Using an Environmental Data Warehouse to Integrate Analytical Data, GIS, and the Web Presentation

Prepared by
Kentucky Research Consortium for Energy and Environment
233 Mining and Minerals Building
University of Kentucky, Lexington, KY 40506-0107

Prepared for
United States Department of Energy Portsmouth/Paducah Project Office
Acknowledgment: This material is based upon work supported by the Department of Energy under Award Number DE-FG05-03OR23032.

April 2005
Using an Environmental Data Warehouse to Integrate Analytical Data, GIS, and the Web Presentation

Prepared by
Kentucky Research Consortium for Energy and Environment
233 Mining and Minerals Building
University of Kentucky, Lexington, KY 40506-0107

Poster Presentation
Steve Cordiviola
Geological Society of America Meeting
Knoxville, Tennessee

April 2005
USING AN ENVIRONMENTAL DATA WAREHOUSE TO INTEGRATE ANALYTICAL DATA, GIS, AND THE WEB

David Korns, SAIC
Steve Hampson, UK-KRCEE
Steven Cordiviola, UK-KGS
John A. Volpe, UK-KRCEE
Bruce Phillips, Navarro Engineering
What Is The Problem?

• Large, complex sites/facilities have a multitude of reports, extensive data, technical information and drawings related to:
  – Engineering,
  – Environmental Sciences
  – Biological Sciences
  – Safety
Examples

• CERCLA/RCRA (superfund) Sites
• Defense Installations
• Industrial Sites
• Hazardous Waste Storage Facilities
• Disaster Areas such as New Orleans
What are the Issues?

- Sites have many years of activity and research.
- Multiple contractors, researchers, and regulatory agencies hold “pieces” of complete data sets & technical information.
- Accumulated knowledge not readily accessible

As knowledge is gained so do concerns
Engineering Challenges

- Site & Facility Plans are not created equal
  - Details vary from one technical drawing to another depending on
    » scopes of work for individual projects
    » or areas of interest and
    » contractor

- Variety of coordinate systems
  - Evolves over time

- Symbols and scales are not uniform
  - Different ways to represent same features
Scientific Challenges

– Data Generated from a Wide Range of Studies
  • Geotechnical and geophysical
  • Surface and Ground Water Studies & Modeling
  • Environmental Analyses
  • Emergency Response
  • Fauna and Flora Habitats
  • Land Use
  • Risk Assessment
Scientific Challenges

– Unique databases
  • From index cards to high-end relational dbs
  • Variety of field names representing same type of data or data collection locations
  • Unique naming of the same features

– Different levels of reporting standards
  • Always improving detection limits
  • Data Quality erratically reported
  • Data Validation erratically reported
  • Each researcher handles & reports data quality “exceptions” differently
Security, Safety, and other Regulatory Challenges

Federal, State, and Local Agencies

- Data Reporting Requirements differ between regulatory agencies
  - Multiple submissions of same data in different formats

- Security Issues
  - New levels of security bureaucracy in Post 9-11 era
  - New security rules in Post 9-11 era
  - Access to classified data?

- Paper, paper, paper
  - Submission, tracking & storage via traditional reporting mechanisms
SOLUTION?
The Concept of an Integrated Data Management/Retrieval System

• A systematic and consistent approach
  – Ability to retrieve and display data, maps, and models in a consistent and easy-to-use format
  – Automatic and customized reformatting of data from a variety of inputs
  – Appropriate Security Access to data depending on security clearance
Who Will Benefit?

- Site/Facility Owners
- Contractors
- Researchers
- Regulators
- Public
Components of an Integrated Data Management/Retrieval System

• Data Warehouse
• Geographic Information System
• Web access
  – Intranet
  – Internet
Data Warehouse

- Analytical data from all known sources of data integrated into a single database.
  - Spatially-enabled data tables (likened to a master sample location feature table in GIS)
  - Unified parameter names, units, dates, and location names.
  - Consistent, rule-based loading of data applied
    - Detection limits,
    - Missing data,
    - Non-detects
Geographic Information System

- Uses geodatabase concept
  - Centralized spatial and attribute data storage
  - Line and polygon topology models
  - Easy-to-use customizing and validation rules
  - Available Standards for a variety of features
Geographic Information System

• Integrate features from a variety of GIS and CAD datasets
  – Consistent coordinate system
  – Layers converted to feature classes
  – Versioning abilities
  – Include raster datasets
WEB Access

• Multiple query options
  – By form (drop-down lists)
  – By Map (point and click)
  – SQL queries (text-based)

• Multiple Views
  – Documents
  – Tabular data
  – maps
WEB Access

• INTRAnet OR INTERnet
• Variety of export features
  – To spreadsheets
  – Predefined models or applications
  – To reports
• Secure Access
  – Username and Password
  – “Public” access
Input & Conversion

Disparate Data Sets
- Environmental
- Geotechnical

Non-tabular Data
- GIS
- CAD
- Models

Paper Sources
- Reports
- Forms
- Photos

Rules-based conversion of data into a consistent format

GIS Analyst

Data Warehouse

Tabular data
X, Y Locations
Geodatabase format

Hyperlinked PDFs & TIFs

Extract Location Data
End-User Access

Data Warehouse

WEB server

Access Control

GIS Analysts, modelers, data entry, etc

Site Managers, Technical staff, contractors, regulatory agencies, contractors, public, etc.

Requests

Read/Write

Read Only

Firewall & Access Control
WEB Products
SEARCH TOOLS

By Form

By Map
WEB Products
Display Results

Tabular

Including Graphs

Graphical
WEB Products
Ancillary Output Possibilities

Spreadsheet

PDF Viewer

For Modeling

Graphing
Advantages/Disadvantages

• Minuses
  – DW requires routine updates
  – Conversion to geodatabases
  – Need to update security as users come and go
  – Very sophisticated system

• Pluses
  – 1 Stop Shopping for Site/Facility Data
  – Ends redundancy of Site/Facility data mining activities
  – Multiple Data format Output Capability
    • Each user does not have to convert data
    • Consistent formats
  – Uniform interface