



## Real-time Data Discovery and Notification of Restoration Progress through Automated Electronic Data Delivery

Michael F. Barinek, DePaul University, School of Computer  
Science, Telecommunications and Information Systems

Mitchell K. Beard, EarthSoft, Inc.

David E. Dougherty, Subterranean Research, Inc.

David A. Wilson, U.S. Environmental Protection Agency,  
Region 5

## Project Research Goal

- To answer the question:
  - *How to manage environmental electronic data deliverables to facilitate data feature discovery and notification?*

## Electronic Data Deliverables

- What types of electronic data?
  - Physical, chemical, geology, and biological data related to your site or project
  - Metadata
  - Location Data
  - *Geographic information*
    - *Topography, geological, geophysical data*
    - *Points of Interest*
    - *Aerial photographs, satellite imagery*

## Discovery and Notification

- What is data discovery and notification?
  - Data Discovery
    - The system discovers problems at your site
    - Examples may include: new analytes detected, results above action level, pumping wells under performing
  - Notification
    - The system immediately notifies you about these problems
    - Includes a variety of channels for notification delivery

Superfund, RCRA,  
Phase II, Phase III  
Investigations

Web-site,  
email,  
ftp,  
others

*Environmental Information Agents*  
*'Push' Reports*

## Benefits

- Benefits

- Know about problems faster
- Gain this knowledge routinely
- Find solutions faster

Reduces the risk from not knowing potential problems at a site

- Net results

- Shift from reactive to proactive
  - The system informs you of potential problems
- Better risk management and communication

Accelerates site closeout

*Individual sites or across multiple sites*

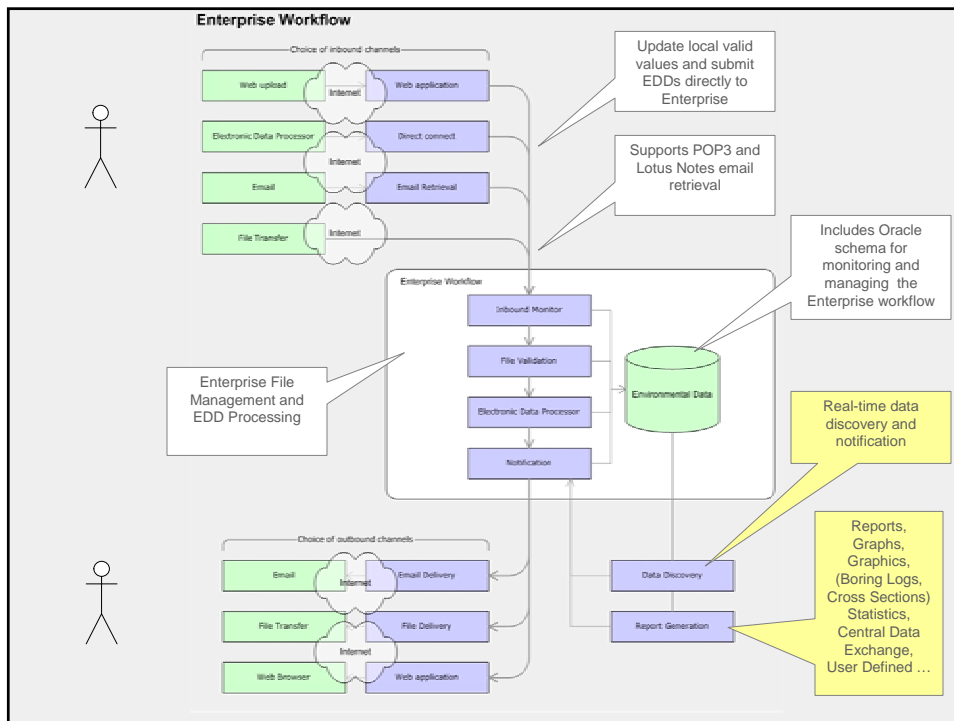
## Research Goal Revisited

- How do you achieve these benefits?

*Provide environmental enterprise software for managing environmental electronic data deliverables permitting, discovery, and notification*

# Enterprise System

- Include a formal, open interface specification describing electronic data exchange guidelines
- Push the responsibility of submitting data to the data provider
- Provide software to the Data Provider for preparing, authenticating, checking, and submitting EDD
- Provide an automated system for managing inbound and outbound files
- Provide automated Electronic Data Processor (EDP)
  - (Closed Loop: same code used by both the data provider and enterprise system)*
- Provide a mechanism for receiving notification of site restoration progress



## Interface Specifications

- Formal, open interface specifications describing electronic data exchange guidelines
  - Describes types of data that can be submitted in the EDD format
  - What information to be included with each EDD submittal
  - EDD file naming and table formatting conventions, data integrity rules, and valid data types
  - Checks for completeness and correctness

*EPA Region 5 Groundwater Evaluation and Optimization System (GEOS) EDD*

*Other EDD formats for individual sites, facilities, or agencies, ie ERPIMS, ERIS, SEDD, Multimedia EDD*

## Data Submissions

- Pushes the responsibility of submitting correct/complete data to the data provider
  - The data provider submits an EDD directly to the automated system
  - Forces compliance with the EDD format
  - Simple projects (USTs) can live with simple formats, but Superfund and Base Closure programs need more data
  - Common concern
    - “It’s too difficult to meet the specification”



*Provide software to the Data Provider for preparing and submitting EDDs*

## Data Provider Software

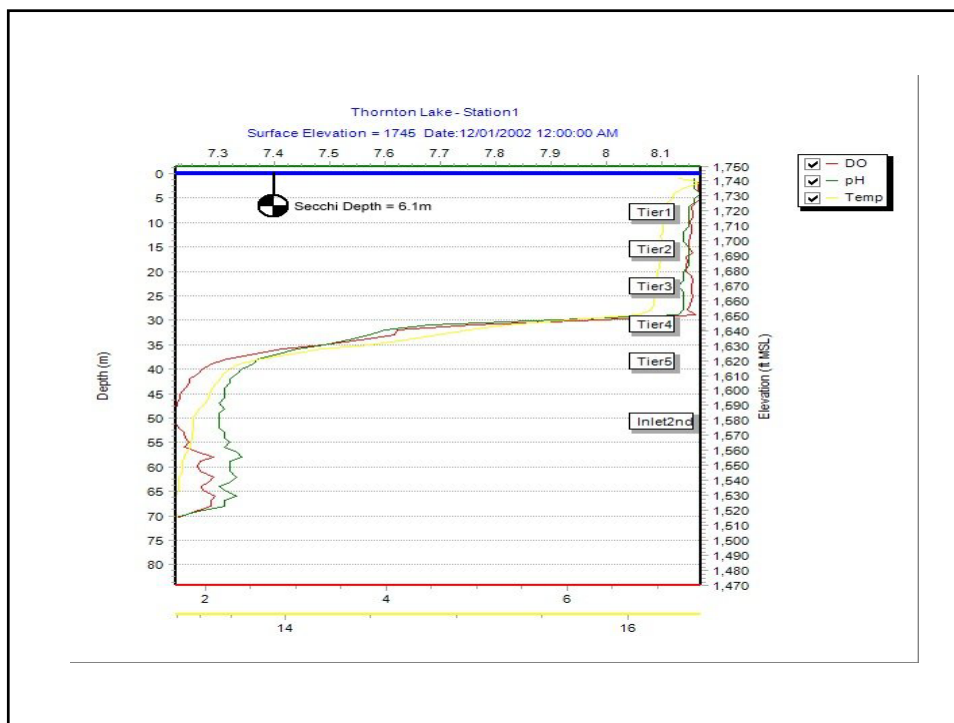
- Disconnected checking for the Data Providers for preparing and submitting EDDs
  - Desktop Electronic Data Processor (EDP)
  - Provide disconnected syntax and reference value checking
    - Local XML Schema for formats
    - Local XML Database for reference values
  - Includes a mechanism for receive local XML updates from the Enterprise system

## Closed Loop EDD Checking

- Desktop EDP is **same code** as Enterprise EDP
- Ensures same checks on both sides of data transaction:
  - Desktop EDP offers privacy, advanced checks, color coded error designations, can be called programatically from LIMS or other data collection software.
  - Enterprise EDP offers web or email interface, high throughput, automated responses with complete error log with EDD rejections, automated data loading into EQUIS
- Ensures complete, correct transaction

# Data Discovery

- Server-side 'Intelligent Agents'
  - Triggered
    - New data, new 'hits'
      - For example: If we have a new Arsenic 'hit', then generate:
        - » New Arsenic contours
        - » New Arsenic trend charts
  - Scheduled
    - Weekly, monthly regular events
  - On-demand
    - Ad Hoc
- Automated Generation of reports, graphs, graphics, models, statistics, visualizations, exports, etc.



## Summary

- Enterprise software for managing environmental electronic data deliverables permitting discovery and notification
  - Where we are in the development process
  - Enterprise and Desktop EDP available now, Intelligent Agents available in early Fall
- Our Partners
  - EPA, Region 5 GEOS
  - EarthSoft, Inc.
  - Subterranean Research, Inc.
- Demonstrations available.

