EPA United States Environmental Protection Agency

EPA OSWER Data Management

Current Efforts and Issues

Jean M Balent
US EPA
Technology Innovation & Field Services Division
Balent.jean@epa.gov
703-603-9924

Overview

- Data in OSWER
- TIFSD's Experience
- Data Management in HQ and Regions
- Data Partners and their Efforts
- Issues
- Wrap Up

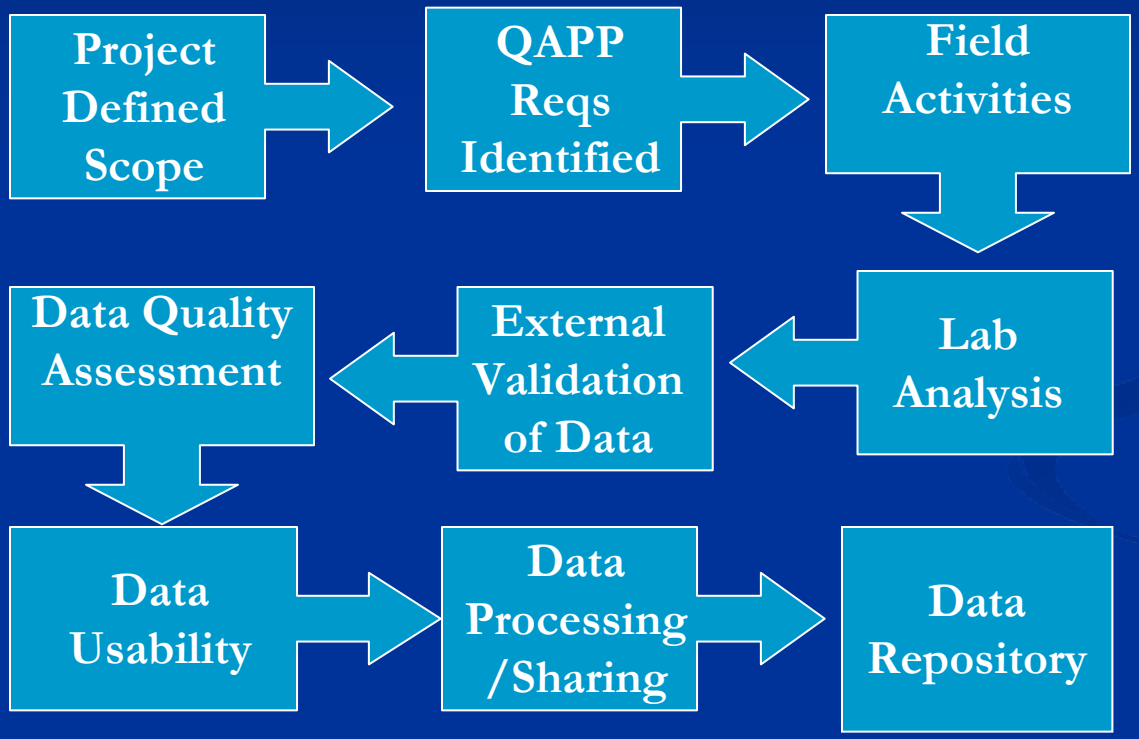
What data?

- Superfund program collects, reviews, and works with large volume of data for decisions at different scales
- Program Data
- Financial Data
- Congressional Metrics Data
- Site-specific data (EPA, RP, state, other data partners)
  - Geologic, Hydrogeologic data
  - GIS data
  - Contaminant data
  - Monitoring and remediation system data
  - Data used to construct a Conceptual Site Model (CSM)

OSWER Data Management

- Decentralized management approach
- Program management, financial, congressional targets collected and tracked nationally (HQ)
- Site-specific data management delegated to regions
  - Specific approach/tools unique to individual RPM
  - Many different approaches for managing data

Typical Data Flow for a Project



All data flow happens in different ways for each region and often each site

When the data flows, often times is not accessible for other purposes

Data flow isn't always linear: constantly collecting and refining CSM

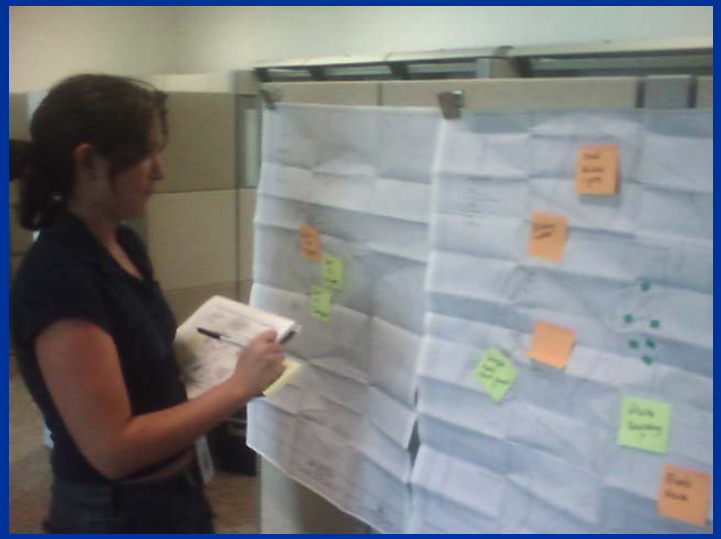
EPA United States Environmental Protection Agency

Technology Innovation & Field Services Division

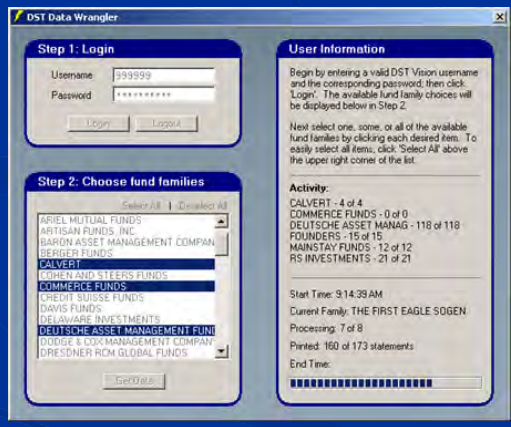
- Offer technical support to SF, BF, and other sites
- Assist with optimization, remedy design reviews
- Run the Contracts Lab Program
- ERT activities and tools (Scribe/Scribe.net)
- Collect and share information on hazardous waste clean up and site characterization

What we've seen…

Data visualization
Conceptual Site Model



Data storage techniques

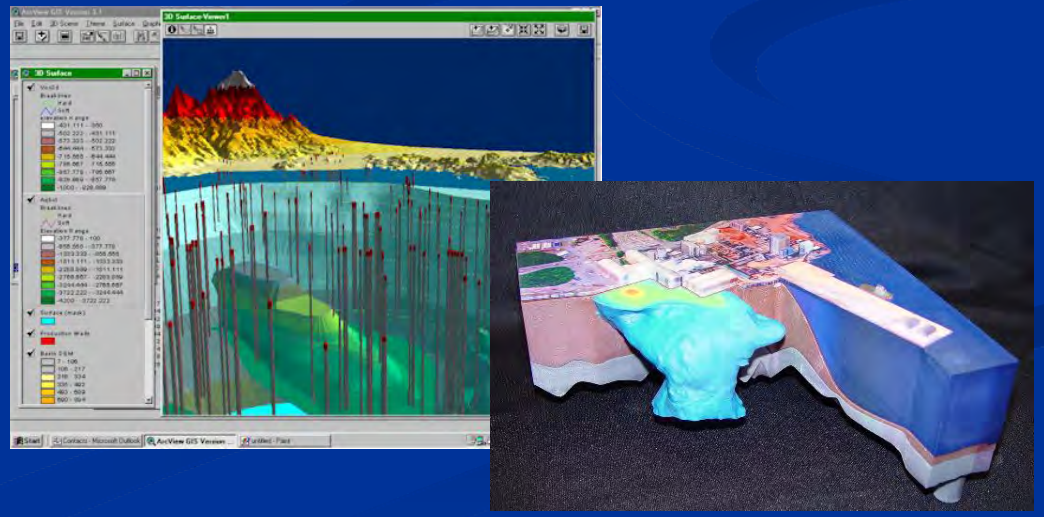


Data management

Often a contractor-owned or proprietary system is used to house and analyze data

What we've ALSO seen

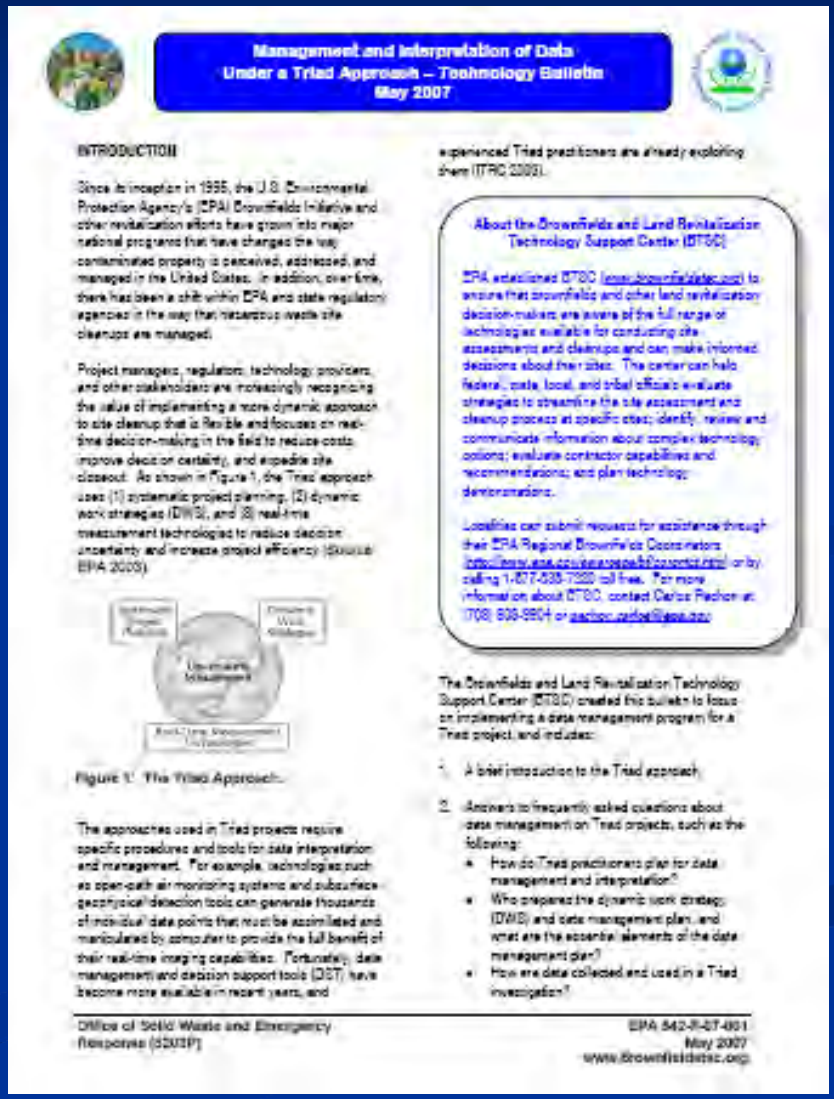
- Data management systems deployed in regions
- 3D models and realtime visualization tools
- Applications to ensure sampling design is sufficient
- Use of web-based project management system to provide realtime access to site data



TIFSD Publications & Resources

Several Case Studies and Bulletins available documenting smart data management approaches & tools

- www.cluin.org
- www.triadcentral.org



Management and Interpretation of Data Under a Triad Approach – Technology Bulletin May 2007

INTRODUCTION

Since its inception in 1995, the U.S. Environmental Protection Agency's (EPA) Brownfields Initiative and other revitalization efforts have grown into major national programs that have changed the way contaminated property is perceived, addressed, and managed in the United States. In addition, over time, there has been a shift within EPA and state regulatory agencies in the way that hazardous waste site cleanups are managed.

Project managers, regulators, technology providers, and other stakeholders are increasingly recognizing the value of implementing a more dynamic approach to site cleanup that is flexible and focuses on real-time decision-making in the field to reduce costs, improve decision certainty, and expedite site closeout. As shown in Figure 1, the Triad approach uses (1) systematic project planning, (2) dynamic work strategies (DWS), and (3) real-time measurement technologies to reduce decision uncertainty and increase project efficiency (Source: EPA 2003).

Figure 1. The Triad Approach.

The approaches used in Triad projects require specific procedures and tools for data interpretation and management. For example, technologies such as open-path air monitoring systems and subsurface geophysical detection tools can generate thousands of individual data points that must be assimilated and manipulated by computer to provide the full benefit of their real-time imaging capabilities. Fortunately, data management and decision support tools (DST) have become more available in recent years, and experienced Triad practitioners are already exploiting them (ITRC 2003).

About the Brownfields and Land Revitalization Technology Support Center (BTSC)

EPA established BTSC (www.brownfieldstsc.org) to ensure that brownfields and other land revitalization decision-makers are aware of the full range of technologies available for conducting site assessments and cleanups and can make informed decisions about their sites. The center can help federal, state, local, and tribal officials evaluate strategies to streamline the site assessment and cleanup process at specific sites; identify, review, and communicate information about complex technology options; evaluate contractor capabilities and recommendations; and plan technology demonstrations.

Localities can submit requests for assistance through their EPA Regional Brownfields Coordinators (http://www.epa.gov/brownfields/contacts.htm) or by calling 1-877-838-7350 (toll free). For more information about BTSC, contact Carlos Pachon at (708) 608-5804 or pachon.carlos@epa.gov.

The Brownfields and Land Revitalization Technology Support Center (BTSC) created this bulletin to focus on implementing a data management program for a Triad project, and includes:

1. A brief introduction to the Triad approach.
2. Answers to frequently asked questions about data management on Triad projects, such as the following:
   - How do Triad practitioners plan for data management and interpretation?
   - Who prepares the dynamic work strategy (DWS) and data management plan, and what are the essential elements of the data management plan?
   - How are data collected and used in a Triad investigation?

Office of Solid Waste and Emergency Response (5203P)

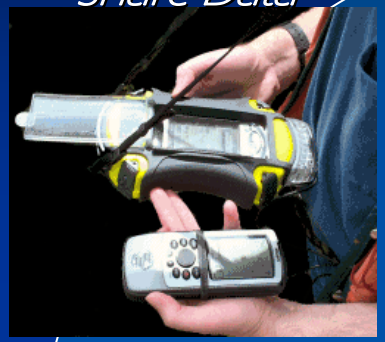
EPA 542-R-07-001
May 2007
www.brownfieldstsc.org

Data Flow & Tools

QA/QC

Collect Data

Share Data



Field Data

Lab Data

Scriblets

Scribe.net

Forms II Lite

SEDD

STORET/WQX

Scribe

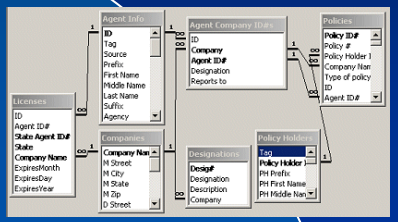
EQuIS®

CLPSS

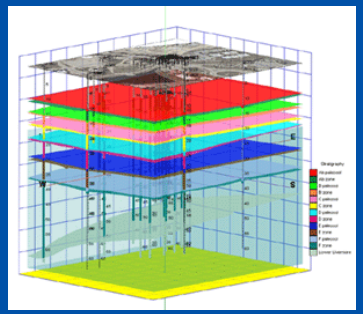
ADR

Store Data

Process Data

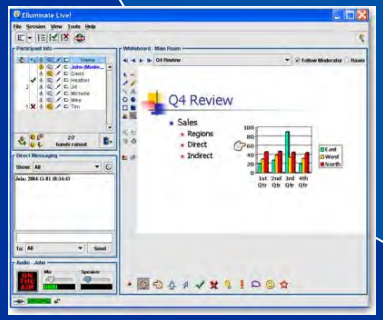


Database



Evolving Conceptual Site Model

Communicate



Quickplace

Collaboration Pages

WebOSC

Sametime

Web Conferencing

Documentation

Records

Distance Collaboration

VSP

SADA

DST Matrix

Make Decisions

Decision Support Tools

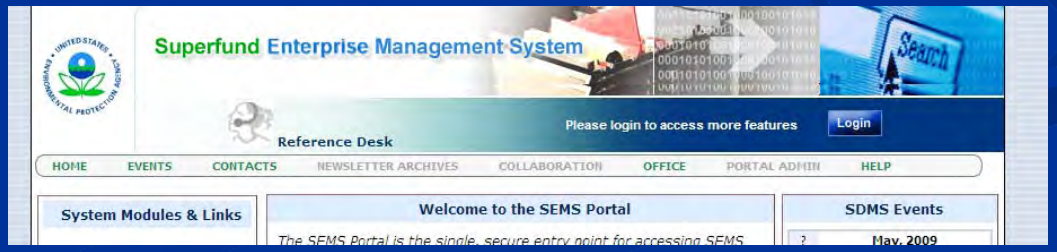
Data Visualization Tools

OSWER Data Systems

- **Several systems used to collect and manage information at national level:**
  - Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) → Program data
  - Institutional Controls Tracking System (ICTS)
  - Superfund Document Management System (SDMS)
  - Superfund E-Facts → Government Performance and Results Act Data
  - Contract Laboratory Program Support System (CLPSS) → analytical data from CLP
- **What about information for site-level decisions? Program wide reviews site conditions, cross program studies?**

Data Partners & Efforts

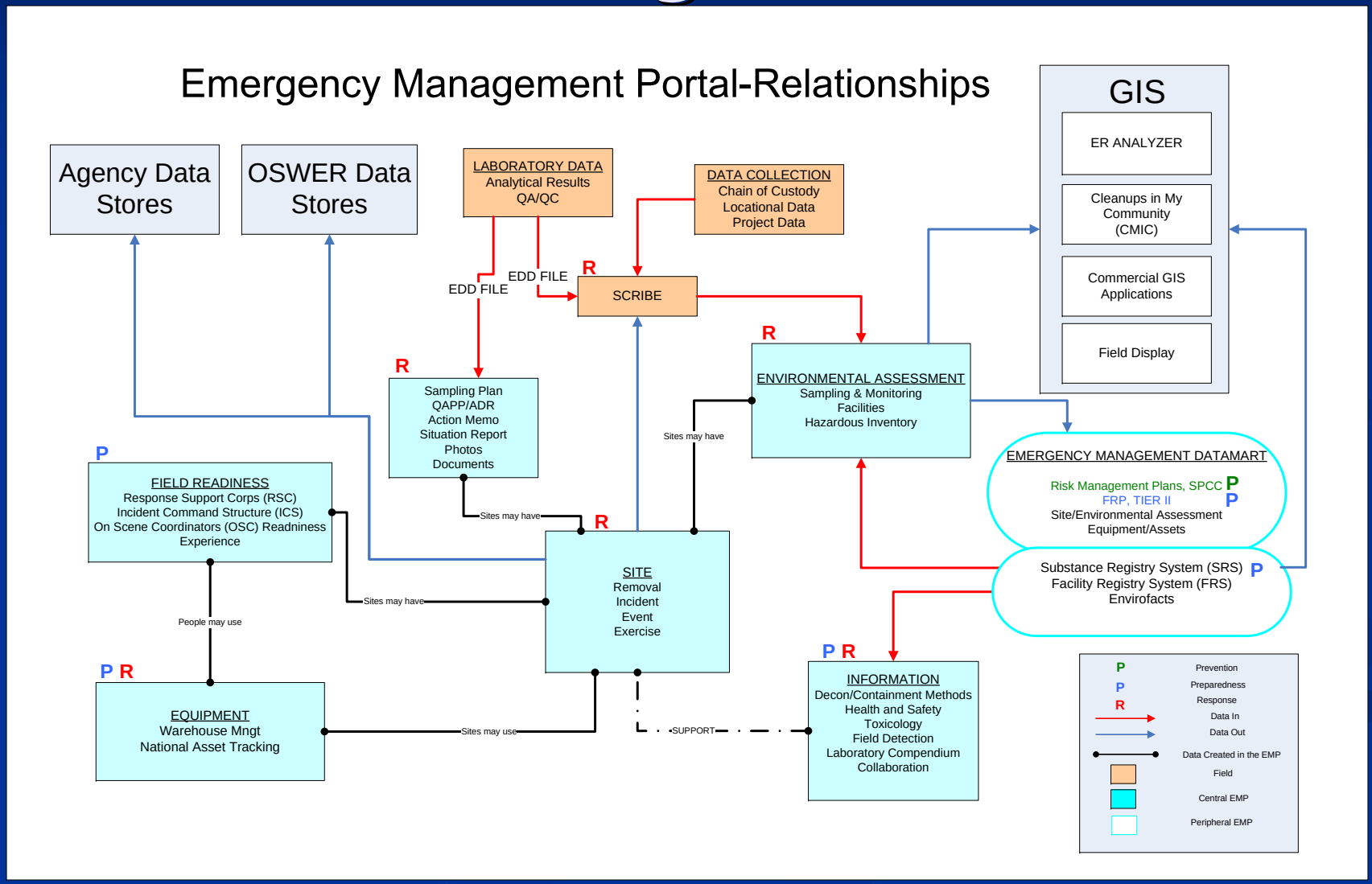
- OSWER is developing the **Superfund Enterprise Management System (SEMS)**
  - Integrate three primary Superfund data collection, reporting and tracking systems: CERCLIS, SDMS and ICTS
  - May include analytical data from CLP as well as Removal program
  - Issue: format of data, validation of data, method/format of storing data



Data Partners & Efforts

- OEM is building the **Emergency Management Community on the EPA Portal**
  - http://portal.epa.gov
  - Will encompass the WebEOC (emergency operations center) as well as include removal site data from/with Scribe/Scribe.net
  - Possible inclusion of other site data (SF remedial, BF, RCRA, UST)
  - Under construction, may have operational pieces in '09
  - Issues: coordinating efforts and data formats between ERT and OEM

Map Data Flow Regional Remedial Programs

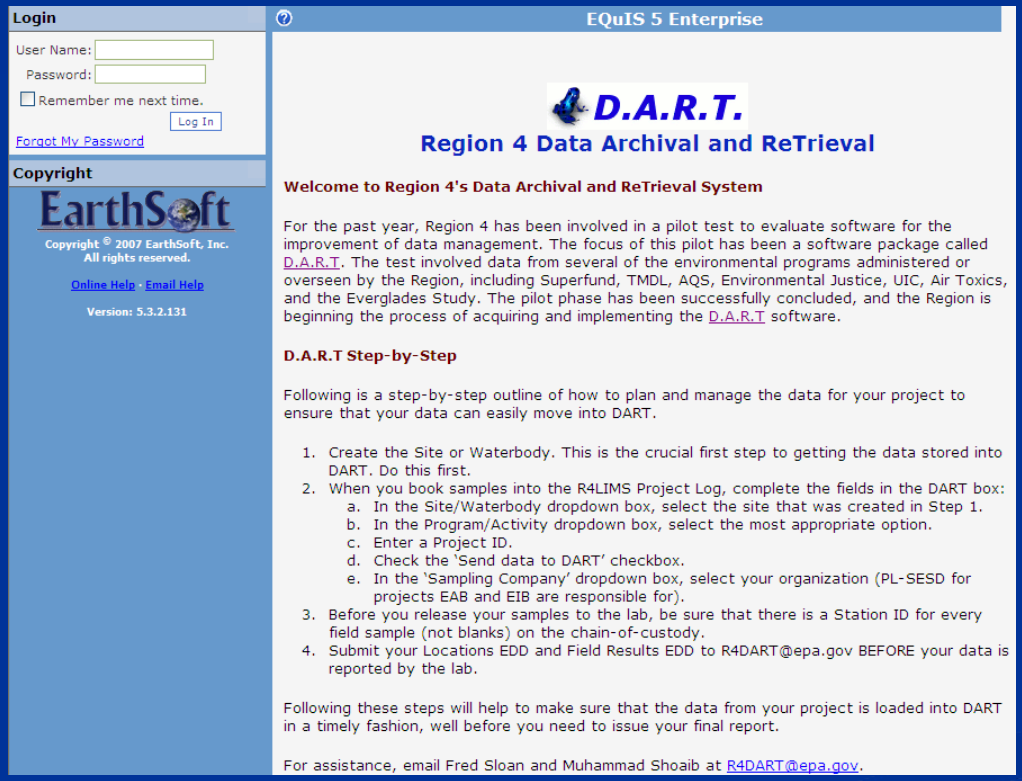


Data Partners & Efforts

- **Regions have adopted wide variety data management solutions**
  - R2, R4, R5 → EQuIS
    - 5 regions have EQuIS licenses
    - All at varying stages of deploying
    - Goal for some to use for managing ALL data from ALL programs
    - Plan/willing to export data to other system when required
  - Other regions using and exploring STORET/WQX and/or Scribe/Scribe.net for same purpose
    - R8, R9, R10
  - Contractor databases widely used in many regions
    - R1, R5, R3, R9, R10
  - Issues: variation in format of data and individual needs of each region, wide difference in effort to support data management in regions

Data Partners & Efforts

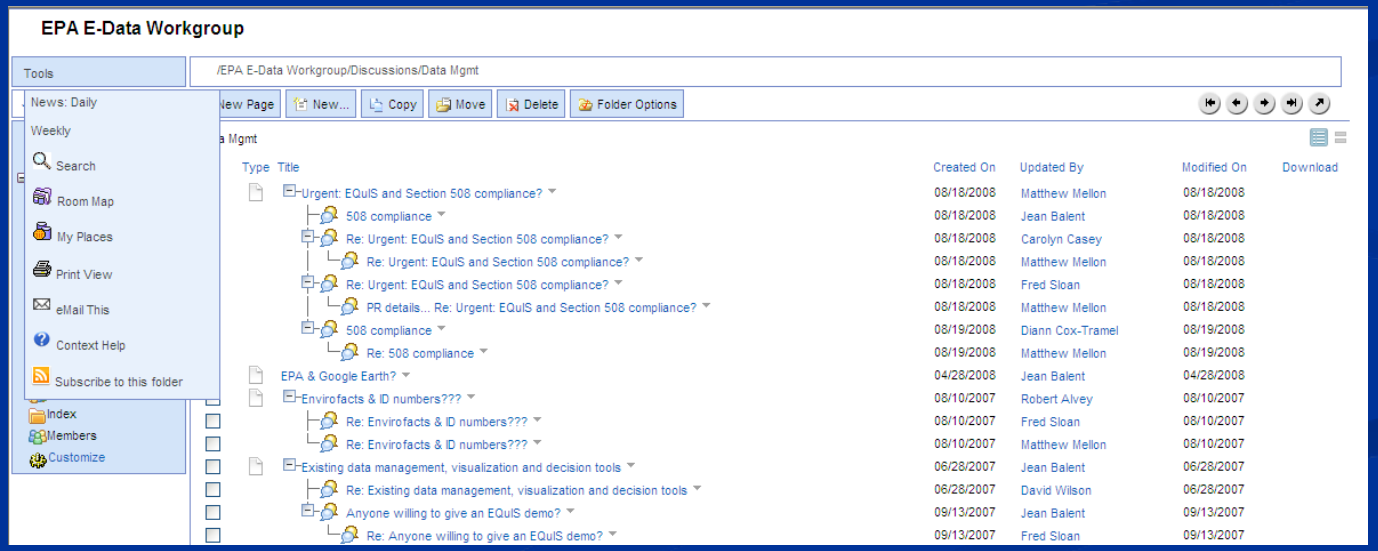
- **Region 4: DART**



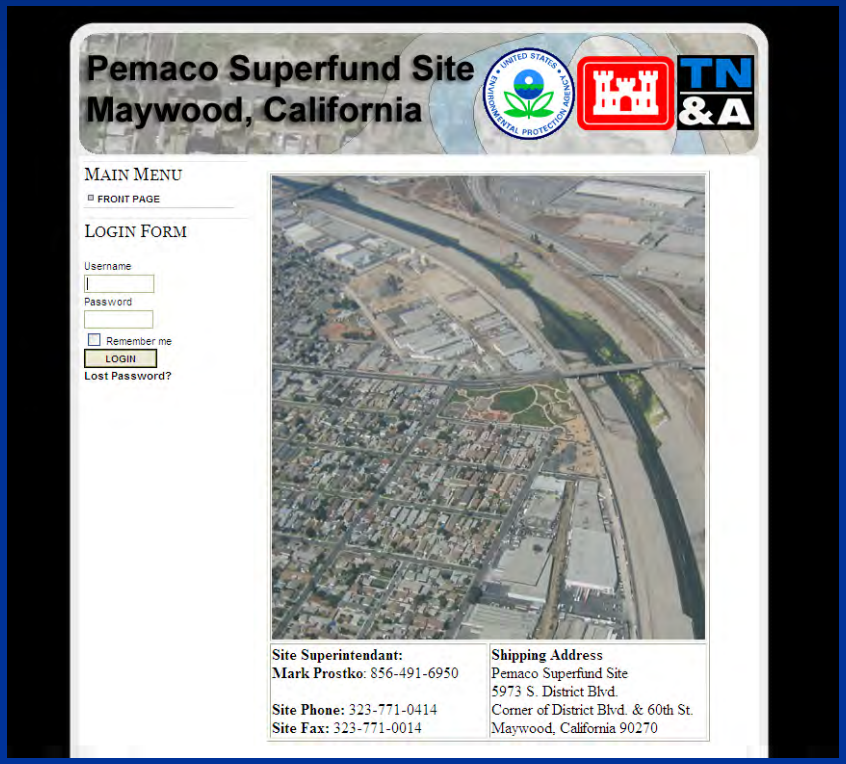
For the past year, Region 4 has been involved in a pilot test to evaluate software for the improvement of data management. The focus of this pilot has been a software package called D.A.R.T. The test involved data from several of the environmental programs administered or overseen by the Region, including Superfund, TMDL, AQS, Environmental Justice, UIC, Air Toxics, and the Everglades Study. The pilot phase has been successfully concluded, and the Region is beginning the process of acquiring and implementing the D.A.R.T software

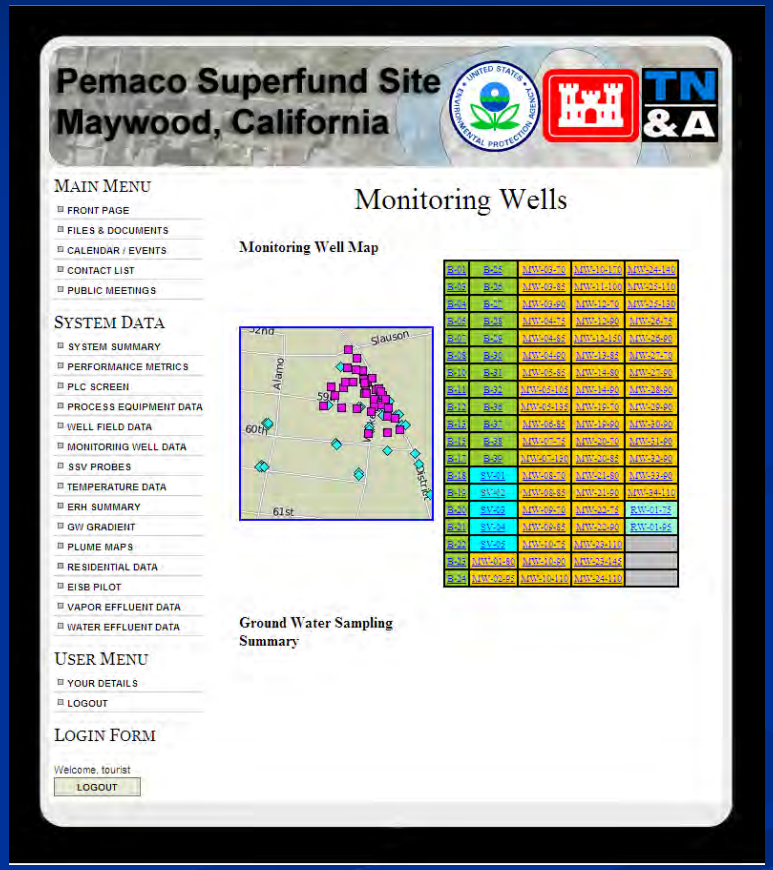
Data Partners & Efforts

- **E-datawork group**
  - Grass-roots collection of regional & HQ staff
  - collaborate and share experience with issues related to electronic environmental data



Data Partners & Efforts

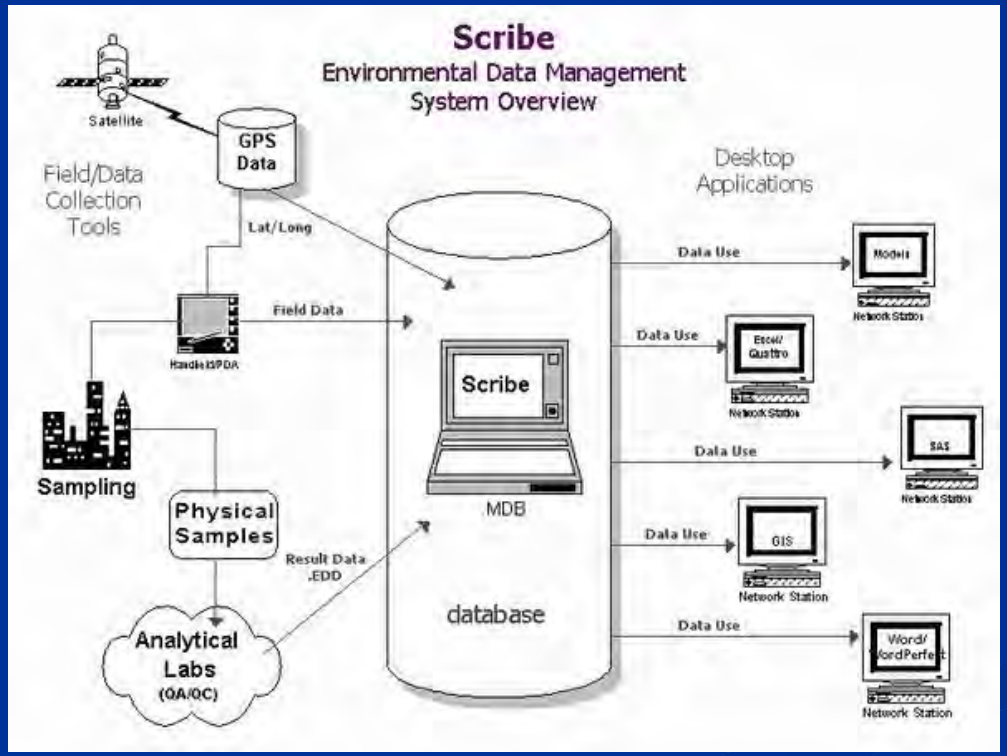


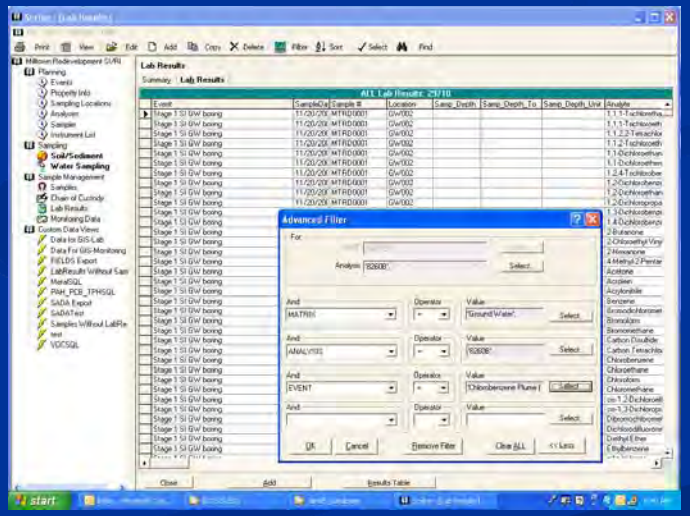


Custom, project-specific regional websites and databases for storing data, visualizing and exchanging information

Data Partners & Efforts

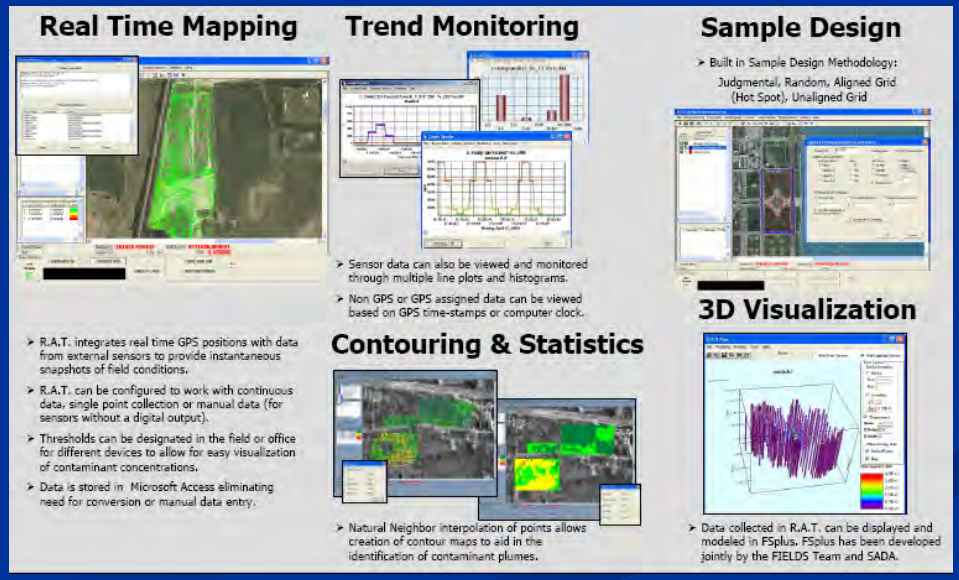
ERT's Scribe & Scribe.net used to provide site-specific, flexible data collection, management, and exchange solution used by OSCs and RPMs





Data Partners & Efforts

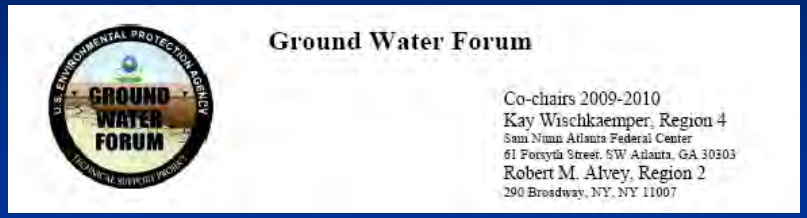
- **FIELDS (FIeld EnvironmentaL Decision Support) Team**
  - Collaboration of Region 5 employees and Research Associates
  - Assisted in the characterization of 50+ sites
  - Rapid Assessment
  - Tools (RAT) Software



Data Partners & Efforts

- **Other Data Sponsors/Partners have their own tools and formats for site data**
  - Multiple EDDs exist among the regions and SF Lab programs for various functions
    - **Lab EDDs to transmit data (200 EDDs)**
    - EDDs to transfer data to storage systems
    - SEDD for data review/validation (5 stages)
    - Contractor EDDs
  - Other agencies and states require or have their own EDDs which regions are working to comply with
  - PRP EDDs also exist
  - Field Data may also have their own EDD

Data Partners & Efforts



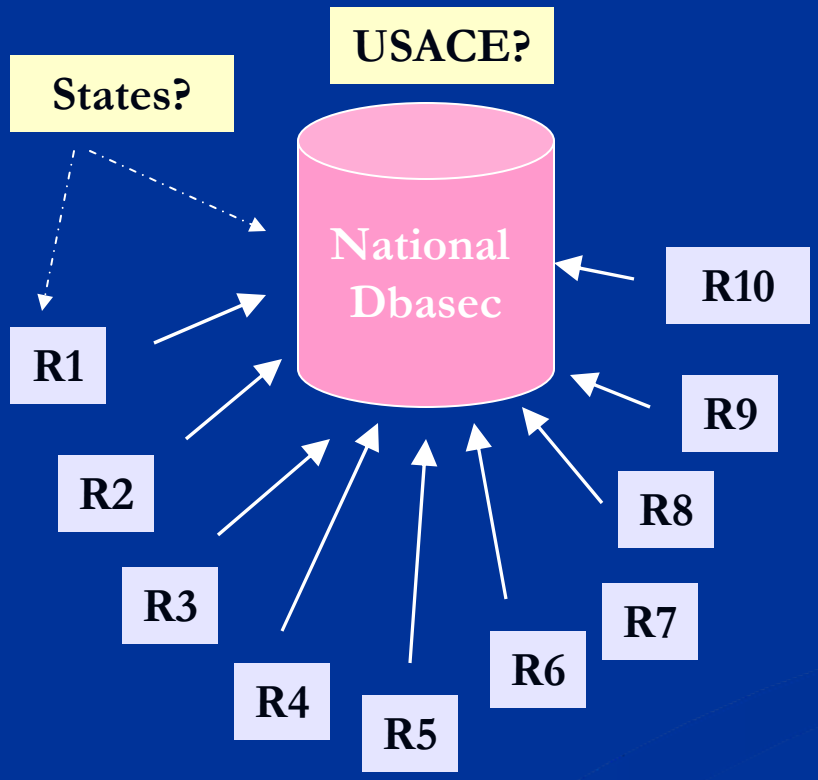
- EPA GW Forum believes this should be a priority for Agency
  - an Agency-wide initiative to expedite implementation of a **uniform data storage and management system for all chemical, biological and supporting hydrogeologic and locational information** collected at sites as part of routine EPA activities, starting with Superfund
- Other regions have expressed similar views to EPA management in the past

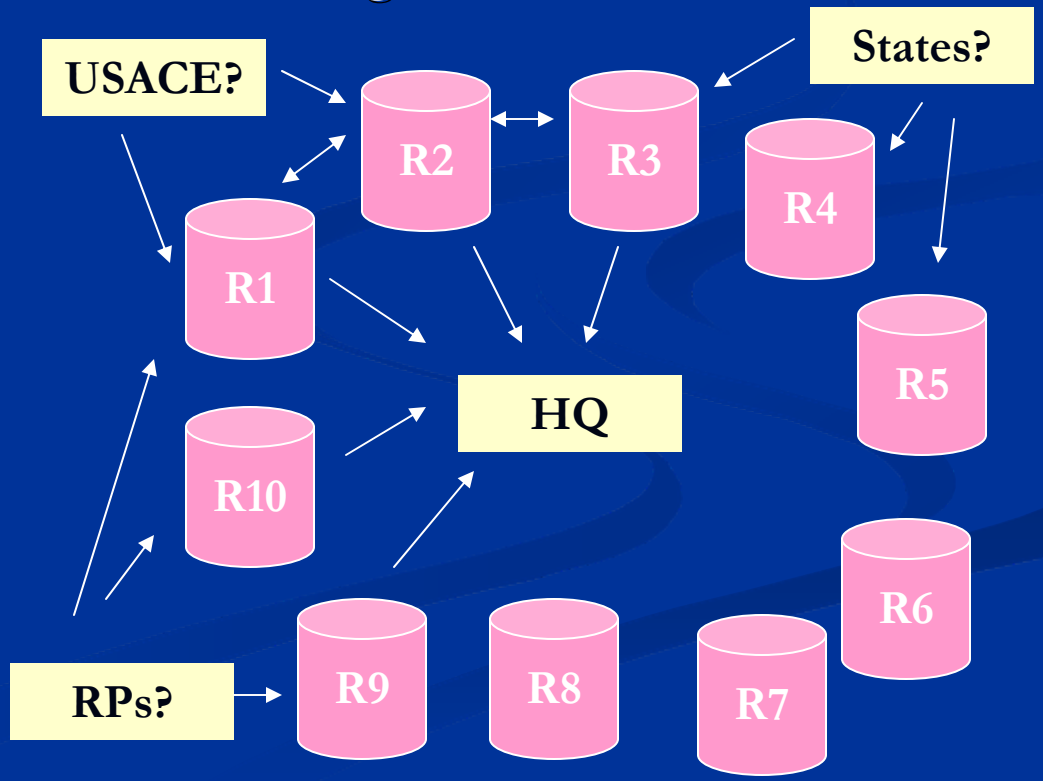
The Problem

- Increasing demand for readily accessible data
  - Recognized **growing volume** of data
  - Maturity of SF program (years of data, staff turn-over)
  - Accountability for decisions made with data
- EPA Superfund offices need access to better methods and resources to consistently store, access, interpret, analyze, visualize, and assure quality of environmental data

Is the Solution…

- **National Repository**
  - All report to one
  - HQ has ownership
- **Regional repositories**
  - Communicate between
  - Regions are data owners





Key Elements to the Solution

- **EDD (electronic data deliverables)**
  - Format in which data are stored and exchanged
- **Repository**
  - Place/object used to store the data
- **Tools**
  - Applications to review, analyze and understand data
- **Ownership**
  - Who controls the data, responsible for it

Issues surround each element!

Other issues

- Concerns to open access
  - SF Program does not have shape files for site boundaries
- Flexibility to regional needs
  - state well survey, new field in regional database to create unique well IDs
- How will data be used
  - If we build it, will they come?
  - Training/application needs for proper usage of information

Possible Next Steps for TIFSD

- Identify those who are "doing it better" and document their successes
- Work with these data "champions" to support an information exchange network
  - E-data WorkGroup
- Share information on these approaches, tools, and contacts with other regions and offices
- Look for new areas to improve
  - Push for "standard" approaches, consistent formats
  - Creation of a Data Support Center?
  - Deploy certain tools agency-wide?
  - Ongoing Needs-Analysis for Program?

Summary

- OSWER does not have a comprehensive solution for site-level data access and usability
- Many parties are involved and working on possible/partial solutions
- No one sole leader exists to coordinate all efforts and lead communication between players
- Is the answer:
  - Centralized Data Warehouse?
  - Regional Databases with access for other parties?
  - Standardizing a data format across programs?
- Work with OSWER management & Regions to start dialogs for a program-wide solution

Comments

Jean M Balent

US EPA

Technology Innovation & Field Services Division

Email: Balent.jean@epa.gov

Phone: 703-603-9924