EOSDIS

NASA'S EARTH OBSERVING SYSTEM
DATA AND INFORMATION SYSTEM

Making Metadata Better with the CMR and MMT

Summer ESIP 2016

Jason Gilman

Dana Shum

The material is based upon work supported by the National Aeronautics and Space Administration under Contract Number **NNG15HZ39C**

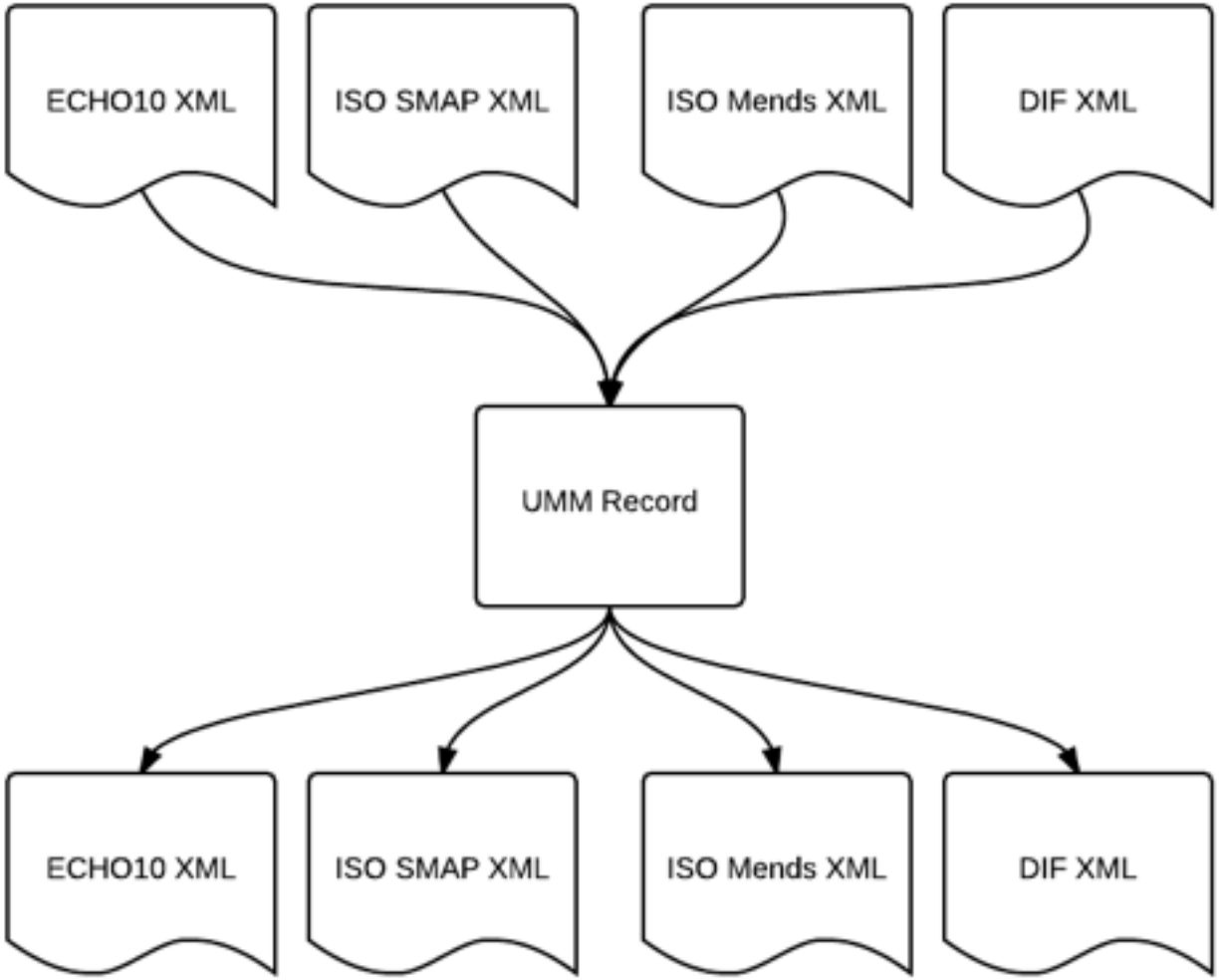
METADATA PROBLEMS

Completeness and Consistency

- Misspellings: “Bioosphere”
- Legacy Terms: “AM-1” instead of Terra
- Inconsistent Names: Processing levels “Level 1”, “1”
- Whitespace around element values
- Missing elements

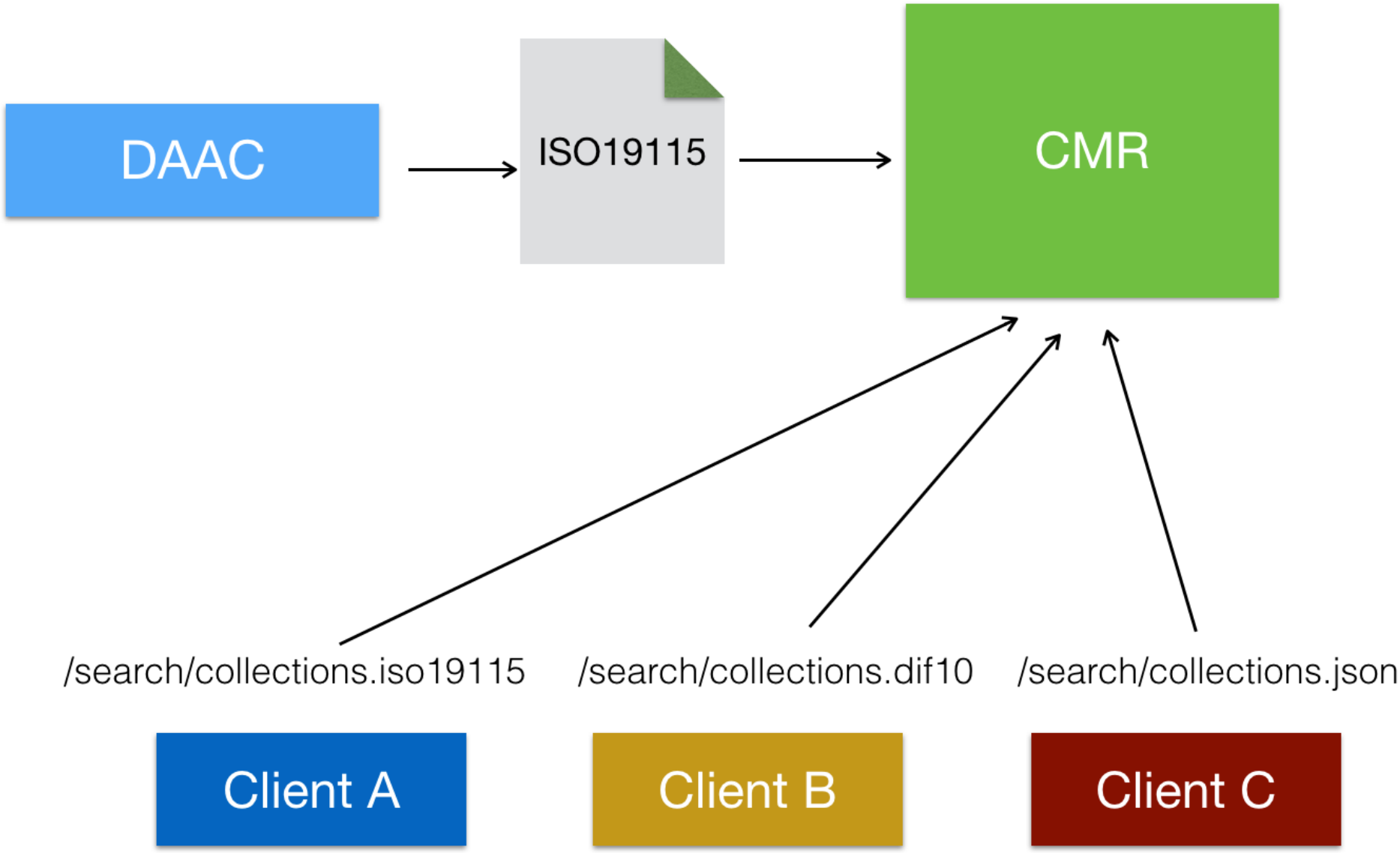
Overview and Benefits

UNIFIED METADATA MODEL

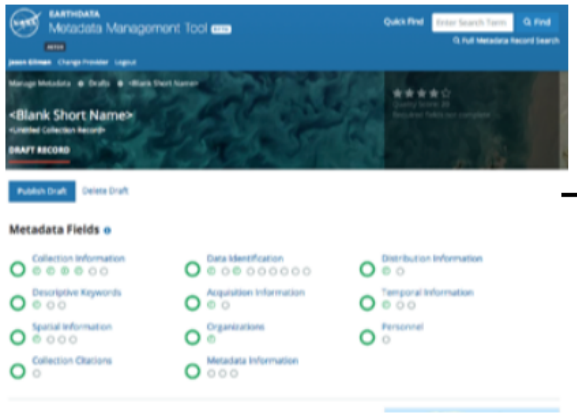

ECHO10 XML
ISO SMAP XML
ISO Mends XML
DIF XML
UMM Record
ECHO10 XML
ISO SMAP XML
ISO Mends XML
DIF XML

UMM Benefits

- Common validations across all metadata dialects
- Consistent features across all dialects.
- Allows conversion from any dialect to any other dialect.


DAAC
ISO19115
CMR
/search/collections.iso19115
/search/collections.dif10
/search/collections.json
Client A
Client B
Client C

MMT



{UMM JSON}

CMR

/search/collections.iso19115

/search/collections.dif10

/search/collections.json

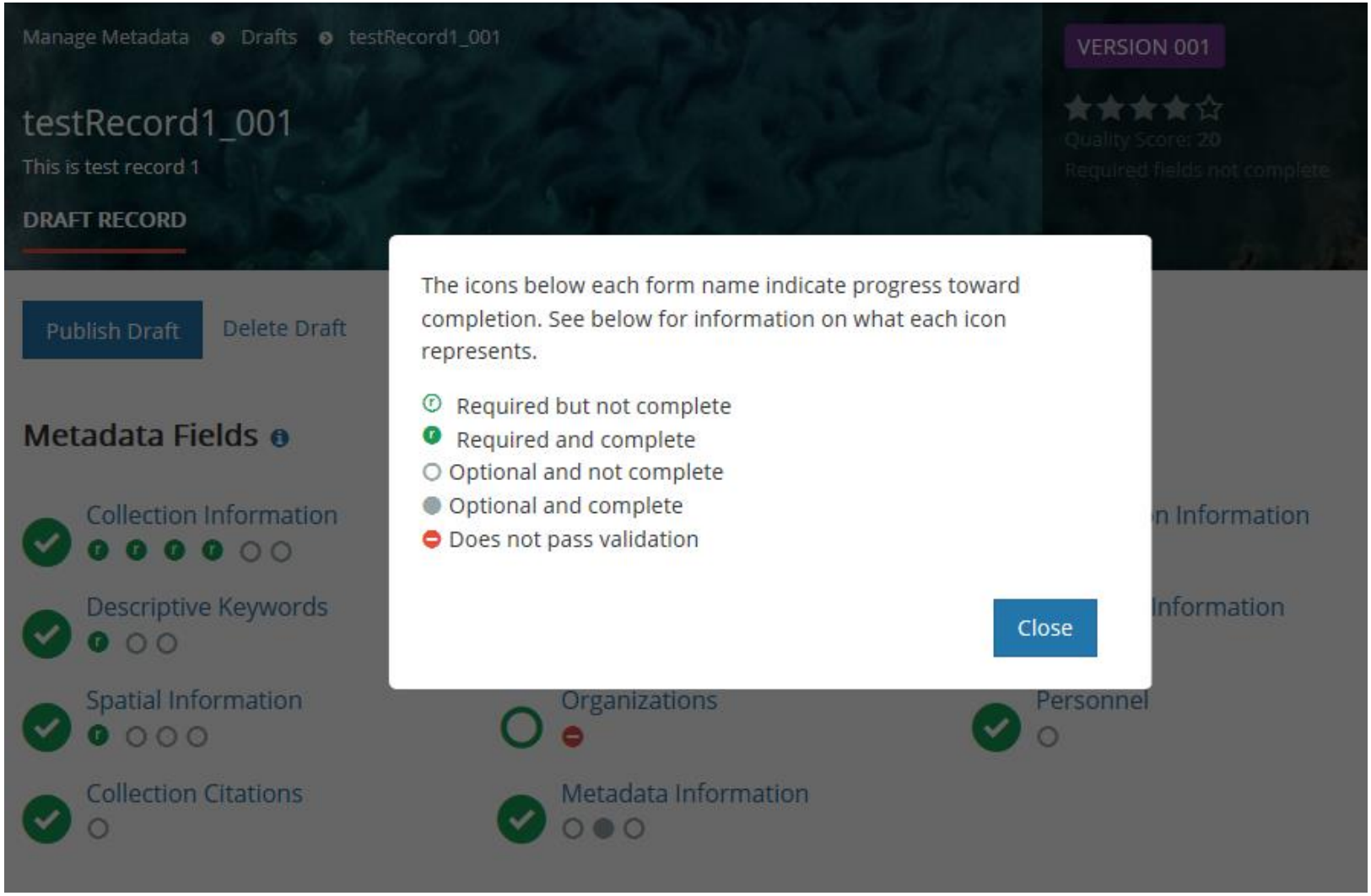
Client A

Client B

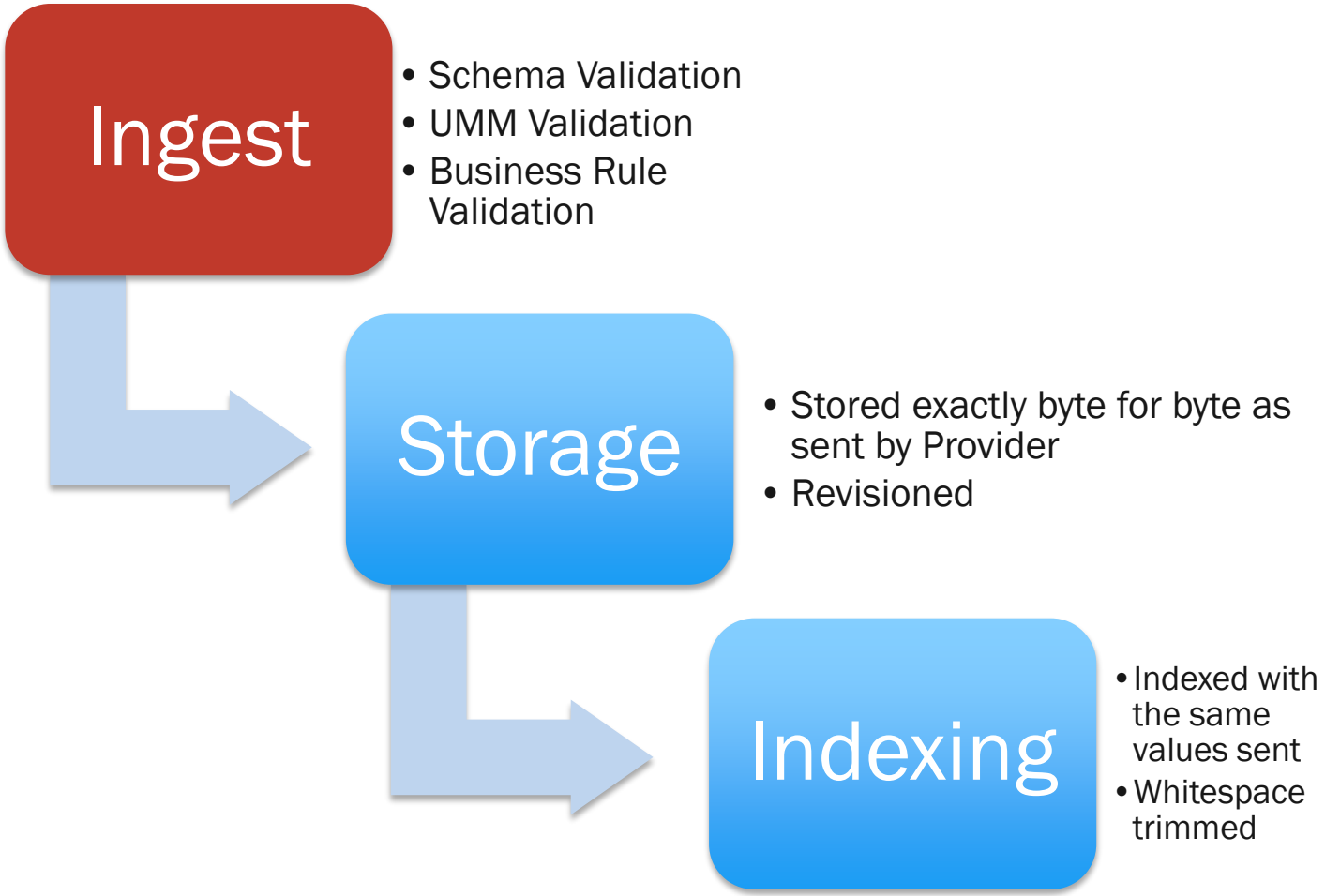
Client C

COMPLETENESS AND CONSISTENCY ENABLERS

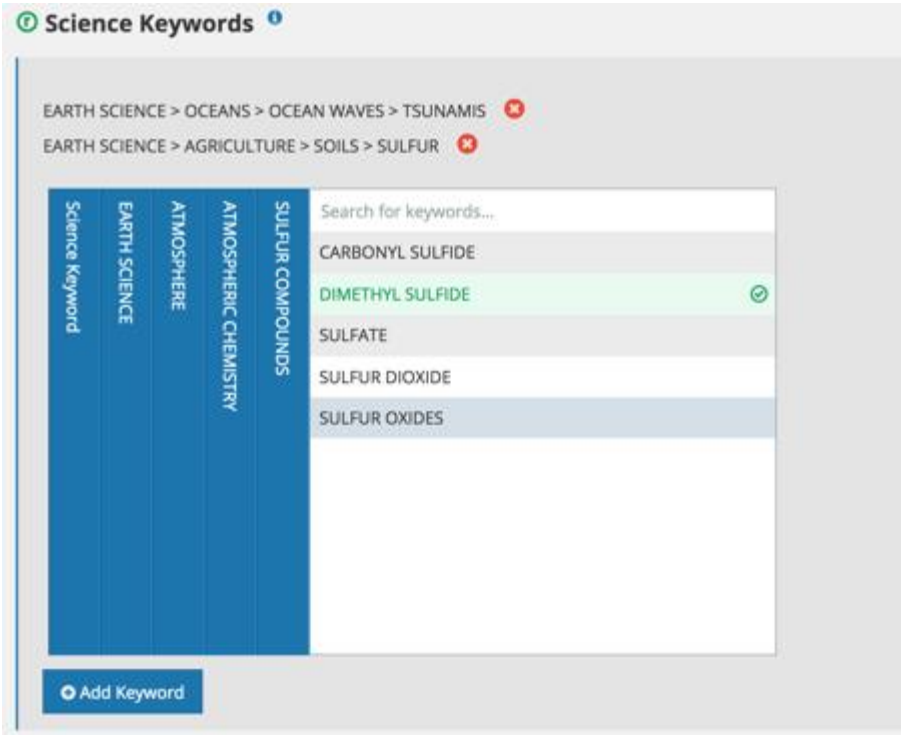
MMT Pre-Ingest Validation



CMR Ingest Validation

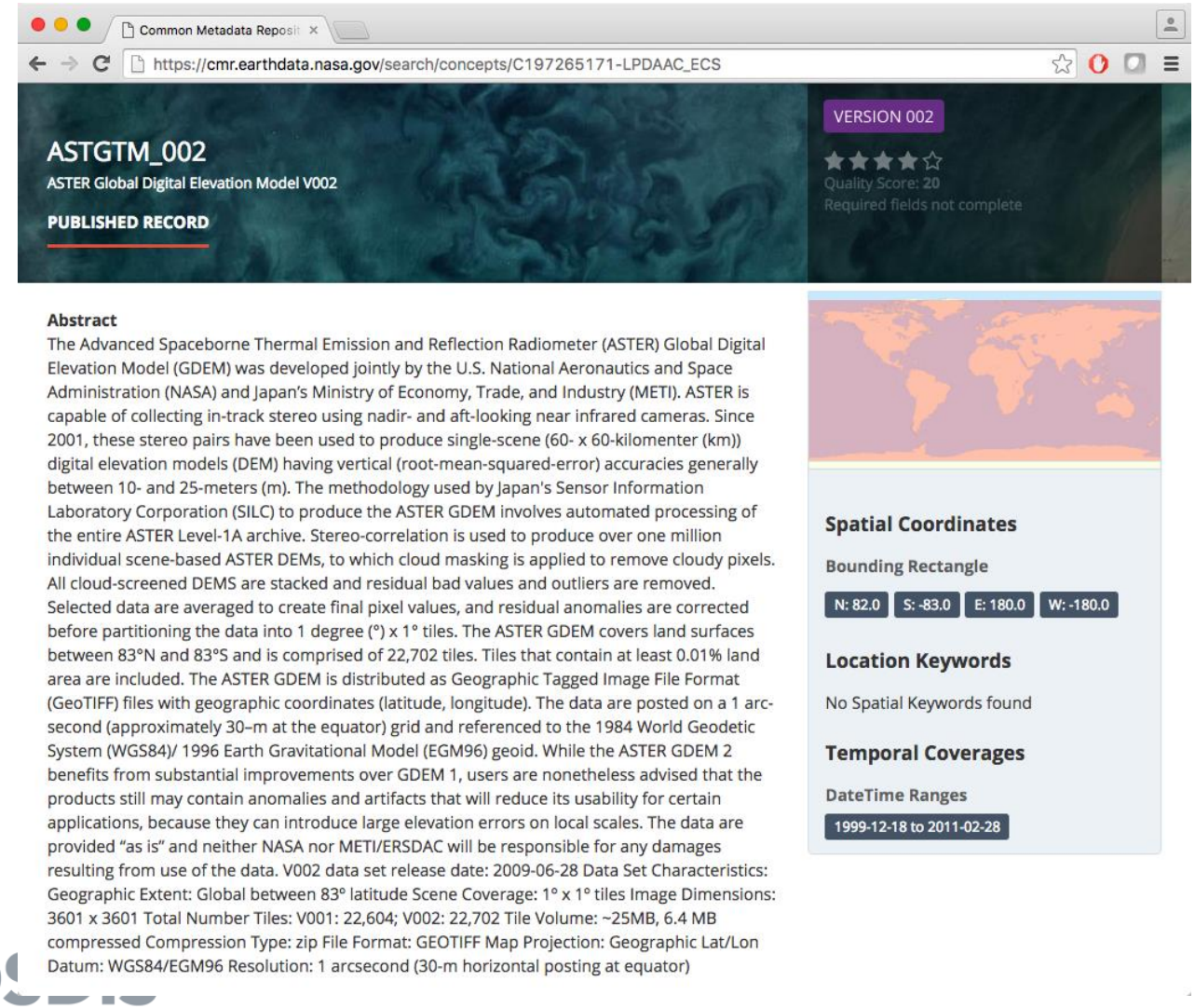


MMT and CMR Keyword Validation

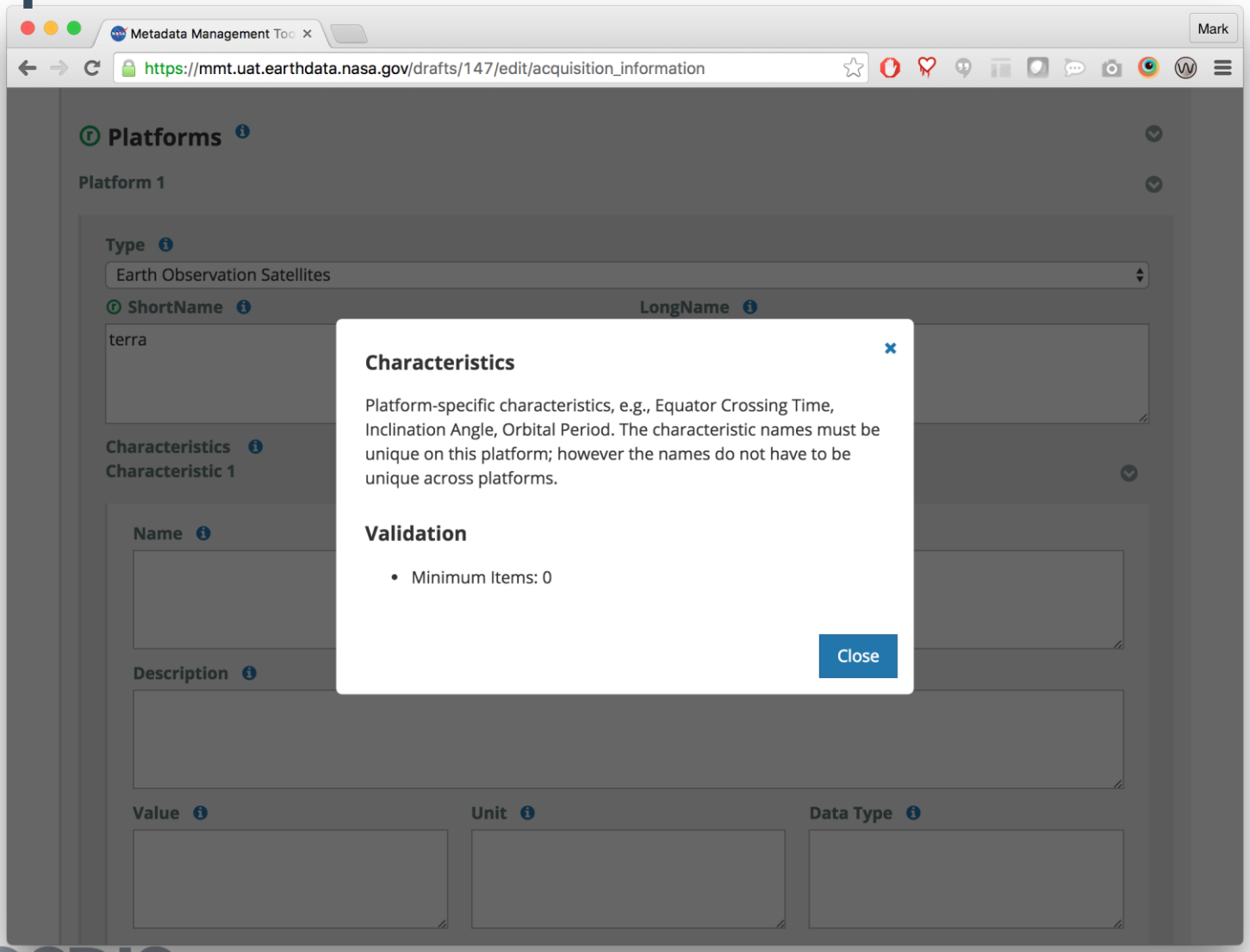


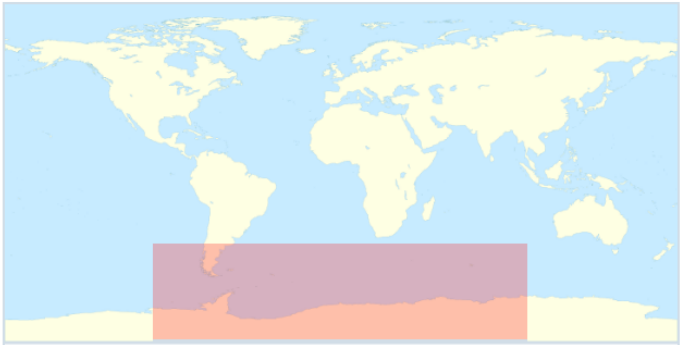

```
curl -i -XPOST … -H "Cmr-Validate-
Keywords: true" \
https://cmr.earthdata.nasa.gov/ingest/pro
viders/LARC_ASDC/validate/collection/some
NativeId -d \

 "<Collection>
  …
  <ScienceKeywords>
    <ScienceKeyword>
      <CategoryKeyword>EARTH
SCIENCE</CategoryKeyword>

<TopicKeyword>BIOOSPHERE</TopicKeyword>
      <TermKeyword>SOILS</TermKeyword>
    </ScienceKeyword>
  </ScienceKeywords>
</Collection>”

HTTP/1.1 422 Unprocessable Entity

<path>ScienceKeywords/0</path>
<error> Science keyword Category [EARTH
```

Preview Panel / Collection HTML



Help Text


Metadata Management Too
Mark
https://mmt.uat.earthdata.nasa.gov/drafts/147/edit/acquisition_information
Platforms
Platform 1
Type
Earth Observation Satellites
ShortName
terra
LongName
Characteristics
Characteristic 1
Name
Description
Value
Unit
Data Type
Characteristics
Platform-specific characteristics, e.g., Equator Crossing Time, Inclination Angle, Orbital Period. The characteristic names must be unique on this platform; however the names do not have to be unique across platforms.
Validation
Minimum Items: 0
Close



Spatial Coordinates

Bounding Rectangle

N: -38.0 S: -90.0 E: 100.0 W: -100.0

Temporal Coverages

DateTime Ranges

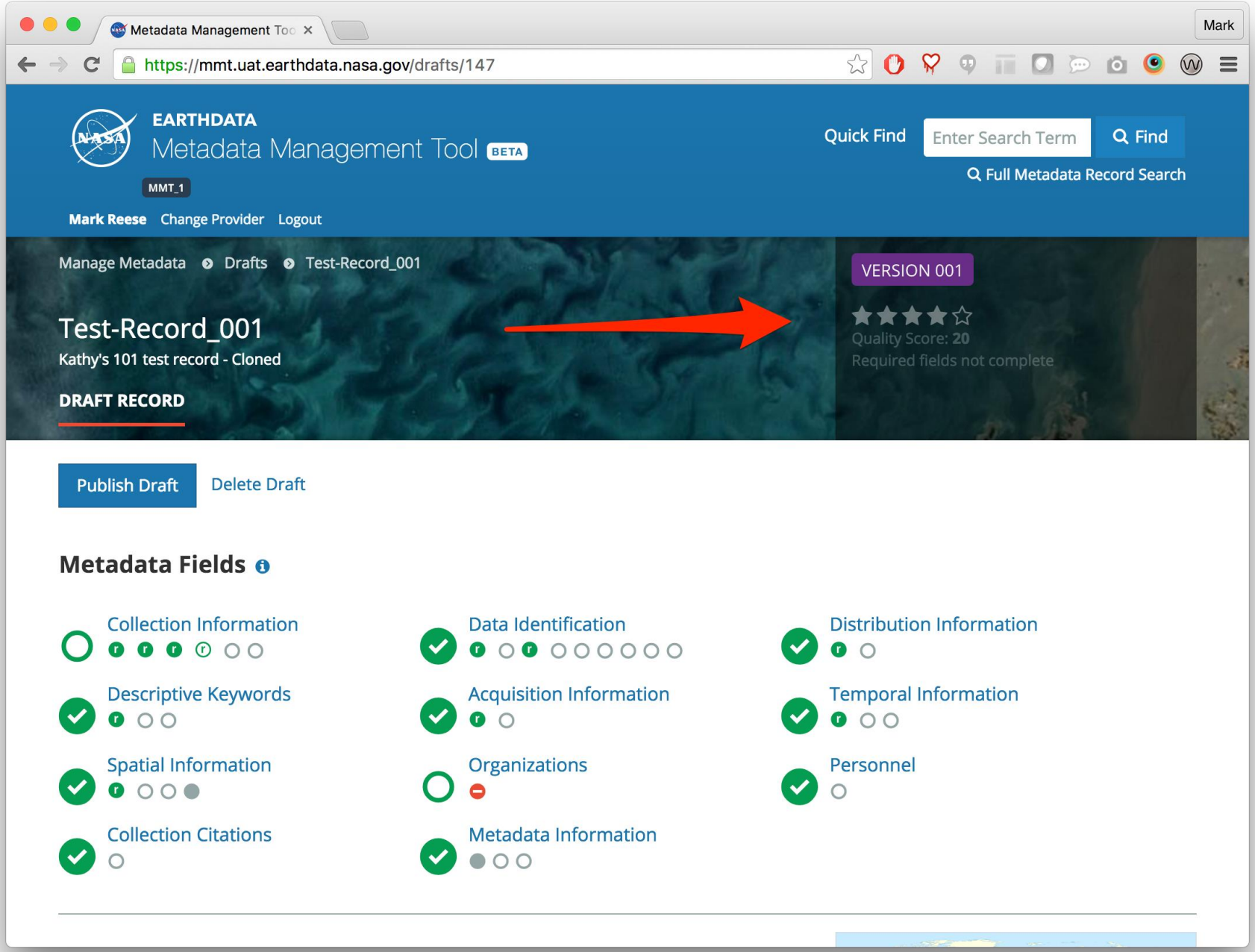
2014-06-17 to 2016-06-09

EOSDIS

METADATA QUALITY RUBRICS

Metadata Quality Rubrics

- Provide indication of completeness of Metadata based on different recommendations (UMM-C, CSW, DataCite…)
- Automatically evaluated and tagged in CMR
- Badges and guidance displayed in MMT.
  - Encourages better quality metadata
- Future
  - Reports, aggregations, search relevancy.


Metadata Management Tool
Mark
https://mmt.uat.earthdata.nasa.gov/drafts/147
EARTHDATA
Metadata Management Tool BETA
MMT_1
Quick Find
Enter Search Term
Find
Full Metadata Record Search
Mark Reese
Change Provider
Logout
Manage Metadata
Drafts
Test-Record_001
VERSION 001
Test-Record_001
Kathy's 101 test record - Cloned
DRAFT RECORD
Quality Score: 20
Required fields not complete
Publish Draft
Delete Draft
Metadata Fields
Collection Information
Data Identification
Distribution Information
Descriptive Keywords
Acquisition Information
Temporal Information
Spatial Information
Organizations
Personnel
Collection Citations
Metadata Information

EOSDIS

HUMANIZERS

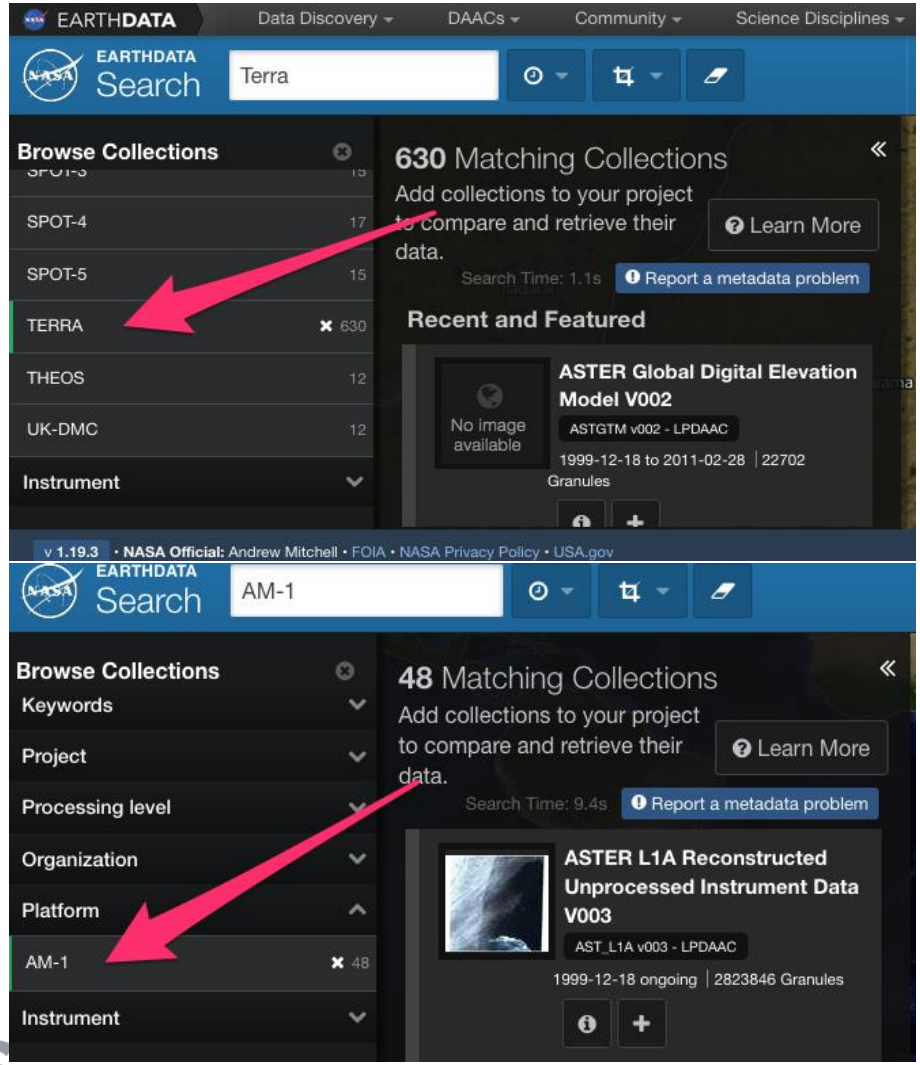
Humanizers Prototype

- Give clients clean, consistent facet data without having to change the underlying metadata.
- Temporary solution to help users while metadata is cleaned up.
- Humanizers can inform changes that are needed to metadata.

Fixing Problems in Facets

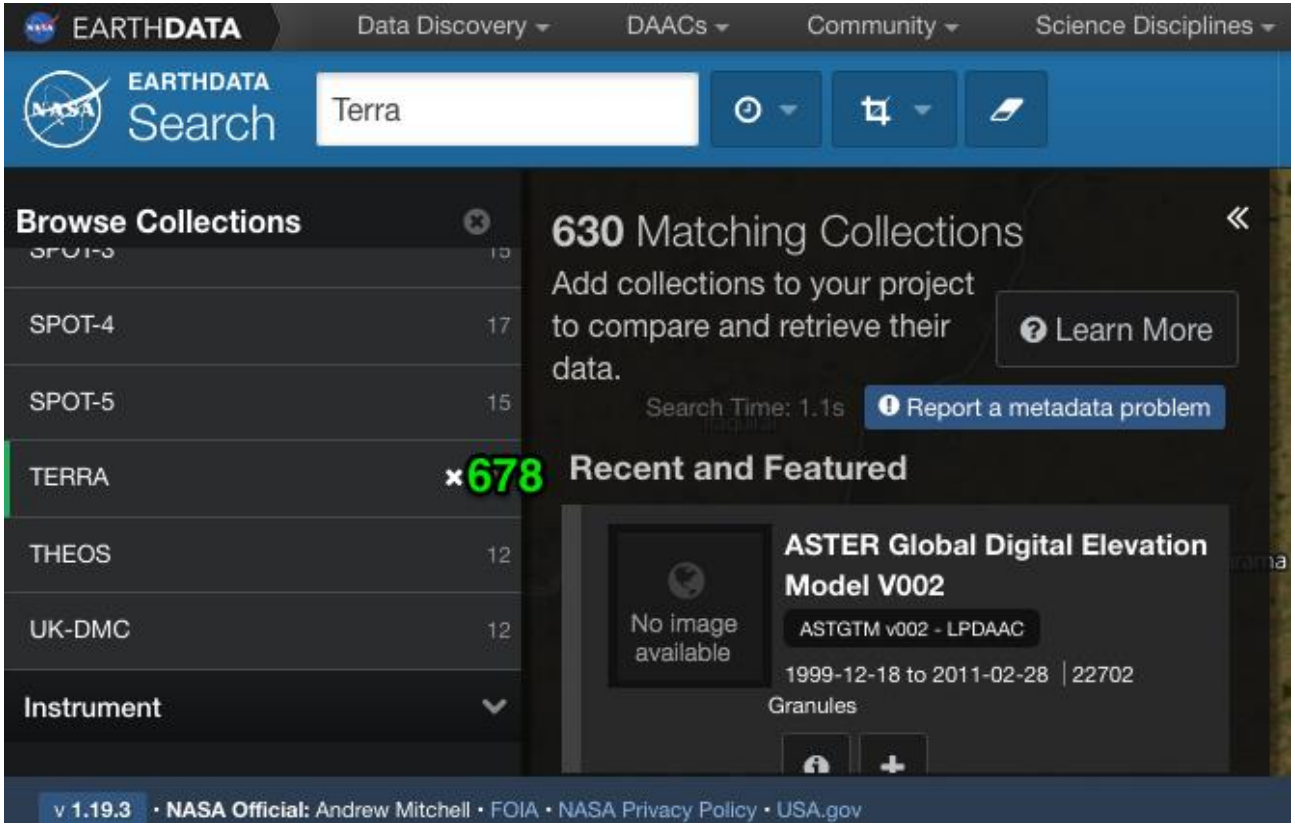
- Misspellings: “Bioosphere”
- Legacy Terms: “AM-1” instead of Terra
- Inconsistent Names: Processing levels “Level 1”, “1”
- Whitespace around element values
- Use normal case when appropriate

Before Humanizers



EOSDIS

After Humanizers



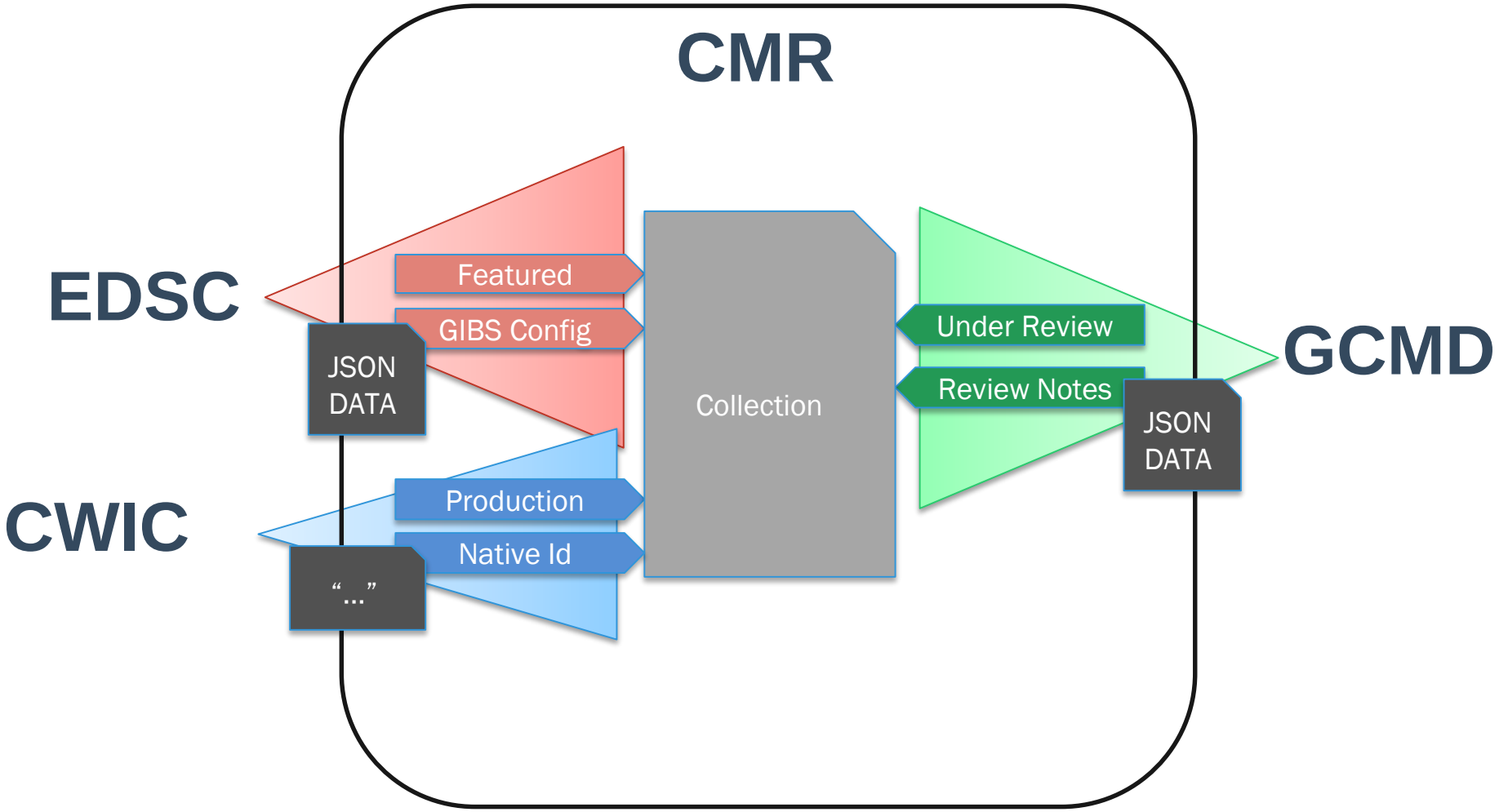
How it works

1. Admin creates humanizer instructions in CMR
2. CMR indexes impacted collections with additional humanized fields
3. Clients optionally request facets with humanized values.

Underlying metadata and existing fields are **not modified.**

TAGS

Tags
=
Enhanced Metadata



EOSDIS

SESIP-0716-JG

What is a Tag?

- Key
  - “edsc.extra.gibs”
- Description
  - “Used to associate GIBS configuration for a collection”
- Category (optional)
  - “EDSC_display”

What is a Tag Association?

- Connects a Tag with a Collection
- Tag Key
  - “edsc.extra.gibs”
- Collection Id
  - “C1000000560-NSIDC_ECS”
- Data
  - Any arbitrary JSON up to 32K. Single strings will be made searchable.

Tag Data can be Retrieved

```
curl "https://cmr…./search/collections.json?\
        tag_key=edsc.extra.gibs\
        &include_tags=edsc.extra.gibs”
…
"tags" : {
        "edsc.extra.gibs" : {
          "data" : [ {
            "maxNativeZoom" : 5,
            "antarctic" : false,
            "product" : "GHRSST_L4_G1SST_Sea_Surface_Temperature",
            "geo" : true,
            "arctic" : false,
            "title" : "Sea Surface Temperature (L4, G1SST)",
            "source" : "Multi-mission / GHRSST",
            "match" : {
              "time_start" : ">=2010-06-21"
            },
            "format" : "png",
            "resolution" : "1km"
          } ]
        }
…
```

Tag Data can be Searched

Find all collections:

- tagged with “org.ceos.wgiss.cwic.granules.native_id”
- with data “JPL-L2P-MODIS_A”

```
curl "https://cmr…/search/collections.json?\

tag_data[org.ceos.wgiss.cwic.granules.native_id]=JPL-L2P-MODIS_A"
```

EOSDIS

Tags Enable Many Use Cases

- Allows layering on additional information to each collection.
- Examples
  - Mark “featured” collections.
  - Categorize collections (Reviewed, approved, needs work, etc).
  - Store visualization information
  - Add a client specific id to collections.

This material is based upon work supported by the National Aeronautics and Space Administration under Contract Number **NNG15HZ39C.**

Raytheon

EOSDIS