

SRRTTF Data Management Workgroup

Draft Notes for Meeting: 8 April 2015

Meeting Objectives: Identify key data management needs and options. Identify next steps.

Purpose of Workgroup:

Establish a data management work group to discuss larger data management issues.
Identify path forward for data management.

1. Review workgroup purpose statement
2. Data management needs: What do we want to do with the data we have?
 - Data warehouse storage
 - Identify key decisions to be made regarding data management
 - What questions do we need to answer?
 - Example – City of Seattle uses the EQUIS system to help with source tracing and manage the sampling workflow. EQUIS also has a module that can be used to prepare data for upload to EIM.
3. Types of data to be managed
 - Sample results, congener data, conventionals, sample characteristics and location, QA/QC
 - Identify how to handle:
 - Emerging/recent data
 - Data from external parties

Examples of data that is collected:

- River and synoptic studies (TF and EAP)
- Toxics Management Plan studies (dischargers)
- DMR reports (don't contain PCB data); PARIS is not linked to EIM.
- LimnoTech collected data from original data call and is stored in an Access database

Matrices

- Water
- Groundwater
- Stormwater
- Water column
- Sediment
- Biota
- Air

4. Review databases and capabilities, strengths and weaknesses.

General comment: all data systems will require personnel to review, qualify, and enter data. Also, all data systems require technical support. Availability and cost of support is a consideration.

The Ecology PARIS and DMR systems are separate from EIM but eventually will be similar to EIM, capable of managing receiving water and groundwater data and linking to PARIS.

Should also consider whether the system is web-based (data is stored on off-site server) or downloaded to users (like an Access database). There are some open source options, with costs associated with support. Web based interface has some advantages for keeping users up to date.

- EIM
 - EIM has required data. Ecology data coordinators work with uploading, data checks, etc.
 - It is a good repository for data but does not have some of the tools that other programs have. Data is of good quality and needs to be qualified before entry.
 - Data can be submitted from lab by EDD. There are templates already spelled out to accept data so that labs can deliver to EIM.
 - Meta data needs to be in a common format for uploads.
 - The My EIM is used by the Toxics Cleanup Program and has some limited tools
 - Ecology can't build custom tools (like a commercial program could)
 - If considering EIM as a database, make a list of what can be uploaded and uses for information
 - Discharge source tracing (Brandi Lubliner)
 - Toxics Management Plan data (possibly)
 - Products are a separate database (Josh Grice)
 - Pre 2013 PCB data in EIM is different because the PCB data template was changed then.
 - EIM has "end user" qualifiers but lab QA/QC data not stored in EIM in order to manage the size of the database. Supporting information and links with end user qualifiers are maintained.
- EQUIS
 - Used by City of Seattle.
 - Works well for their purposes.
 - Capable of uploading to EIM but takes time for data to be processed.
 - Has a good reputation, costs associated with purchase and ongoing licensing
- SFEI
- DRBC

5. Next Steps

- Conduct a requirements assessment (DDilks) and share at the May TTWG.
 - This would be used to evaluate existing systems and/or develop a purpose based system.
 - Prepare list of data management needs and decisions; send to DD – what tools are needed to manage data?
 - Mapping
 - Graphing
 - Etc.

- Evaluate existing database options and fill out table
 - EIM
 - SFEI (APB)
 - DRBC (RMCC – contractor will be looking at this in June or July).
 - EQUIS