**Water Information and Data Subcommittee**

**Data Exchange Workgroup (Workgroup #4)**

Minutes – February 22, 2012

**Attendees**: Steve Tessler (USGS), Dharhas Pothina (TX), Sara Larsen (WSWC), Laura Paeglis (NE), Steve Malers (Riverside), Jeff Hogan (NE), Becky Fulkerson (Reclamation), Dwane Young (WSWC/WestFAST)

**Next Call:** Week of March 5th, look for a Doodle Poll

**Administrative:** The minutes from the prior call were approved. Several of the workgroup members said that they can’t see the workgroup 4 folder in the Google Docs. Dwane said that he would look into that and resolve it. **Action Item: Dwane make sure everybody has access to the Workgroup 4 documents.**

**Review Issues Document:** Dwane and Sara reported on their visits to New Mexico and Oregon. They discussed the security issues with both of these states. In both cases, the states felt that the data could be open to the public (since a lot of it already is). They also reported that in Oregon, the state felt that if the WSWC could define a standard approach for sharing data, that they would be able to build their own components to provide their data via web services.

Security: The workgroup continued discussion of the issues document. Steve Malers asked about how we would deal with provisional data or data that had not received a thorough quality check. Discussion ensued on what agencies put out provisional data. USGS marks provisional data as such, but still makes the data available. NRCS does not. They wait until the data have been reviewed before making the data public. The group felt that so long as the schema allowed for data to be flagged accordingly, there wasn’t really a problem with provisional data if the states wanted to publish it. Sara asked TX how they dealt with the local caching of data from their data partners. Dharhas said that the biggest issue has been whether or not the partner has the IT infrastructure to run the services and not unduly burdening their partners. In many cases they don’t. They have been working with TCEQ to have them host the services to serve up their data. Under an ideal situation, the partner will run the services directly off of their database. However, if that’s not an option, then TX will either get a copy of their data, or retrieve it off of a website. They then will run the services from a local cache.

Dwane asked the group what would be needed to allow for a data partner to build out their own components. Dharhas responded that with the CUAHSI implementation they created a specification that defined how requests would be made, what services would be implemented, and what format (schema) would be used to exchange the data. For the REST implementations, they’ve also found value in defining what the REST interface looks like (how does one construct the URLs to retrieve the data). For SOAP services a WSDL would be developed that everyone can conform to. Dwane asked how the group felt about SOAP vs. REST services. Without the need for a robust security model, REST services could probably meet the needs of this project, and would be easier to implement. The group agreed that we should build towards REST services as opposed to SOAP services.

Steve Malers stated that one of the challenges that he’s faced in using services is that the catalogs are not well defined. He felt that one of the benefits of this project would be to make some progress on the basic catalogs.

Dharhas also reported that the WaterML 2 schema is close to being through the OGC process, and will soon be a standard.

Data Exchange: The group continued through the document. The next topic was data exchange. Dwane said that this topic has already been covered by the group in the discussion of the overall architecture of the system.

Implementation: Sara discussed her current experience in working with both Exchange Network components and the TX services that were shared by Dharhas. So far, her experience has been that it will be easier to implement the TX services. Dwane said that his original thinking on leveraging the Exchange Network was based on the idea that the states would be able to leverage the components that have already been developed in the state environmental quality offices. All of the states in the west already have an Exchange Network Node. However, after further discussion with some of the states the impression is that it would be difficult to make those Nodes work for other agencies in the state. Additionally, the Exchange Network provides a very robust security model that may not be necessary for this data exchange. With those considerations, the group feels that the best approach would be to learn from these other data exchange efforts, but to define our own standards for communication. We will leverage existing standards where appropriate.

Maintenance: The group discussed how the system would be maintained. The issue is that once the pilot system is done, WSWC will likely not be the long-term home for the central portal, but may be the home for the standards that are developed. There are a couple options. The first will depend on what USGS does with the Water Census. One of the components of the Water Census is to develop a central portal that is consuming relevant data. If that portal were to become a reality, it could replace the central portal being developed by the WSWC. Another option would be for a state to host the central portal. Dharhas suggested that whatever is developed be developed in an open source way to help with the long-term maintainability of the system. It was also mentioned that the larger issue will be getting the partners to continue to provide data. Dwane mentioned a concept that is used by the Exchange Network called a Trading Partner Agreement (TPA) where both parties sign an agreement to continue to share data at specified frequencies. Dharhas said that TX has a similar document that they call a Memorandum of Agreement (MOA). He mentioned that this increased the program’s visibility to upper-management groups. The group felt that this would be a valuable piece to the data exchange to have on paper that the parties agree to share data.

Governance: This will define how changes will be made to the specifications and data standards that are developed. Dwane described the very robust governance structure that is part of the Exchange Network. He described two levels of governance. One at the management level, and one at the technical level. He also discussed a third level of decision making that happens at the project level in an Integrated Project Team (IPT). He felt that for this project, we are currently in an IPT –type governance, with the workgroups being individual IPTs. As the project matures, we may want to consider formalizing a more robust governance group. The Water Information and Data Subcommittee within the WSWC could serve as the management level governance. We would need to form technical level governance. Dwane agreed to write up a proposed approach for the issues document to governance. **Action Item: Dwane will write up a proposed approach for the governance section.**

As a homework assignment, Dwane asked the members of the workgroup to consider the question: “What incentives need to be offered to encourage participation by states and federal agencies?” We will discuss this further during our next call. **Action Item: All workgroup members will consider what incentives would need to be offered to get people to participate in the data exchange (i.e. how do we answer the question “Why would my state do this?”).**

Dwane also asked for the workgroup member s to be thinking about what the services would be that we would define for the system. A couple examples are in the back of the issues document. **Action item: All workgroup members consider what services would need to be defined for the system.**

**Call Schedule:** The next call would be scheduled for the week of March 5th. Dwane will send a Doodle Poll to the workgroup to check on their schedules.