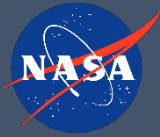# Service Metadata Refactoring
- Status Update -

CEOS WGISS-48
Hanoi, Vietnam
October 8-11, 2019
Hosted by VNSC VAST

Valerie Dixon
NASA EOSDIS
Valerie.dixon @nasa.gov

WGISS = Working Group on Information Systems and Services
CEOS = Committee on Earth Observing Satellites
EOSDIS = Earth Observing System Data and Information System

VNSC = Vietnam National Space Center
VAST = Vietnam Academy of Science and Technology
NASA = National Aeronautics and Space Administration

Initial Issue - Recap

- Single Unified Metadata Model – Services (UMM-S) was becoming too complex to manage well
  - Difficult for Curators; Required fields… for which use-case?
- Took a closer look at some key use cases…
  - End-to-End (E2E) services
  - Smart Handoffs
  - Downloadable Tools
  - Service Entry Resource Formats (SERFs)
- … and found that the Smart Handoff and Downloadable Tool use cases overlapped substantially re: needed metadata fields
  - SERFs run the gamut; to be sorted per their usage
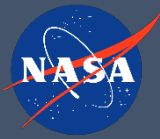
Selected Solution: Two Models

- New UMM-Tools for "Front-End" (Web User Interface) tool access
    - Downloadable Tools (basic)
    - Smart Handoffs (advanced)
- Streamline UMM-Services for "Back-End" (End-to-End) services
    - Sub-setting, Re-projections, Reformatting, Open Geospatial Consortium (OGC) Services
- Timeline
    - ESDIS Standards Office (ESO) Review: July 8 - August 30
    - Schema completion planned by end of 2019
    - Implementation in EOSDIS tools by early 2020
        - Common Metadata Repository (CMR), draft Metadata Management Tool (dMMT), Metadata Management Tool (MMT), Earthdata Search, International Directory Network (IDN)

ESDIS = Earth Science Data and Information System
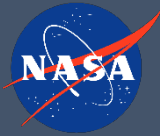
UMM-Tools - Details

- The "Front-End" Unified Metadata Model – Tools (UMM-T)
  - web user interfaces (downloadable tools, smart handoffs)

Handoff to a Tool will be surfaced in Earthdata Search / IDN Portal Collection information (if applicable):

Tool associations will be made available in a Collection's Dataset Information page (MMT, Earthdata Search, IDN Portal):
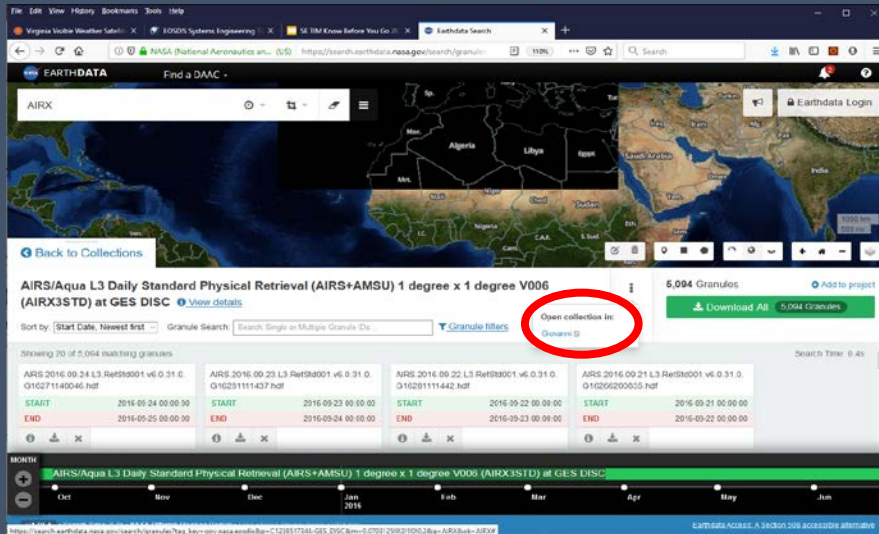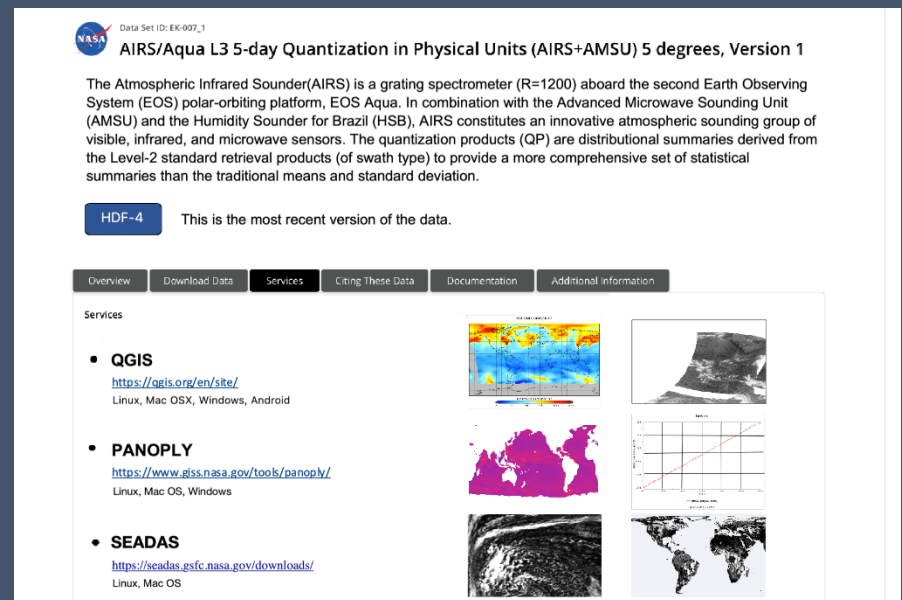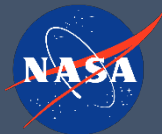

Image From https://search.earthdata.nasa.gov↑

Tools Tab mock-up courtesy ESDIS Summer Interns Alyssa Kaewwilai and Katheryn Hobart
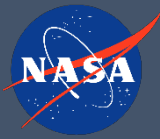
UMM-Tools - Expectations

- Providers will have to support a new model to describe their web user interfaces and downloadable tools

- MMT will support the curation of UMM-T records

- Collection Dataset Information Pages will get a 'Tools' tab

- A preliminary "ToolMatch" exercise completed by ESDIS interns associated (matched) a dozen or so visualizations tools to a sample set of data collections
  - These Tools and corresponding Collection Associations will be added to CMR once UMM-T is implemented
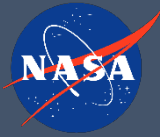
UMM-Services - Details

- The "Back-End" Unified Metadata Model – Services (UMM-S)
  - E2E services (APIs, OGC Web Services, OPeNDAP, THREDDS, etc.)
- Model Changes (Top Level Fields):
  - Deprecated Fields:
    - Science Keywords: Infer relationship using associations to collections
    - Platforms: Infer relationship using associations to collections
  - Revised Fields:
    - Type (R): Types Restricted to just "Service Types"
    - Service Organizations (R): Removed ContactPerson and ContactGroup
    - Contact Persons: Removed RelatedURLs
    - Contact Groups: Removed RelatedURLs
    - Related URL: Restrict URLContentType to CollectionURL, PublicationURL, and VisualizationURL
    - UseConstraints: Added sub-fields for license URL or license text
  - New Field:
    - URL: New field where you can directly access the back-end service

API = Application Programming Interface
OPeNDAP = Open-source Project for a Network Data Access Protocol
URL = Uniform Resource Locator
THREDDS = Thematic Realtime Environmental Distributed Data Services

UMM-Services - Expectations

- Existing E2E UMM-S records will not be impacted

- Providers will have to support the new version of UMM-S
  - There will be backwards compatibility with the previous version
  - A value for the new required field (URL) will have to be added or mapped

- MMT will support the new version of UMM-S

- The 17 existing records that describe front-end tools will be migrated from UMM-S to UMM-T
  - Metadata Associations to Data Collections will be preserved

Special thanks to the EOSDIS team

Erich Reiter
  erich.e.reiter@nasa.gov

Michael Morahan
  michael.p.morahan@nasa.gov

Doug Newman
  douglas.j.newman@nasa.gov

Alyssa Kaewwilai

Simon Cantrell
  simon.cantrell@nasa.gov

Tyler Stevens
  Tyler.B.Stevens@nasa.gov

Scott Ritz
  Scott.A.Ritz@nasa.gov

Katheryn Hobart

Thank you!
If you have any questions, please reach out:
 *valerie.dixon@nasa.gov*

Backup Slides

Tentative
Look Ahead to UMM-S/T evolution

- Whether, how, and where to handle secondary tools
  - e.g. if you convert Collection X with Tool Y, then Tool Z can work with Collection X; should we link Tool Z to Collection X?
- Whether, how, and where to handle coder support tools
  - e.g. python libraries, Jupyter Notebooks, drivers, etc. for data visualizations
  - As point of reference: Pyviz.org, holoviz.org
- MAAP project efforts to develop a UMM-Algorithm
  - Overlaps with coder support tools item above?
- UMM-Visualizations
  - to better map visualization layers to data collections
  - Current prototype utilizes UMM-S in yet another way
  - Further discussion re: Worldview planned for late 2019

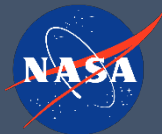MAAP = Joint ESA-NASA Multi-mission Algorithm and Analysis Platform
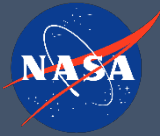
WGISS-47 Action 11

- Ad-hoc team led by Valerie Dixon (Michael Morahan, Andrea Della Vecchia, Chris Lynnes, Cristiano Lopes, Richard Moreno) to define proposal for metadata model for services addressing a set of high priority use cases. Proposed model to be circulated to WGISS-all for review and approval
    - June 6: Valerie noted a good deal of analysis and a path forward to a back-end integrated service model, and further work needed for the front-end model. The team is on track to have an updated UMM out for review by end of June.
    - July 17: SLT start reviewing UMM-S and Yves Coene shared OGC TB-15 report   ]
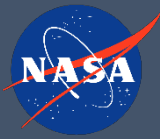
Top Level Fields for UMM-T

- Name (R)
- Long Name (R)
- Type (R)
- Version (R)
- Description (R)
- ServiceKeywords (R)
- Organizations (R)
- URL (R)

- DOI
- SupportedOutput Formats
- SupportedInputFormats
- SupportedOperatingSystem
- SupportedBrowsers
- SupportedSoftwareLanguage
- Quality
- AccessConstraints
- UseConstraints
- AncillaryKeywords
- ContactPersons
- ContractGroups
- RelatedURls
- SearchActions

A Key Use Case to Consider:

End-to-End Services (E2E)

- Enables a data transformation service to be applied to data sets behind-the-scenes:
  - User finds and selects a data set
  - User selects the desired service option
    - e.g. subset by [variable]
  - Metadata linking dataset and applicable services calls the selected service to perform upon the selected dataset
    - Done on the backend, transparent to user
    - Enabled via HTTPS API endpoint, like OPeNDAP or EGI
      - e.g. OGC Web Map Service
  - User receives transformed data set

HTTPS = Hypertext Transfer Protocol Secure   EGI = ESI (EOSDIS Service Interface) Gateway Interface

End-to-End Services (E2E):

Customizable Collection in Earthdata Search

Configure & Download Granule

E2E Customization Options
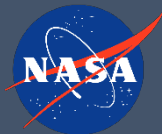
A Key Use Case to Consider:

Smart Handoffs

- Enables transferal of search results and/or context to another tool
  - User finds and selects a data set
  - User wants to view or manipulate the dataset in another tool
    - e.g. Giovanni, State Of The Ocean (SOTO)
  - User selects option to transfer search context to associated, enabled tool
    - e.g. dataset name, geospatial or temporal search parameters, etc.
  - User is redirected to the desired tool with previous search context pre-populated
    - via a well-defined API endpoint
    - via a WMS/WCS/THREDDS call
  - User carries on with their analysis in their desired tool

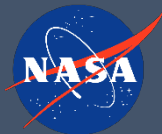WMS -= Web Map Service   WCS = Web Coverage Service
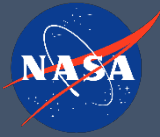
Smart Handoffs:

Collection in Earthdata Search
with a Smart Handoff
metadata association…

…to a more specialized data
extraction tool like Giovanni:

https://search.earthdata.nasa.gov/search

https://giovanni.gsfc.nasa.gov
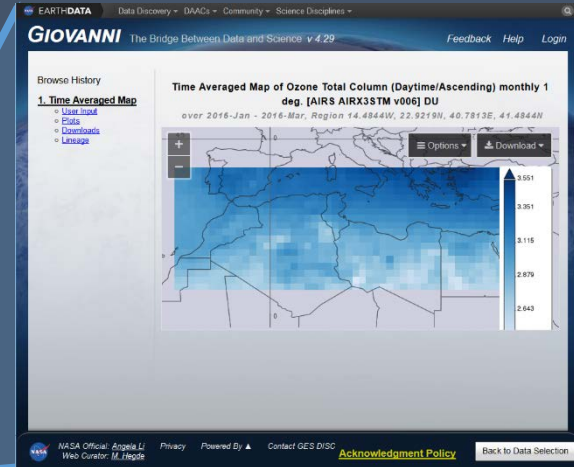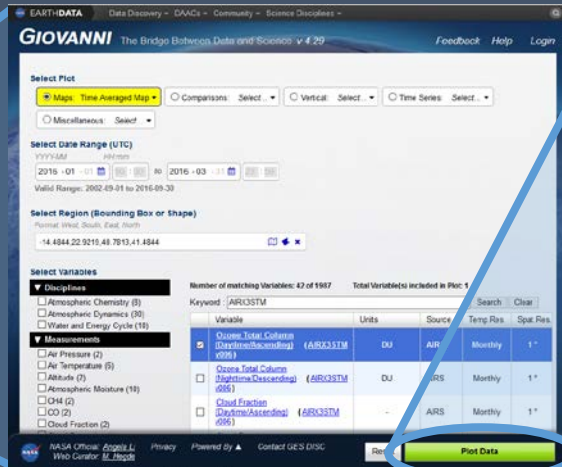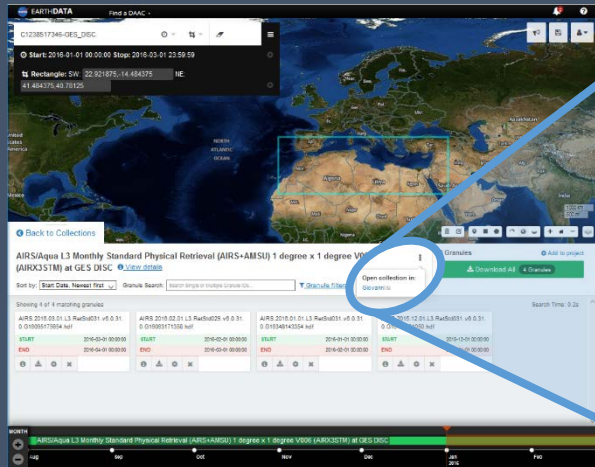
A Key Use Case to Consider:

Downloadable Tools

- Associates datasets to applicable tools which must be downloaded to the user's processing platform to use them locally
    - User finds and selects a data set
    - User wants to see what Downloadable Tools are applicable to the selected data set
        - e.g. Python, R libraries, Panoply
    - User selects desired Local Tool and is redirected to a primary splash page, containing:
        - Information about the tool,
        - How to install and use it,
        - A download option
    - User downloads the dataset to perform local processing
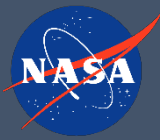
A Key Use Case to Consider:

Service Entry Resource Formats (SERFs)

- Legacy records from the Global Change Master Directory (GCMD)
  - Describe tools and services, primarily from the IDN community
  - Migration effort recently concluded
    - Identified and updated still-valid SERF records
    - Migrated valid SERF records to UMM-Service
    - Archived remaining legacy SERF records
    - Deprecated SERF DocBuilder tool

- SERF Curation concerns
  - Access to Metadata Management Tool (MMT) restricted
  - New tool for anyone to curate UMM records is coming soon!
    - The Draft Metadata Management Tool (dMMT) will allow users with an Earth Data Login credential to propose new or modifications to existing collection or service metadata records
  - Until then, send requests to support@earthdata.nasa.gov

- With Refactoring, UMM-S SERFs will need to be reallocated to applicable use-case sub-models