

## ESDS-RFC-020 Technical Working Group Final Report

The ESDS-RFC-020 Technical Working Group (TWG) has conducted a review of the Earth Observing System (EOS) Clearinghouse (ECHO) Metadata Model Request for Comments (RFC) document with the following conclusion:

**That the Standards Process Group recommends ESDS-RFC-020 for endorsement as a NASA technical note**

### History

ESDS-RFC-020 was submitted by the ECHO project from NASA GSFC as a potential NASA Earth Science Data Systems Working Group (ESDSWG) Standard, in March, 2010.

Over a period of approximately five months, the ESDS-RFC-020 Technical Working Group (TWG) conducted a review designed to determine the implementation and operational suitability of the submitted RFC to NASA Earth Science systems.

A set of review questions was adapted from previous reviews conducted by the Standards Process Group (SPG). The review questions and community responses can be found on the SPG web site <http://www.esdswg.net/spg/rfc/esds-rfc-020>.

The review responses were positive with respect to both the value of the ECHO metadata model and the need for a metadata standard within NASA systems; however the overwhelming majority noted that it would be prudent to consider instead the emerging metadata standard for Earth science data from the International Standards Organization – ISO 19115. Being ISO compliant would allow for NASA metadata catalogs to be more visible to contexts outside their native domain. This would foster interoperability and encourage data sharing. Based on this consensus view, the TWG's recommendation is to publish the ECHO RFC as an SPG technical note<sup>1</sup> and revisit the RFC when the ECHO project completes their proposed mapping of core attributes to the ISO 19115 standard and/or to consider the ISO 19115/19115-2 standards for endorsement.

### Recommendation

The TWG bases its recommendation on an analysis of the following factors.

---

<sup>1</sup> A technical note, in SPG parlance is a document that contains useful information but is not a "standard". It must nonetheless, be relevant to the domain of NASA Earth Science data systems, serve a useful purpose, be technically of high quality, be well written and undergo a process of review from relevant stakeholder communities to prove its relevance.

**Strengths:** ECHO is a highly successful operational, enterprise-level metadata repository with over 3000 data collections and about 87 million data granules, and currently contains metadata holding exported by all NASA DAACs.

ECHO provides common mechanisms for provider communities to publish their data and service offerings and mechanisms for consumers to discover, understand, and access those resources. This enables ECHO to provide uniform search semantics across provider holdings for users and relieving some of the secondary responsibilities of providers, of supporting discovery and usage of that data.

The ECHO metadata model has its heritage from the well-established EOS Data and Information System (DIS) Core System (ECS) data model and is robust and able to represent most NASA's earth science data types.

**Weaknesses:** There was one perceived weakness in the ECHO Metadata Model cited.

While this concern had no direct relevance to the metadata model review, a reviewer felt that the use of the Alternative Query Language (AQL) as the primary query language was unwieldy to use. The generation of these XML (DTD) queries was not easy to create automatically and needed an element of manual 'hand-holding'.

**Applicability:** The ECHO metadata model has been optimized for remote sensing data, and therefore may be less suited to other NASA data types currently being cataloged, such as in situ data acquired in field experiments.

**Limitations:** From the review community's perspective, the single biggest hindrance to the adoption of ECHO as a proposed standard seems to have been that this would encourage the NASA earth science data systems community to "standardize" on a home-grown metadata model, rather than move toward an existing international standard for earth science metadata. As one reviewer stated, ECHO "is only used within NASA, and it is optimized for remote sensing data".

## **RFC Overview**

ECHO is a middleware solution which provides a Service Oriented Architecture (SOA) environment for the Earth Observing (EO) community. This environment serves both the providers and consumers of EO resources (data and services). ECHO provides common mechanisms for provider communities to publish their data and service offerings and other mechanisms for consumers to discover, understand, and access those resources.

By providing an SOA platform, ECHO allows the legacy providers of large amounts of data to focus on their contribution to the EO community to manage the availability of their information resources. These legacy providers can be freed from some of the secondary responsibilities of supporting discovery and usage of that data. ECHO works

along with its Earth Science Data Partners to gather metadata representing each partner's data holdings in a process called "metadata ingest." This metadata is then made available through a published Application Program Interface (API) which exposes the necessary functionality for data discovery. ECHO also acts as an order broker between its Data Partners and end users.

The submitted RFC defines metadata requirements and recommendations for ingesting earth science metadata into the ECHO system. ECHO's metadata model was derived from the ECS metadata model with some extensions. ECS data model in turn was developed in parallel to the FGDC model. Metadata conformant with ECS or generally with FGDC can be mapped to ECHO.

The following three metadata constructs are utilized by the ECHO system:

- Collection - A grouping of science data that all come from the same source, such as a modeling group or institution. Collections have information that is common across all the granules they contain and a template for describing additional attributes not already part of the metadata model.
- Granule - The smallest aggregation of data that can be independently managed (described, inventoried, and retrieved). Granules have their own metadata model and support values associated with the additional attributes defined by the owning collection.
- Browse - An image which provides a high-level view of the associated granule or collection metadata item. Browse images are not spatially enabled, but are very useful during data discovery and cross-referencing to other granules or collections.

For each metadata type, the minimum metadata fields required to validate against the ECHO Ingest schema are outlined. In addition, a list of recommended metadata fields that should be included in data ingested into ECHO is provided. These additional fields will facilitate the searching and ordering of data by the Earth Science community.

The RFC described above, as well as supporting documentation and community reviews, can be downloaded from:

<http://www.esdswg.net/spg/rfc/esds-rfc-020>