

ECHO NRT Metadata Field

Prepared by: Kathleen Baynes

1 Overview

This paper proposes the addition of a new, non-required boolean metadata element in the ECHO 10 Collection schema (e.g. `<NearRealTime >true</NearRealTime>`) to identify NRT products.

2 Background

ECHO has started incorporating Near Real Time (NRT) metadata into its catalog. In order to generate data products within 3 hours of observation time, a number of changes have been made to standard processing approach to expedite the availability of input data sets. NRT data is replaced in the ECHO catalog by Science Quality products and is aged out of the system after a period time.

Because of the transient and expedited nature of these NRT products, there needs to be an easy way to distinguish these holdings from archive data within ECHO. This should be easily determined both via the ECHO API and via the Reverb web interface.

3 Proposed Strategy

ECHO should introduce a new element to its metadata at the collection level. This element should be of type `xs:boolean`. A granule within this collection should inherit this property from its parent, so is not needed in the granule level schema.

With the introduction of this new field, data providers can specify NRT products and those products can easily be made available for searching via ECHO. In addition, these NRT products can be easily visually distinguished via end-user interfaces such as Reverb.

The element to be added is described as follows and should precede the Orderable and Visible elements in the Collection.xsd. This is not a required field.

```
<xs:element minOccurs="0" name="NearRealTime" type="xs:boolean">
  <xs:annotation>
    <xs:documentation>
      Indicates whether this collection contains Near Real Time Data
    </xs:documentation>
  </xs:annotation>
</xs:element>
```

An example instance of this element as it would appear in the Collection xml document is as follows. Again, this is not a required field in the ECHO 10 Collection metadata.

```
<NearRealTime>true</NearRealTime>
```

In addition, ECHO should expose this field for discovery via the catalog-rest API by adding a **near-real-time** parameter for searching collection metadata. An example of a catalog-rest API route using this parameter is as follows:

```
http://{server}:{port}/catalog-rest/echo_catalog/datasets?near_real_time=true
```

The above query would return all ECHO collections marked as Near Real Time. This could be used by Reverb to expose these collections via the web.

4 Data Partner Impacts

This should only impact data providers actively supplying NRT data to ECHO. These providers will need to re-ingest their collections to include the NearRealTime element set to true. Because the new field is not required, no other data providers should be impacted.

5 End-User and Client Impacts

Reverb users should experience no impacts because of this change. Developers who are working with the ECHO API to perform searches should be made aware of the new search parameter and should also be made aware that any data they are returned could have this new field included. If they are doing Collection schema validations of the returned results, they will need to ensure they are pointing at the latest version of the Collection.xsd file. The schema file at the following location will be updated to reflect this update:

```
http://www.echo.nasa.gov/ingest/schemas/operations/Collection.xsd
```

6 Future Work Impacts

The Metadata Architecture Study part 2, currently underway should consider adding a Near Real Time field to its Unified Metadata Model (UMM). This would ensure that future metadata repositories expose this field as searchable from the outset.

Date	Version	Brief Description
May 2013	1	Initial Draft

Table 1 Document Revision History