

The NOAA Data Center Metadata Enterprise: *IT Solutions*

NOAA National Coastal Data Development Center
NOAA National Oceanographic Data Center
NOAA National Geophysical Data Center
NOAA National Climatic Data Center

26-27 August, 2009
Asheville, NC, USA

Enterprise Functions

The Data Center Enterprise Metadata System

Metadata Manipulation Functions

Import

Convert

Export

Validate

Publish

Edit

Management/Admin Functions

Manage Controlled
Vocabularies

Manage
Components

Manage DM
Data

Manage
Workflow

Support Internal and
External Queries

Generate
Reports

Control Data and
User Access

Link to Archival
Storage

Handle
Versions

Minimize
Duplicates

System-Wide or Cross-Cutting Functions/Requirements

Support Human (GUI) and
Machine (API) Interfaces

Support
Standards

Support Collections and
Granules

Metadata Enterprise

Overall Assumptions

- All identified functions are important
- Categorization determined by Shareable vs User-Specific functions
- Enterprise functions can be categorized as both
- Some functions overlap

Metadata Enterprise Functions

Enterprise (Shareable)	Non-Enterprise (User Specific)
<ul style="list-style-type: none">• Validate (FGDC, ISO, customized)• Convert (FGDC, ISO, customized)• Control Data Access• Publish<ul style="list-style-type: none">• Manage Versions• External Queries<ul style="list-style-type: none">• Generate Reports• Versions• Manage Controlled Vocabularies• Minimize Duplicates• Support Standards• Support Machine (API) Interface• Manage DM Data (some aspects)• Manage Components	<ul style="list-style-type: none">• Edit• Import• Export• Control User Access• Internal Queries<ul style="list-style-type: none">• Generate Reports• Minimize Duplicates• Manage DM Data• Link to Archival Storage• Support Standards• Support GUI• Manage Workflow• Manage Versions• Manage Components

Metadata Enterprise

Questions/Clarification

- Refine/clarify requirements and priorities
- Clarification of “collections” and “granules” is needed
- Persistent storage format is still in question
- How do we search records?
- Unique identifier selection is required to reduce duplicates across the system

Metadata Enterprise

Design Concepts

- Communicate with REST services
- Flexible to accept other standards for validate & convert
- Transfer via XML within the Enterprise
- Mandatory publish of public records to Web Accessible Folder (WAF) for single location
- Require best practice tags in metadata record
 - Capture if and where data are archived within a metadata record, but not within the enterprise
 - Require data delivery mechanism within the record
- Unique identifier needed to reduce/eliminate duplicates, improved tracking

Metadata Enterprise

Assumptions

- FGDC to ISO crosswalk completed
- Component communication architecture selection defined
- Address security concerns for read/write access of catalog
- Select controlled vocabularies from existing
- Metadata catalog (TOC, more?) design defined
- Metadata catalog persistence mechanism determined
- Leverage WAF for metadata catalog prototype

Metadata Enterprise

Short-Term Efforts (1 of 2)

- Refine/clarify requirements and priorities
- Review MERMAid requirements
- Reinvigorate identification of vocabularies for cross-center adoption
- Pool resources to discuss approaches to Minimize Duplicates (e.g. unique identifiers)
- Collaborate on technical solutions for communication (e.g. service semantics)

Metadata Enterprise

Short-Term Efforts (2 of 2)

Proof-of-Concept:

- Leverage MERMAid pre-2.0 Release Validate and Convert capabilities
 - Flexible Validate and Convert services to allow user-defined schema & mapping
- Leverage GeoNetwork search capability
 - GeoNetwork harvests NOAA collections from MERMAid, NMMR, etc.
- Leverage WAF for metadata catalog prototype
- Versioning via subversion of WAF

Enterprise Functions

The Data Center Enterprise Metadata System

Metadata Manipulation Functions

Import

Convert

Export

Validate

Publish

Edit

Management/Admin Functions

Manage Controlled
Vocabularies

Generate
Reports

Manage
Components

Manage DM
Data

Support External
Queries

Control Data
and User Access

Manage
Workflow

Link to Archival
Storage via metadata

Handle Versions
(limited)

Minimize
Duplicates

System-Wide or Cross-Cutting Functions/Requirements

Support Human (GUI) and
Machine (API) Interfaces

Support Standards
(XML, FGDC)

Support Collections and
Granules

Metadata Enterprise

Mid- to Long-Term Effort

- Communication with system users, IT developers & upper management
- Develop formal requirements
- Component Management
- Control Data Access
- Publish
 - Manage Versions
- External Queries
 - Generate Reports
 - Versions
- Manage Controlled Vocabularies
- Minimize Duplicates
- Support Standards
- Support Machine (API) Interface
- Manage DM Data (some aspects)
- Manage Components
- FGDC to ISO crosswalk completed
- Component communication architecture selection defined
- Address security concerns for read/write access of catalog
- Select controlled vocabularies from existing
- Metadata catalog (TOC, more?) design defined
- Metadata catalog persistence mechanism determined