Big Earth Metadata Initiative: Catalog Services for the Web Analytics Scenario

Ted Habermann, John Kozimor, Sean Gordon
Earth Science Group
The HDF Group
thabermann@hdfgroup.org

John Farley
Raytheon
**Concept** – Abstract property or aspect of an object.

**Location** – World Wide Web Resources have unique locations that are expressed as Universal resource locators (URLs).

Documenting a World Wide Web Resource requires a location expressed as a URL:

**Spiral** – A collection of concepts required to support a particular documentation need or use case.
Recommendation – A set of concepts that a group/organization believes are required for achieving a documentation goal.
Attribute Convention for Data Discovery (ACDD)
Catalog Services for the Web (CSW)
Directory Interchange Format (DIF)

Recommendations
Subject: CSW Support
To: Data / Metadata System Manager
From: Above

I recently learned that it is critical that we support something called CSW for our metadata catalog. I think it is related to the OGC (not sure what that is).

We need to get this done ASAP! Presentation tomorrow!

Thanks,
The Management
Question #1: What is CSW?

Initially:
- Data Discovery – 8 recommendations
- Data Accessibility/Usability (Services) – 3 recommendations
- Data Understanding – 6 recommendations

Recommendation comparison tool helps identify concepts that are included in multiple recommendations: high priority concepts.
The impact of a recommendation on an organization depends on the metadata dialects that are used in the organization.
Recommendation and Dialect Maximums

- **Recommendation**: Max = 7
- **ISO**: CoreQueryables = Max = 7
- **ECHO**: CoreReturnables = Max = 7
- **FGDC**: Recommendation (Spiral)
Collection Analytics

Half of the DIF records are at dialect max.
These are good examples.
Two Improvement Strategies

Strategy #1
Achieve Dialect Max with existing systems and tools

Strategy #2
Achieve Recommendation Max with new dialects / systems / tools

This approach identifies
1. specific actions to improve metadata,
2. examples demonstrate benefits, and
3. quantitative metrics for measuring improvement.
Many USGS records are at dialect max for Core Queryables. These are good examples.

Dialect Maximum:
Core Queryables = 7
Core Returnables = 14
Future Directions

Wiki Updates
   Improved guidance (Elli)
   Updated technical content (John)
   Collection analytics results with connections to guidance

DOI – Two new recommendations
   ESDSWG
   NOAA

UMM – * Analytics

USGEO / Other Partners

ESIP Session