We have the capability to show less than whatever that range might be like five part per billion.   So we can capture that information.

Susan Holdsworth
I think what's important to note is that if you are reporting a non-detect as a result, you have to report the detection limit associated with that result as I understand.

Charles Kovatch
That's a good point.   Yes.

Michael Brennan
Great.   I guess moving on we could switch back to the portal.   Is legacy data, data that was submitted to the EPA before 1999?   Is that included in the portal and are there any plans to include it in the future?

Nate Booth
I'm not sure if that was directed at me.   This is Nate, but I'll answer it.   Legacy STORET is not part of the portal currently.   That is something that we have heard from the community as a request.   It is not in our current work plan but it is

something that has been discussed and so if that is of interest to your organization we would like to hear about that.

Susan Holdsworth
And this is Susan.   I will add one of the barriers we have been struggling with the legacy data is that much of that data does not have the metadata documentation with it.   The, what I kind of call the data quality documentation to allow things like method, detection limit, much of the legacy data does not have that documentation with it.   And that prevents it from meeting the data standard that was adopted by the monitoring council and is the basis for sharing data from these two major warehouses.   It would probably involve us going back to those data owners and seeing if they are willing to dig up the documentation about those data or figure out some other workaround for including data that is not of documented quality.   Which kind of goes against our principles here with the data element based sharing of data of documented quality.

Michael Brennan
Thanks, Nate and Susan.   Another question for Nate.   Nate, you provided some examples of using our studio in your presentation.   And we had a question whether any of the can codes that you referenced would be available for the public.

Nate Booth
Yes, definitely.   So I am not sure, I think the slides will be provided and within the slides you will see the URL for the package that I showed.   So that is available publicly.   It is work in progress so the material available on the web currently has been approved and is out there but there will be additional functionality in the future.

Michael Brennan
Thanks, Nate.   Also, can you comment on other options for downloading Water Quality Data Portal information besides getting it in Excel or a text file like options to receive information through REST services?

Nate Booth
Yeah, sure.   So REST is a particular web service protocol.   And if you were to look at the Web Services User Guide it will give several examples of how to construct REST web service request for accessing the portal.   And besides tab delimited in Excel format as well as KML you can access the raw WQX XML data format out of the portal using those web service calls.   And that's the way that many of those applications that I showed access the portal currently.

Michael Brennan

Thanks, Nate.   And let's push it back to Charles.   Charles, can we assume that any data currently in a state STORET system is automatically updated to the EPA STORET and hence to the Water Quality Portal?

<u>Charles Kovatch</u>
The answer to that question is no.   We cannot assume that any state data is automatically populating STORET and then being made available on the portal. In fact, that is one of the things that we are working on is trying to get more state data to be available through the portal.   As you saw from that map which was the representation of the data that is available from the portal, there are states that were darker colors and some states that were lighter colors.   Those states with lighter colors, we are trying to reach out to them to try to help them submit that data.   For it to be automatic, that is definitely a vision, maybe I don't know with increased technology that might be more available.   But right now the answer is no.   Because of just the different states or different state partners have been able to submit that data and others we are still working with them.

<u>Michael Brennan</u>
Great, thanks.   And Charles can you comment on the types of data that WQX will submit?   Will it submit the sediment data?

<u>Charles Kovatch</u>
Yes.   And in the pull-down menu in one of the lookup tables that we showed in one of the media was water.   You can also select sediment as a media.   So the answer to that is, yes.   And we have sediment data within STORET.

<u>Michael Brennan</u>
Great.   One more question on WQX.   So with this adoption of the portal, does that mean that existing WQX STORET page will disappear and does that means this is getting phased out?

<u>Charles Kovatch</u>
Well, that is a really good question.   I think right now we have been happy to work with USGS and the National Water Monitoring Council to create this great product, the Water Quality Portal.   STORET webpages, we still have a lot of information on there that is relevant to WQX and WQX Web.   Those will remain. We still have our warehouse query capabilities on our webpage and for now those will remain.   And we still have a lot of information on our webpage about the different web services that are available for users to query data out of STORET.   Right now those are going to remain.   So this is still a new product that we are very excited to be making available.   And we still have a lot of information on our webpage that is very relevant in order to help people to participate in the portal.   So for now all of those capabilities will remain.

Susan Holdsworth
I think I would also like to point out that the portal and STORET warehouse are not yet completely redundant.   Maybe if we get to that point then we could talk about perhaps really merging them and retiring the STORET warehouse.   But right now the biological community types of data and the habitat characteristic types of data are not available through the Water Quality Data Portal that the Council is hosting.   And data submitted through WQX of that type is only accessible through the STORET data warehouse.

Michael Brennan
Great.   I think that about covers it for questions for this webinar.   If we did not get to your question today, we will follow up with you afterwards.

Susan Holdsworth
I guess I would just like to thank everybody who participated, to thank the speakers, and everybody who asked such great questions to help us clarify the role of the warehouse, the role of the Water Quality Exchange for submitting data, and providing us some insights in how you might use the data.   I thought this was a really valuable exchange and again, I would like to encourage folks to visit the portal on the Council website, select a link to getting data in, how to upload data, and that will take you over to the Water Quality Exchange site and to contact any of us with questions.   And remember we do have user support services for WQX in the portal that you can access.   I will turn it back to you, Anne.   Thank you.

Anne Weinberg
Okay.   At this time I would like to conclude today's webcast.   Thank you, Susan, Nate, Charles for participating today.   And of course, thanks to everyone who joined us.   That ends our webcast for today.   Thanks again.