

Hydroinformatics

Assignment 4

Database Implementation and Loading Data

Due Date: October 13, 2016

Learning Objectives

1. Store, retrieve, and use data from important data models used in Hydrology such as the CUAHSI Observations Data Model (ODM)
2. Develop data models to represent, organize, and store data
3. Design and use relational databases to organize, store, and manipulate data

Computer and Data Requirements

1. Data required for this exercise are available in a zip file on the Canvas website for this assignment.
2. You may use any database management system to complete this project. However, the detailed instructions are provided using MySQL.
3. Although not specifically required, you will likely want to use MySQL Workbench to load data and examine the contents of your ODM database.
4. The ODM Tools Python application for Windows and Mac can be downloaded using the links on the Canvas website for this assignment.

The Problem

You are the data manager for the Logan River Watershed GAMUT network. You need a way to organize and store all of the observational data that is streaming in from sensors that you have installed in the watershed as well as other data that you have retrieved from additional sources. One of your colleagues has suggested using the Observations Data Model (ODM) developed by the CUAHSI Hydrologic Information Systems project.

Database Implementation Option1 – ODM on MySQL:

Using the data and metadata provided for monitoring sites in the Logan River and the attached instructions, prepare and load the Site, Variable, Method, and Source information into an ODM database in MySQL. Then, load time series data from the data file for the Logan River monitoring site at Main Street into the ODM database. Connect to your database using the ODM Tools Python software and explore it using the visualization functions to make sure that you have loaded your data correctly.

To test and explore the generality of ODM beyond the Logan River sensor data, find and load data and metadata for at least one additional site and variable into your ODM database (you may use any convenient tools for doing this). Examples might be discharge data for a USGS gage in the watershed or weather variables from a nearby weather station.

Database Implementation Option 2 – Your Schema on RDBMS of Your Choice:

If you choose to implement your own database structure using a different relational database management system, you will need to organize the tables of your database, load your data into the database, and present a discussion of the results. Were you able to successfully create tables that store all of the data you were provided? Were you able to create relationships to link your tables?

To test and explore the generality of YOUR database beyond the Logan River sensor data, find and load data and metadata for at least one additional site and variable into your database. Examples might be discharge data for a USGS gage in the watershed or weather variables from a nearby weather station.

ODM Analysis (Everyone)

Whether you implement your own database or use ODM, please explore the ODM design and provide a brief evaluation of ODM. In this assessment, you may consider:

1. Does the ODM database organize and store the Logan River sensor data, its associated metadata, and that of the additional data that you loaded?
2. What deficiencies in ODM did you observe when you loaded the subset of data into the database (e.g., is there information about the sites, variables, methods, sources, etc. that is missing in ODM)?
3. What are the advantages of using a standard data model like ODM to store your data?
4. What are the disadvantages of not using a standard data model for storing your data?
5. Were you able to complete your evaluation with the available ODM software tools? How could they be improved?

Deliverable

Provide a one-page briefing for your colleagues explaining the database system you used, the data you loaded and any challenges encountered. Provide a brief evaluation of ODM as a data management database for sensor data from the Logan River Watershed GAMUT network. In an appendix provide screen shots of the software used to conduct this project. For example, you could provide plots of the data that demonstrate the capabilities of ODM and the ODM software tools for your colleagues. If you use another relational database management system, you can provide other visual representations of the data as loaded in your database.