Approaches to the Development of a Comprehensive Ecological Data Management System

Barry Baker[1], Vincent Pellerito[1] and Craig Bartlett[2]

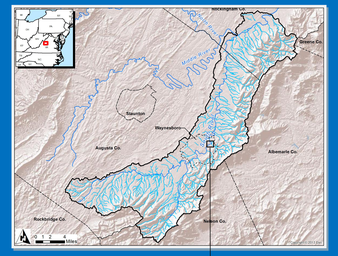
[1]URS Corporation, 335 Commerce Drive, Suite 300 Fort Washington, PA 19034; DuPont Corporate Remediation Group, [2]Chestnut Run Plaza, 715, 4417 Lancaster Ave, Wilmington, DE 19805

INTRODUCTION

The Program

The South River Program includes a comprehensive, multiyear, multidisciplinary data collection effort for a watershed in Virginia. Data collection has included a broad array of environmental data that must be managed in a way that permits complex statistical analyses, data integration and decision-making. Typically, standardized approaches to ecological data management have entailed custom developed systems that may meet the requirements of one ecological context, but may be inappropriate for another.

An ecological data management system should include certain features in order to effectively manage, analyze, and visualize data. To meet these requirements, these data management needs were described and appropriate software tools were developed and purchased that would allow for a seamless experience for the 'average' data user.



Project	Sample Name	Date Sampled	Sample Matrix	Fat Mass	Sample Mass	Fat Method	% Fat
LCLAC	LCS-25-21	10/17/2002 1:40:00 PM	Bulk Sediment	0.06	8.92	160.5	1UTOR
LCLAC	LCS-25-21	10/17/2002 1:40:00 PM	Bulk Sediment	0.66	8.82	UFGMBG	2296T-SO-0
LCLAC	LCS2-41	10/17/2002 1:40:00 PM	Bulk Sediment	0.07	8.13	UFGMBG	2296T-SO-0
LCLAC	LCS-25-21	10/17/2002 1:40:00 PM	Bulk Sediment	0.07	8.13	160.5	1UTOR
LCLAC	LS-04-01	10/17/2002 1:40:00 PM	Bulk Sediment	0.15	8.13	UFGMBG	2296T-SO-0
LCLAC	LS-04-01	10/17/2002 1:40:00 PM	Bulk Sediment	0.15	8.2	160.5	1UTOR
LCLAC	LS-50-53	10/17/2002 1:40:00 PM	Bulk Sediment	1.64	8.1	160.5	1UTOR
LCLAC	LS-50-53	10/17/2002 1:40:00 PM	Bulk Sediment	1.64	8.1	UFGMBG	2296T-SO-0
LCLAC	LS-50-55	10/17/2002 1:40:00 PM	Bulk Sediment	0.26	8.15	160.5	3UTOR
LCLAC	LS-50-55	10/17/2002 1:40:00 PM	Bulk Sediment	0.26	8.15	UFGMBG	2296T-SO-0

Sampling and Analytical Data

- Media chemical concentrations
- Quality control information
- Biological sample maturity and sex
- Biological sample growth metrics
- Species trophic levels
- Taxonomy
- Reproductive status
- Sub-sample blood and/or tissue information

Substrate Data

- Sediment Texture
- SAV Density

Spatial Data

- Studies of river geomorphic characteristics
- FEMA flood zones
- Hypsographic data products (such as elevation contours and DEMs)
- Potentiometric contours
- Land use/land cover
- Species diversity
- Cadastral boundaries
- Project boundaries

DEVELOPMENT NEEDS AND CHALLENGES

The Problem

The ecological data management system had to include certain features in order to effectively manage, analyze, and visualize data. It had consider not only whether the schema is sufficient to store and manage the data, but also whether the data served meets the expectations of data users., for example:

⇒ Does the data schema accurately describe the semantics of the data collection program?

⇒ Do query forms allow for extensive filtering and data aggregation options to precisely meet a data inquiry?

⇒ Are there multiple export formats available to users?

⇒ Is there sufficient support and training for the software?

⇒ Is there built-in user security that allows for database administrators to determine appropriate data access for users?

Management Needs

- An extensible relational data model
- Event oriented to associate spatial, temporal, and analytical data
- Ability to store metadata
- Standardized reference lists
- Updatable using various EDD formats
- Appropriate user security

Analysis Needs

- Ability for multiple levels of data aggregation
- Flexible querying for *ad hoc* ecological modeling needs
- Inherent statistical analysis capabilities

Visualization Needs

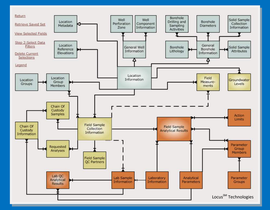
- Live data visualization
- Inherent mapping module for spatial data
- Appropriate visualization of time-series data
- Extensible reporting features

SOLUTIONS

An Evolving Solution

The data management solution for the South River Program has evolved from a custom developed, stand-alone solution using Microsoft® Access to a Cloud based 3rd Party based system using Locus℠ Environmental Information Management system (EIM)

The current solution represents a complete data management system including built-in mapping visualization, statistical analysis, and secure data access. With native data analysis and mapping capabilities, data users are more satisfied with data access for their scientific workload.

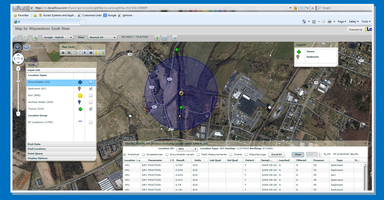
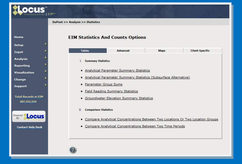


Management Solution

- Complete data management solution including data visualization (e.g., mapping)
- Event oriented to associate all pertinent spatial and temporal data
- Extensible biological data model
- Extensible reference lists including recursive classification lists
- Web-enabled allowing for universal software access
- Built-in validation and EDD error checking
- Customizable permission levels and menus, with secure access

Analysis Solution

- Flexible data aggregation and querying
- Inherent statistical analysis capabilities
- Accommodation for spatial and temporal analysis of the data

Visualization Solution

- Live data mapping and graphing
- Extensive visualization options including trend analysis
- Extensible reporting features