

# Proceedings of the 9th Annual Federal Depository Library Conference

October 22 - 25, 2000

---

## Gateway to the Earth: An Integrated Approach to USGS Information Delivery

David L. Govoni, U.S. Geological Survey  
Reston, VA

---

### Where We Were: Survey Science

- Discipline-specific scientific surveys:
  - Geology, Cartography, Geography, Hydrology, Biology, etc.

### Information Is Usually Grouped and Accessed by Organization

- Biologic Information
- Geographic Information
- Hydrologic Information
- Geologic Information

### Where We Are Going: Integrated Science

#### Gateway to the Earth Is Designed to Foster Appropriate Integration

#### Customers Require Information from Many Points of View

- *Organization* -- Science Center, District Office, Science Team, Mapping Center, Field Office
- *Discipline* -- Biologic, Geographic, Geologic, Hydrologic
- *Theme* -- Hazards, Resources, Environment
- *Geography* -- Place, Ecoregion, Physiography, Stream/Watershed, Latitude/Longitude, Congressional District
- *Time* -- Date, Epic, Time Series

#### Access Should Be Integrated and Interactive

#### Scope of Gateway to the Earth

- Gateway to the Earth encompasses all of the ways that users find, get, and use natural science information
- Initial focus on USGS Web sites and Web-based services



## Gateway to the Earth Project Elements

- Customer needs and expectations
- Standardized lexicon and metadata tagging
- Information architecture and graphical design
- Robust search capability
- Advanced geospatial search
- Information infrastructure support

## Customer Needs and Expectations

- Ensure utility for target audiences
- Identify primary and secondary customers and their information and support needs
- Analysis based on extensive customer interaction including focus groups

## Standardized Lexicon and Metadata Tagging

- Support meaningful, consistent content description (metadata) and labeling
- Permit more accurate search results
- Assemble a broad high-level USGS thesaurus with links to deeper thesauri as needed
- Tag all appropriate content with controlled metadata

## Information Architecture

- Provide a high-level framework for consolidating and presenting our Web-based knowledge assets in a way that is:
  - Understandable
  - Logically structured
  - Appropriately integrated
  - Consistently labeled
  - Easy to navigate and use
- More effective thematic browse interface

### **Information Architecture**

- Logical physical or virtual grouping of like information promotes less redundancy, more efficient navigation, and more meaningful cross-linkages

### **Information Architecture Design Strategy**

- Analyze current Web architecture
  - Identify principal knowledge assets and linkages - Identify potential improvements
- Define Gateway information architecture
  - High level organization of information assets
  - Principal topics and pathways for thematic browse
- Browse trees reflect rather than dictate information architecture

### **Web Browse Interface**

- Create visual metaphors to provide context and organization of content tailored for:
  - Audience (scientist, planner, student, etc.)
  - Learning styles (non-linear, visual, etc.)
  - Purpose (informational, visualization, decision support, e-commerce, etc.)
- Labeling system closely linked to thesaurus
- Tight integration of browse trees with automated, directed search tools

### **Robust Search Capability**

- One interface for querying all USGS information sources
- Improved information discovery
  - High probability of finding important, relevant resources
  - Aggregated, meaningful results
  - Ability to integrate related USGS information:

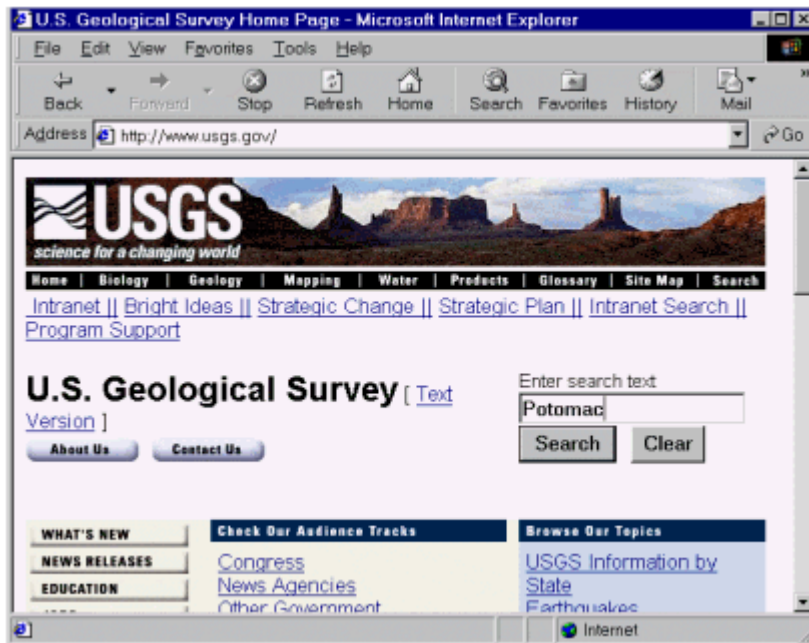
"Find other stuff like this!"

"Tell me more about this place!"

## Improved Search Engine Technology

- Ultraseek search engine
  - Adjusted rankings for major themes
  - Standard keywords from thesaurus
  - Geographic search capabilities
- Data Bases Search Engine (DBSE)
  - Focus initially on major USGS data bases
  - Search to granularity of each data base

## Standard USGS Search



## Standard USGS Search Result – Today

Found 1235 results – pages or documents

Pretty good for Web pages and on-line reports



## Separate Searches of Other Data Bases Yield More Information

- National Water Information System
  - 24 streamgaging stations along the Potomac River
- Map-List Data Base
  - <- 2 1:24,000 scale topographic maps named "Potomac"
- National Geologic Maps Database
  - 36 references to "Potomac"
- National Biological Information Infrastructure
  - 1361 references to "Potomac" from various sources

## Search Results We'd Like to See

- List of ALL databases that contain potentially relevant information
- Links that "drill down" into each database

## Advanced Geospatial Search Capability

- Permit fully integrated geographic and thematic searches of all USGS content
- Initial capability

- Search points and bounding rectangles of latitude/longitude
- Implement in Ultraseek
- Advanced Spatial Search Engine
  - Search polygons
  - Consider factors like footprint, shape, and data density to establish relevance ranking

### **Information Infrastructure Support**

- Extensive network upgrades
  - Failure-resistant Web services
- When data access is most critical it is also most vulnerable -Content mirroring  
 -Geographically dispersed servers at major Internet gateways

### **Integrated Information Is the Hub**

- Land and Water Sustainability
- Coastal Habitat
- Environmental Change
- Land and Water Restoration
- Hazards
- Species and Habitats
- Land Resources
- River Management
- Nonrenewable Resources

### **Gateway to the Earth Is the Key**

- Provides a common enabling structure so anyone can find, get, and use the data and information they need to:
  - Integrate relevant data and information to solve scientific problems - Make decisions - Educate and learn - Create new products and services - Use USGS scientific information in ways we have not envisioned