

**Ground-Water Status and Trends Analysis & Reports Meeting
San Antonio, Texas, April 20-22, 2005**

Data Warehouse Session

This summary written by Ken Skach
(Session presented by Ken Skach, Terri Arnold,
Pete McMahon, and Kelly Warner)

Objectives of the Data Warehouse session:

- 1) Familiarize the PA authors with the power of the Nawqa Data Warehouse (DWH), possible approaches and successful strategies for retrieving desired data from the DWH, and potential pitfalls and quirks that commonly cause problems with DWH retrievals.
- 2) Present a proposed system of PA Netcodes to update previously retrieved DWH datasets, and present a proposal for writing a Principal Aquifer Fact Sheet.
- 3) Decide whether the PA teams want/need "summary report type" graphs to compare PA data to a "national" dataset.

Results and Guidance, by Objective:

Objective 1) Familiarize the PA authors with the power of the Nawqa Data Warehouse (DWH), possible approaches and successful strategies for retrieving desired data from the DWH, and potential pitfalls and quirks that commonly cause problems with DWH retrievals.

About half of the session was devoted to this objective (see the session's PowerPoint presentation). Topics presented include: approaches to data warehouse queries, common pitfalls and surprises, crosstabs and rollups, GIS-derived ancillary data sets, the data warehouse's GW-POSV Patch tools, and the process for fixed data to be updated to the data warehouse.

Decisions and Recommendations:

Consensus: Instead of publishing "Citable Queries to the DWH" (which would retrieve updated data as years go by), we decided that each PA team should **freeze a static dataset** for future reference, to be maintained outside of the DWH, for each report product.

Consensus: Coordinate updates for Principal Aquifers data (i.e., fixes that must be made by study units) **through Lori Apodaca.**

Recommendation: GIS-derived Ancillary data is driven by needs. If you need ancillary data more current than what is presently served in the DWH or on Kerie Hitt's GIS-Ancillary

data web site, **contact Kerie** to make your needs known, and to learn how feasibly your data needs may be satisfied.

Recommendation: Ken will work with Kelly Warner, to develop data checks for aquifer codes, in the DWH.

Recommendation: Since sub-queries are important, the Wisconsin DWH team will continue to **press Oracle** to make sub-query functionality available in the Applet version of Discoverer. As a work-around, the Wisconsin DWH team can **pre-build a sub-query** into sample-level flags (such as flagging samples containing a particular schedule-indicating parameter) if such a flag would be commonly useful. Also, the **Windows Client version of Discoverer** (which does have sub-query functionality) is downloadable from the internal DWH page; see the "Support Information/Install Help" link (then the "Client-Server Software Install Instructions") on the Internal DWH web site.

Handout: General DWH Info and Web Sites:

DWH TRAINING:

- **DWH Cyber Training: June 1-3, 2005**
 - 1:00 – 3:00 EDT, each day for 3 days

WEB SITES:

- **Internal DWH:**
 - http://infotrek.er.usgs.gov/nawqa_www
- **Public DWH:**
 - <http://water.usgs.gov/nawqa/data>
- **Kerie Hitt's GIS-Ancillary Data Web Site:**
 - <http://water.usgs.gov/nawqa-only/gw.html>
 - (Datasets documented via GIS metadata)
- **POSV Coding Guidance, found with NAWQA Sample Coding Guidance, at:**
 - <http://ar.water.usgs.gov/nawqa/sample-coding/outline.html>

Objective 2) Present a proposed system of PA Netcodes to update previously retrieved DWH datasets, and present a proposal for writing a Principal Aquifer Fact Sheet.

The proposed list of Principal Aquifer Netcodes was found unsatisfactory to the group at large. Many teams will need more than one netcode, to store more than one grouping of sites for their different purposes, including their different report products. Terri Arnold also proposed a Principal Aquifer Fact Sheet.

Decisions and Recommendations:

Consensus: A committee will re-work the PA Netcodes list, and see if the larger group of PA teams will approve of it. The committee includes **Terri Arnold, Ken Skach, Wayne Lapham, Kelly Warner, and Chip Hunt.**

Consensus: a group was formed to finalize the PA Fact Sheet: **Terri Arnold, Wayne Lapham, Kelly Warner, a Study Unit chief (yet to be chosen), a Principal Aquifer Team Lead (yet to be chosen), and Pixie Hamilton.**

Handout: proposed PA Netcodes:

The proposed netcodes are not included here, since they were deemed unsatisfactory and must be re-vamped.

Objective 3) Decide whether the PA teams want/need "summary report type" graphs to compare PA data to a "national" dataset.

A Mock-up graph (see the session's PowerPoint presentation, the next-to-last slide) comparing High Plains data to national data, was presented by Pete McMahon. Discussion followed concerning whether all the PA teams would want some comparison graphs in their report(s), and if so, whether we need to ensure that they all have the same "look & feel," and that they all compare against the same "national data set." It was pointed out that some PA teams are still collecting data, so a "national data set" cannot be generated in time for the High Plains report this summer. Also, many PA teams were not sure at this point whether they would even want to do comparison graphs.

Decisions and Recommendations:

Consensus: The High Plains team will make their own graphs to meet their own needs. Any PA Team which wants comparison graphs will have to **design their own**, and **generate their own** comparison dataset **for their own needs**. Acknowledging that comparison datasets may not be equivalent between PA reports, each PA Team must document their own graphs completely, and such comparison graphs will have to stand-alone from any other team's graphs.