



SHARED EXPERIENCE: DATA MANAGEMENT

Water Quality Program Name: [Rivers of Colorado Water Watch Network \(River Watch\)](#)
Program Contact Name: Barb Horn, CO Parks and Wildlife Water Resource Specialist
Contact E-mail: barb.horn@state.co.us
Website: <http://wildlife.state.co.us/LandWater/Riverwatch/Pages/Riverwatch.aspx>

A Data Management Plan is a Priority for Success

Submitted by Barb Horn, January 2013

Colorado: River Watch began in 1989 with the goals of providing statewide, quality aquatic data while providing a hands-on science experience to foster environmental stewardship. The program is housed at the Colorado Parks and Wildlife Agency. We partner with a local non-profit that provides professional staff, funding, and programming. River Watch data is the largest data source for surface water chemistry, physical habitat, and macroinvertebrate data in Colorado.



Colorado River Watch Logo

There are 140 monitoring groups that annually supply River Watch with data. We collect data from about 650 stations on some 350 rivers. Water samples are collected monthly for 6 field indicators and 26 metals. We collect biannual measurements for 6 nutrients and collect annual information about physical habitat and macroinvertebrates. When combined, we generate huge amounts of data.

Before we started compiling these data, we were advised to have a data management plan as well as a plan for how to get these data out to targeted decision makers. We are grateful for that advice; it likely saved us time and many headaches. The worst thing you can do is ignore this critical piece. A plan is essential for turning data generated from monitoring into information into a measurable result.

The Colorado Parks and Wildlife Agency validates the volunteer monitoring data, administrates the data management system, and provides these data to decision makers. Thus, state-issued software limits what data management system is chosen. In 1989, that was system was Dbase for PCs. For the first few years, we provide computers to our volunteers to enter and transmit data into our system until PCs became commonly place in most homes. Currently, we have a Sequel server database that is hosted from the Department of Natural Resources. This system allows multiple users simultaneous access. It interfaces with a website to allow volunteers to enter their data and allows laboratories to upload the analytical results. The switch to the new host and data management model was a scary jump. It required that we release all control. It meant we had to hire someone to make

changes when needed, which adds to cost. However, it has allowed us to automate many validation steps and reduce human error in data entry and manipulation.

We are currently getting ready to move that data system into the “cloud”. This decision is driven mostly by political and budget changes, but also that our data is getting too voluminous to fully utilize and store. What this means is that we can retain security and hire a programmer to make changes to our database and user interface, but they cannot access the database. This transition has cost us tens of thousands of dollars more than it should have, but is supposed to save us money in the long run. Currently, we budget about \$10,000 for database operation and maintenance. That should be reduced by this move to under \$5,000. This move is risky if our agency decides it won't support this model in the future. We have navigated these technology changes, making hard decisions along the way. Our data is available on our website, uploaded to EPA STORET, uploaded to a statewide database of which EPA STORET data is a subset (called Colorado Data Sharing Network).



Photo from CORW: Volunteers sample for macroinvertebrates.

Another suggestion: You don't have to be a programmer to get what you need. Make a data flow map. Map all your sources of data, then how, where that data gets entered, validated and stored as raw data. Include all laboratories and all meta-data management. Then make a map of how that data is manipulated, moved and altered to generate analyses, information, dumps, files and reports for decision makers. Usually what is needed for raw data management is different than what is needed to generate information, but the first has to serve the latter. Then document this work flow and plan. You show it to a computer geek and they can design it for you. You will likely need a relational database tool (such as Access or Sequel server database) with a custom-built user interface.