

To the Geoportal and Beyond!

Preparing the Earth Observing Laboratory's Datasets for Inter-Repository Discovery
Sean Gordon(scgordon) – University of Illinois (@illinois.edu), Eric Dattore (edattore), Steve Williams (sfw) – National Center for Atmospheric Research (@ucar.edu)

Problem

Each NCAR lab has a unique repository solution. Each has a different approach to atmospheric research. Currently it's not possible to search all of NCAR's repositories at once.

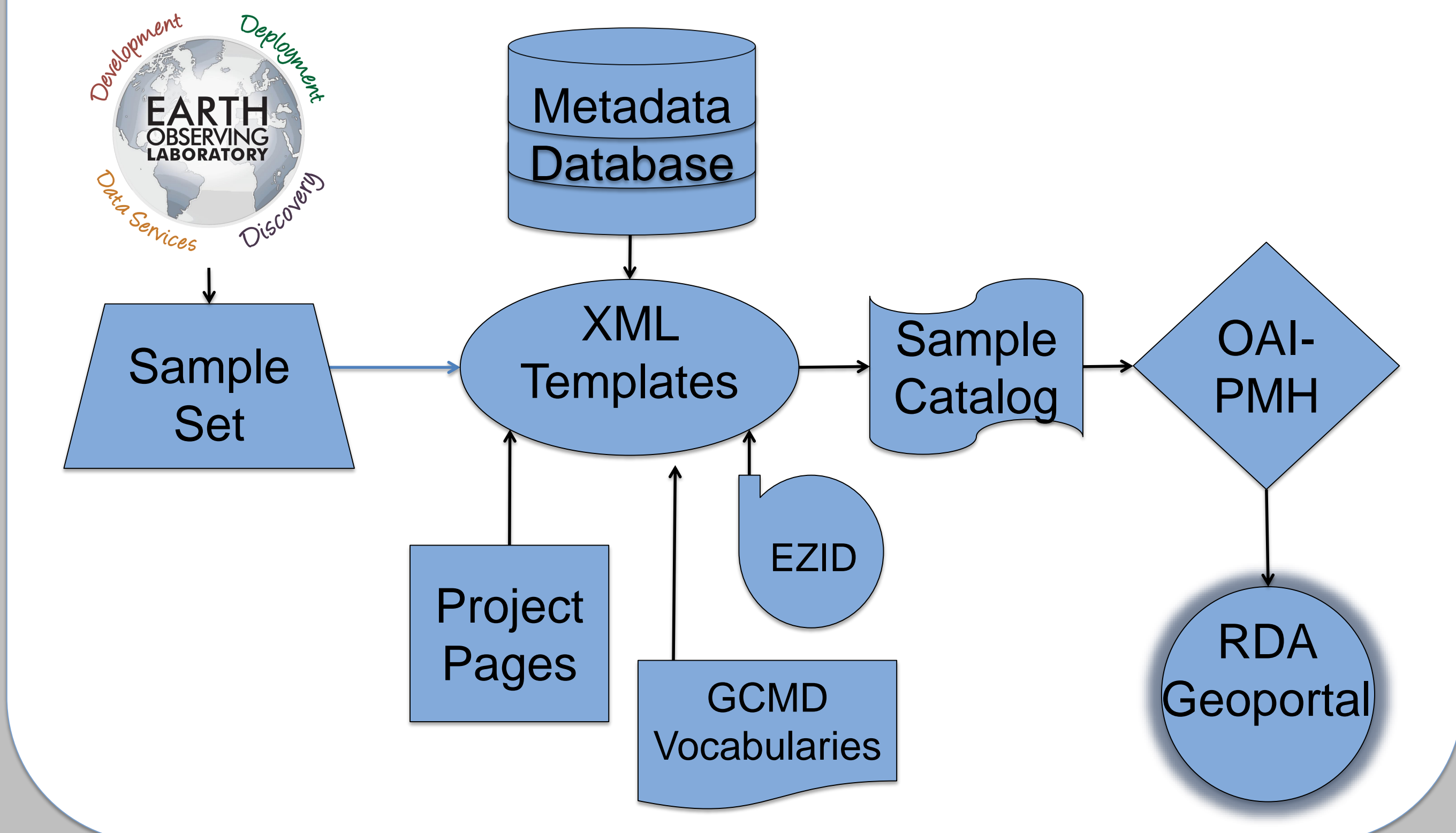
The Earth Observing Laboratory decided to investigate using ESRI's Geoportal as part of an infrastructure solution to facilitate discovery by connecting to the Geoportal that CISL's RDA runs to develop a proof-of-concept model for inter-repository discovery at NCAR.

EOL's metadata was not available in a format that Geoportal could ingest.

Research Questions

- What metadata standards are best to describe EOL's datasets?
- What metadata standards are compatible with Geoportal and the Global Change Master Directory?
- How can EOL's catalog be harvestable by a Geoportal and GCMD?
- What infrastructure is needed to connect repositories?
- How is it possible to automate the process?

Proof-of-Concept



Process

Ethnographic Inquiry

- Interview DMG about previous Geoportal testing
- Interview RDA about successes with Geoportal
- Interview EOL data managers about the ingest and storage of data and creation of metadata

Metadata Development

- Investigate and choose metadata standards
- Identify representative sample set of EOL datasets
- Crosswalk metadata standards to database
- Develop XML templates with a focus on discovery
- Draft a use policy that encourages citation
- Mint DOI's with EZID
- Extract metadata from databases and ingest applications
- Normalize EOL keywords to GCMD keywords, ISO topics
- Manually generate sample catalog of EOL's datasets

Infrastructure Development

- Develop OAI-PMH as a way to serve dataset documentation
- Coordinate with RDA to have their Geoportal harvest the catalog
- Lay groundwork for automating record creation
- Proof of concept model connecting EOL sample set to the RDA catalog at the RDA Geoportal

Key Findings

ISO 19115-2

- Explicit and highly modular standard
- FGDC current recommendation
- Easily extendable
- Commonly used by Earth Science repositories

OAI-PMH

- Enables Geoportal and GCMD to harvest catalogs
- Supports all chosen metadata standards
- Interoperability options for NCAR using repository wrappers

GCMD Science Keywords

- Shared vocabularies for explicit description and enhanced discoverability
- Widely used in Earth and Science repositories

Next Steps

Metadata

- Transform ISO 19115-2 into ISO 19115-1
- Consolidate database/ ingest layer
- Direct download links
- GCMD vocabularies
- Normalize EOL to GCMD
- Revise old project metadata

Infrastructure

- Develop Geoportal meaning file to display additional resources
- Connect with atmospheric repositories and domain intersections
- One NCAR Geoportal many OAI-PMH servers
- OAI-PMH repository wrappers

Automation

- XML record creation
- DOI assignment
- XSL completion rubrics
- Dynamic OAI-PMH

Conceptual Model

