**Water Information and Data Subcommittee**

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

Minutes – March 9, 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:** Combined with Workgroup 3, call to be held March 27th

**Administrative:** The minutes from the prior call were approved. Dwane added everybody individually to the Google Groups to make sure that they had access. Laura and Steve reported that they now have access to the Workgroup 4.

**Review Issues Document:**

Incentives/Participation: Dwane showed some of the incentives that WSWC has come up with for why the states should participate:

* The WSWC has a strong history of protecting the interests of the states.
* States are often bombarded with questions or requests for their data. By making the data available, in a common well-defined format, they can save time in dealing with multiple requests. Dharhas gave a good example of this in Texas where they were responding to requests for drought and lake level data every day. They put the data on the web, which has save them a tremendous amount of time.
* Many other third parties are asking the question of ‘How much water is available’. The states should be the ones giving the answer to that question.
* States are struggling with the same issues of determining the best approach for determining water availability.
* States are struggling maintaining staff to support home-grown models. Need a more sustainable approach.
* Help states get ‘beyond their own boundaries’. Allow for better Regional collaboration.

Dharhas explained how they have river basins that collect data, and the river basin doesn’t have a lot of motivation to share data. What they’ve found is that sharing the data has to be as painless as possible.

* Now is a good time to engage in sharing the data because it could be made as painless as possible.

Steve Malers talked about the resource issues that the states have, and we need to address the issue of how this helps the states, and is not just an extra burden. Part of the message can be that if a state does have a resource issue, and they can’t implement it now, then that’s ok. As long as they participate in the standardization process, and then they can implement it later when they are in a development mode on their system.

Nebraska feels that they don’t see why they wouldn’t do it. Since they’re in the process of re-designing their system, there’s no reason to not engage in the standards process.

Services

Dwane led a discussion on some potential services that would be needed to support a central portal application. He and Sara brainstormed on some of the potential services. Dharhas asked if state, county, and HUC would need to be part of the database, or could it just be calculated. Dwane said that he didn’t see a problem with deriving those values based on locations. Dharhas also stated that although lat/long boxes are useful, it might be more useful to have an input parameter as a polygon. This would be accomplished by providing a series of vertices in order that define the polygon. Steve asked why you wouldn’t pre-calculate the polygons, and then allow the users to select by the polygon. Dharhas answered that in many cases they’re not basins, and it might limit the user.

Dharhas asked if we should also include other types of spatial queries (i.e. climate divisions). Dwane responded that the catalog service is meant to provide summary level data for reporting units. The first service will return the reporting units for a given state. This would allow a user to further query the data based on those reporting units. Dwane stated that the catalog services would most likely always be called for the entire state.

Laura asked for a better definition of what a reporting unit is. Dwane explained that a reporting unit is any grouping used by the state at which the state summaries their information. For example, a state may say for a given watershed or region what the water availability is in that watershed or region. For example, Oregon summarizes their information by Water Availability Basins (WABs). Utah has basins similar to a 10-digit HUC, the USGS Water Use Report summarizes everything by county. A reporting unit is just a basic accounting unit by which information can be grouped and summarized.

The group then began a discussion of ‘time-series’ data, and if the ‘derived’ data as part of this project could be considered time-series data. Steve and Dharhas stated that if the data has a value and it’s associated with a time (for example, consumptive use for a particular month), then it would be considered time-series data. Dwane felt that his view of time series data was closer to what we would see for observed data (i.e. a streamgaging station). He felt that the derived data as part of this project was significantly different from the observed data. Dwane stated that his real concern was that if we call this time-series data, then it could be argued that WaterML 2 would be a reasonable fit for the derived data. Dwane didn’t feel that this data model would really work for the derived data. Dharhas agreed that WaterML would likely not work. However, the group felt that there may be some changes to WaterML 2 that might allow it to work. Dwane said that it would be worth-while to check this out, but he felt that the derived data has enough extra structure, that it would be too much of a stretch to try to stuff that in WaterML 2. Steve suggested that WaterML could be expanded using our own namespace, and that USGS already does this for some of their services. Dwane concluded the discussion by saying that this discussion was straying into more of a Workgroup 3 discussion, and that perhaps the best way to move foreward was to combine Workgroups 3 and 4 for the next few meetings in order to allow a full discussion of the data elements and the services together. The group agreed to this. Dwane said that he would send the group the doodle poll for Workgroup 3.

Dwane continued to explain the services. The WSWC is envisioning 2 sets of services. The first 2 services would be catalog services that would provide a summary of information. The second two services would be detail services that would return individual allocation information as well as individual reporting unit summaries.

As the group continued to discuss the data, Dharas commented that one of the challenges would be that the data may be split between two agencies in the state. He gave as an example that in Texas, the Water Development Board has the planning data, but TCEQ has the allocation data. Dwane stated that this may actually be somewhat common, and agreed that this will present a challenge.

Dharhas asked if the term ‘site’ on the services was necessary. He said that it caused some confusion, because when he thought about ‘site’ he thought about monitoring locations. Dwane suggested that perhaps for the GetSiteCatalog service, we instead call it the GetAllocationCatalog service.

Steve suggested that in order to really understand the data model, we needed to see if we could fit into it some actual data. Dwane agreed with this approach, and stated that at this stage, we’re still trying to figure out the conceptual model. The next step would be to design a schema based on that, and then we would definitely need to test that model and schema against some real data to see how it fits. It would probably be good to try a few different states’ data to see how the model works for the different states.

**Call Schedule:** The next call will be a combined workgroup 3 and workgroup 4 call. It will be scheduled for the week of March 23rd. Dwane will send the Workgroup 3 doodle poll.