

## ***EarthChem/NAVDAT Workshop on MexDB***

University of Kansas, Lawrence, Kansas 66045

April 20, 2007, 8:30am to 2:30 pm

**Report compiled by Doug Walker and Luca Ferrari.**

### Participants:

Doug Walker, Susanne Straub, Luca Ferrari, Jason Ash, Eileen Jones

### Background

EarthChem is a community driven project to facilitate the compilation and dissemination of geochemical data of all types. The project is active at building a home for future data contributions by working with authors, societies, and publishers as well as government organizations. In addition, the EarthChem project responds to community needs to facilitate compiling and serving data.

One of the aspects of EarthChem is to assist contributions of data to the system from existing databases and to provide expertise to other groups assembling databases. One such group is headed by Luca Ferrari and Susan Straub in an effort called MexDB. This is a database project focused on the Trans Mexican Volcanic Belt. The data are being compiled to compare volcanic geochemistry with subduction parameters and eruptive volume. The MexDB workers envision building a relation database and web interface for this system. At present, there is information on about 3000 samples from this zone. Over 2000 of these are currently contained within the NAVDAT database; however, the data for these samples does not contain the detailed subduction and volume parameters and is lacking in some of the metadata for methods and ages.

To help move the MexDB forward, a workshop was held to identify a clear path for interaction of this database with NAVDAT and EarthChem. Because of the close existing connection between NAVDAT and these workers, NAVDAT personnel will play an important role in aiding the creation of the database. EarthChem workers will provide some expertise and ensure interoperability between MexDB and the EarthChem system.

### Goals

The goals of the meeting were to address important issues on the interoperability of the MexDB system with NAVDAT and EarthChem and to identify ways for these groups to interact in a mutually beneficial way.

- a. Identify how much information needed for MexDB currently resides in NAVDAT.
- b. Determine how best to get existing information to the MexDB group.
- c. Establish ways to add additional compiled data.
- d. Plan for interaction of MexDB with EarthChem.

## Workshop Summary

There is clear overlap between the current holdings of NAVDAT and the planned MexDB database. The MexDB has compiled about 3200 samples from the study area. At present, about 2200 of these are in NAVDAT. These samples are missing some critical metadata, including much of the information about analysis methods, analysis institutions, and detailed references for methods. In addition, there is conflicting age information for some samples.

MexDB workers are critically interested in connecting sample information to subduction parameters and vent characteristics. A database for these aspects already exists, but it is not yet linked to the samples. This database contains a unique identifier for each volcanic vent. These, and a MexDB defined unique sample identifier, are also critical data to be linked to the sample data.

The MexDB database will be built and hosted at Centro de Geociencias UNAM, campus Juriquilla, Qro. The system is being created using PHP and MySQL. There are sufficient resources to manage and maintain the database. A prototype of the portal for MexDB is at: <http://dragon.geociencias.unam.mx/mexdb/>

## Approaches

It was decided that the most efficient way for to MexDB workers to create their database and ensure its contribution to the EarthChem is by using the NAVDAT system for several aspects of database development. Participants envision that the workflow given in Figure 1 (see end of this document) would provide the most efficient path for all the participating efforts.

1. The MexDB team will work with the NAVDAT by contributing data as text files for upload. This is quite efficient in that about 75% of MexDB data is already in this format. The MexDB team will add all additional critical metadata (methods, vents, etc. – lower left part of Figure 1).
2. Data will be uploaded after additions and quality control to the NAVDAT database, but be clearly identified as MexDB (as they will in the EarthChem system).
3. The MexDB team will create relational database tables for the vent and subduction parameter information (lower right part of Figure 1).
4. NAVDAT will create the core database tables in MySQL format for the MexDB team (top part of Figure 1).
5. The MexDB team will then link the tables provided by NAVDAT to the vent/subduction parameters table using the vent identifier (center box on Figure 1). This linkage will complete the MexDB database.

Data in NAVDAT will be identified as MexDB. When sent to EarthChem, these data will be labeled as MexDB.

## Next Steps

Considering the approaches listed above, the participants recommended the following next steps to move the process forward. We anticipate that these steps will be done over the next 12 to 18 months.

- 1) Send NAVDAT schema to MexDB programmer. This will ease the process of creating MexDB as well as ensuring that the MySQL files will be easily used. This is also important for adding the level of reference and age information required by MexDB. Marina Manea and Jason Ash will do this task soon supervised by Susanne Straub.
- 2) The MexDB programmer will send the MexDB schema to NAVDAT (once modified) to be used to generate MySQL output from NAVDAT. This can be done by making a backup of the database and sending it to NAVDAT. Marina Manea, Susanne Straub, and Jason Ash will be responsible for this.
- 3) MexDB workers will add the following information to existing files from NAVDAT:
  1. Method information
  2. Institution numbers (from NAVDAT)
  3. Any needed additional reference information for methods
  4. Resolve any conflicting age information (geologic age vs. numerical)
  5. Add the vent number to the samples (if possible, not required for upload)
  6. Add a unique identifier and assign this to each sample (MexDB-ID – not a required item for upload to NAVDAT)

Modified files, and addition files, will be uploaded to NAVDAT. MexDB will have permission level to delete a data file and upload it to the database, but will not have permission to populate files into the database (that will be the responsibility of NAVDAT workers). Luca Ferrari, Susanne Straub, and Teresa Orozco will lead this effort on the MexDB side, and Doug Walker will be responsible for final population into NAVDAT.

4) To ensure the population of and checking of vent information, NAVDAT will provide a table containing georef number, sample ID, location, vent number (if already assigned), MexDB-ID (if available), and IGSN (if available) for all appropriate samples (after population in the step above). MexDB workers will take this table into GIS and assign vent numbers to each sample. If vent is unknown, it will be assigned a point that can contain information equivalent to the vent information. Then the table is sent back to NAVDAT for population of vent number into the database. Jason Ash and Luca Ferrari will be in charge of this effort.

5) NAVDAT will prepare a vent editor tool. This will allow MexDB workers to correct vent assignments as needed. This tool will be connected to the sample profile page and be available to MexDB workers. The editing will be done on a map base presenting

geology, satellite images, and vent information. Eileen Jones will take the lead on creating the tool and Luca Ferrari will build a GIS layer with vent location and attributes.

6) To ensure permanence of the vent number with uploaded data, NAVDAT will create a table that preserves the georef number, vent number, sample number, and MexDB-ID/IGSN (if available) for each sample. In this way, if data are stripped from the database this table is maintained. Upon re-upload, that table is checked to get the edited/correct vent number. Jason Ash will do this task.

7) NAVDAT workers export data to relational tables to be taken into MySQL database in Mexico. MexDB is built/populated from these tables when MexDB workers join additional vent, volume and subduction parameters tables with the imported data using the vent number as the link/foreign key. MexDB maintains its own interface and search functions. Jason Ash and Marina Manea will take on this task.

8) NAVDAT or EarthChem workers will work with MexDB programmers to implement plotting tools (x-y graphs), glossary, and mapping tools (Google maps). Eileen Jones is responsible for this task.

Figure 1. Schematic workflow for interaction of NAVDAT, EarthChem, and MexDB.

