Improving the Value and Utilization of Environmental Health Project Data

James M. Morrison, **MS, CIH, CHMM-** Wisconsin Division of Public Health, Madison, WI 53701

Wouldn't it be nice if environmental health projects data could be made available and evaluated for later use?
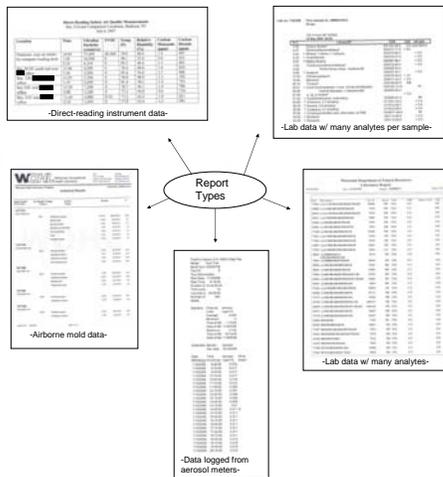
Data value could be increased by helping answer questions such as-

• How do the results I received for this sample relate to samples of other projects analyzed through the same method?

• Is this result indicative of a background value in this area or does it reflect an unknown contaminant source?

• Is this mold species often detected in homes with occupant health complaints?

• How do water and airborne benzene levels relate to incidence of childhood cancer in WI counties?

• How do indoor air quality (IAQ) parameters of moisture, airborne mold and fine particulates relate to incidence of asthma and allergy?
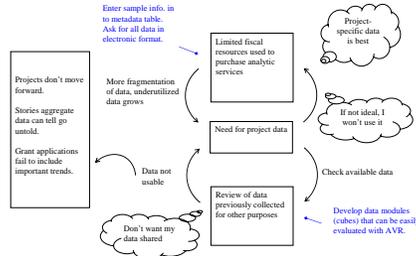
Without being able to retrieve, link and analyze past project data, we're limiting our perspective.

*One challenge is managing data from many different formats.*

-Direct-reading instrument data-

-Lab data w/ many analytes per sample-

Report Types

-Airborne mold data-

-Lab data w/ many analytes-

-Data logged from aerosol meters-

Archetype Description- Shifting the Burden

Enter sample info. in to metadata table. Ask for all data in electronic format.

Project-specific data is best

Limited fiscal resources used to purchase analytic services

If not ideal, I won't use it

Projects don't move forward.

Stories aggregate data can tell go untold.

Grant applications fail to include important trends.

More fragmentation of data, underutilized data grows

Need for project data

Check available data

Review of data previously collected for other purposes

Data not usable

Develop data modules (cubes) that can be easily evaluated with AVR.

Don't want my data shared

Each year, thousands of environmental samples are collected by or on the behalf of the Bureau of Environmental and Occupational Health. However, it is not managed in a way that it can be simply linked and compiled. This dynamic is illustrated in the boxes above. While many would support the idea of managing project data in a way it could later be used for other purposes, there are a number of impediments. Some of these are attitudes or mental models that describe concerns about moving forward with a metadata strategy (reflected in the balloon comments). Fixes are shown in blue.

Advances in data management methods such as Metadata and AVR can offer a solution.

What is Metadata?

Metadata is data about data.
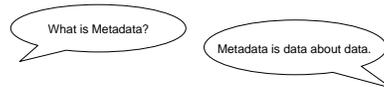
Metadata describes the content, quality, and context of a data resource for the purpose of facilitating identification and discovery. Through descriptive metadata a user can learn the what, why, when, who, where, and how for a data resource. In short, it links data sets and enables data users to evaluate utility of individual data sets.

Metadata helps address another big challenge- collecting and managing attribute information associated with the data sets. This enables researchers to evaluate utility and quality of data.

Analysis, Visualization, and Reporting (AVR) is a software package utilized by Wisconsin Public Health Information Network (PHIN) that provides the ability to analyze, display, report and map accumulated data and share data and technologies for analysis and visualization with other public health partners.

Some reporting types available with AVR-

WI Radon Map

County Level Air Toxics Data Map and Table Demonstrating Mapping Capabilities with Data Table

WI Poison Center Call Data 2001-2007

Drilling Down Further to Elemental Mercury Call Data
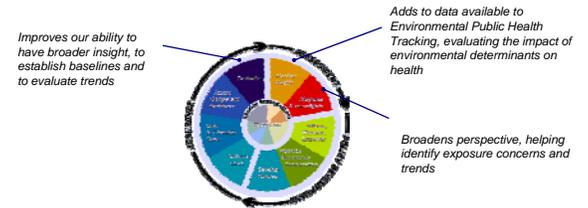

Meeting the Challenges- Proposed BEOH Metadata Table Fields and Definitions

One of the challenges of managing multiple data sets is determining how to enter data for samples that may have dozens of analytes. The intent behind the metadata table design is that CAS #s (Chemical Abstract Service numbers unique to each analyte) can be "cut and pasted" into the metadata entry form. Another challenge is managing multiple report formats. Some reports are received electronically. Other reports such as mold or VOC samples are received in paper form. Field-acquired data is also recorded on paper, but sometimes can be stored electronically. Where possible, links to the source electronic file will be provided. Although access to paper files can be provided in PDF format, there will be limitations as to how easily such data can be linked with other data sets. However, it will at least afford the opportunity to review and access available data and make a determination as to whether the effort to have it translated for use by AVR is worthwhile.

Field	Variable Description
Project ID	The name of the project. May include multiple samples.
Collector Sample ID	Unique identifier assigned by collector of sample
Lab ID	Unique identifier assigned by the laboratory (key linking lab and AVR data sets)
Date	Date in month/day/year format
Sample Type	Grab, time-integrated (e.g. pump and tube or filter), data logged
Begin Date	Date time-integrated sample began or grab sample collected.
Begin Time	Time time-integrated sample began or grab sample collected.
End Date	Date time-integrated sample ended.
End Time	Time time-integrated sample ended.
Sample Media	Air, soil, water
Media Sub-Type	Indoor air, private well, surface water, community well, outdoor air, soil, sediment, food, biota
Method	Method number, e.g. EPA TO15, OSHA, PV2120, EPA 353.2
Mailing Address	Postal address
Collection Point	Specific location at postal address, e.g. 3rd floor northwest corner office, behind garage, outdoor tap
GIS Coordinates	Longitude and Latitude
Depth or Height	Measured in feet
Analyte	CAS#, table developed such that multiple CAS numbers for a single sample can be entered simply and associated with that sample.
Purpose	Reason sample was collected, e.g. investigation of release, background/control, new well, etc.
Keywords	Identifiers relating to context of the sample, e.g. outdoor wood boiler, asthma, school, leukemia
Collected By	Sample collectors name
Agency	Sample collector's agency
Analyzed By	SLH or BEOH
Data Format	.xls, .pdf, doc, etc.
Data File	hyperlink to data file

How does this enhance essential services?

Improves our ability to have broader insight, to establish baselines and to evaluate trends

Adds to data available to Environmental Public Health Tracking, evaluating the impact of environmental determinants on health

Broadens perspective, helping identify exposure concerns and trends

Next Steps:
The following actions are planned to move this project forward.
• Pilot the metadata table with a large private well water sample data set.
• Once the metadata fields and variables are defined, a form to facilitate data entry will be developed.
• Data will be entered through that form using the well water data set.
• Meet with DPH AVR staff to develop a data module (cube) through which the well water sample data can be analyzed and reported based on time and location.
• Meet with the state lab and other agency partners to discuss the pilot. We expect these findings will lead to changes in the lab's data management systems, interagency data sharing and collaboration on metadata practices.
• Identify BEOH data sets that are large enough that data modules (cubes) can be developed to utilize AVR technology to analyze and report on characteristics over time and location. This will open new windows into Wisconsin environmental public health.

For more information, please contact:
Jim Morrison, MS, CIH, CHMM
Chemical Preparedness Coordinator
Wisconsin Division of Public Health
(608) 267-3227 m orrijm1@dhfs.state.wi.us