AGU Paper Number: IN13B-1570

ESDIS DOI System, Approach, and Future Direction

Lalit Wanchoo[1] (lalit.wanchoo@nasa.gov) and Nathan James[2] (nathan.l.james@nasa.gov)

[1] ADNET Systems, Inc., 7515 Mission Drive, Suite A100, Lanham, MD 20706 , [2] Earth Science Data and Information System Project (Code 423), NASA Goddard Space Flight Center, Greenbelt, MD 20771

Introduction

NASA's Earth Science Data and Information System (ESDIS) Project began investigating the use of Digital Object Identifiers (DOIs) in 2010 with the goal of assigning DOIs to various data products. These climate change research data products produced using Earth observing instruments and environmental models are archived and distributed by twelve Distributed Active Archive Centers (DAACs) located across the United States. Each data center serves a different Earth science discipline user community and, accordingly, has a unique approach and process for generating and archiving a variety of data products. These varied approaches present a challenge for developing a DOI solution. To address this challenge, ESDIS has developed processes, guidelines, and several models for creating and assigning DOIs. Initially the DOI assignment process was started as a prototype. Now it is fully operational. In February 2012, ESDIS started using the California Digital Library (CDL) EZID for registering EOSDIS-related DOIs.

Goal

Assign DOIs to over 5000 data products that are archived and distributed by EOSDIS. Provide users a mechanism of citing data and getting to the data irrespective of the archive location of the data products or the responsible data provider for the distribution

Objective

Provide confidence in availability and reproducibility of the referenced data products and ensure that such products and their documentation are discoverable by future data users.

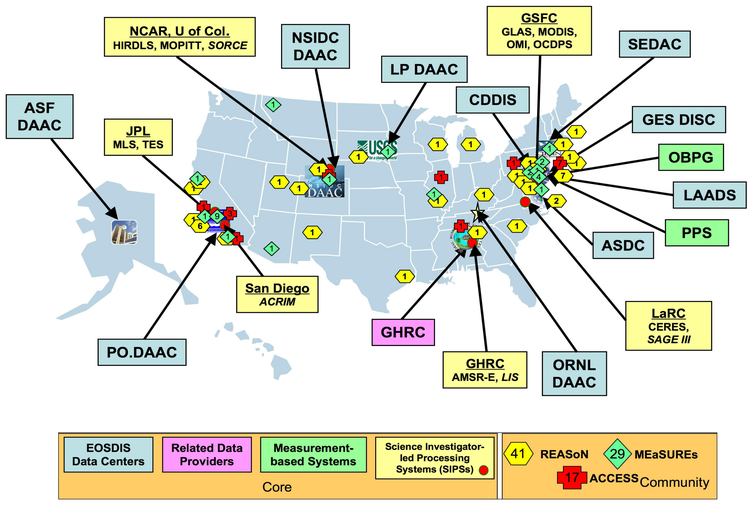
Challenges

1. Numerous discipline-based Data Centers responsible for processing, archiving, and distribution of data products
2. DOI structure (models) that is meaningful and address all possible variations
3. Establish DOI name uniqueness across all data centers.
4. Copies of the same products may be available from different data centers
5. Data products are from currently active missions produced daily and legacy data that has been already processed.
6. Reprocessing leading to multiple product versions. Multiple versions coexist in some cases, but in other cases later versions supersede older versions.

Data Centers

NASA Distributed Active Archive Centers (DAACs) are NASA Data Centers collocated with science discipline expertise. DAACs are data "publishers"; they archive and distribute standard data products produced by the Science Investigator-led Processing Systems.

Map Showing Location of Data Centers



ESDIS DOI Identifier Models

A DOI consists of two part alphanumeric string: doi:[prefix]/[suffix].
The DOI prefix assigned to ESDIS by EZID is 10.5067
The suffix is composed of alphanumeric string which must be unique and should not contain information that might change over time.
DOI Suffix models proposed by ESDIS :

- A) [mission]/[instrument]/data[m][n]
- B) [program]/[measurement group]/data[n]
- C) [mission]/[instrument]/[product identifier]
- D) [EOSDIS/DAACname]/unique numeric string
- E) [unique-numeric-DAAC identifier]/unique numeric string

[m] depicts the processing level of the product
[n] depicts a sequence number that is assigned on a first-come-first-serve basis. New product versions receive new DOI's and are assigned a new sequence number.

DOI Status Definition

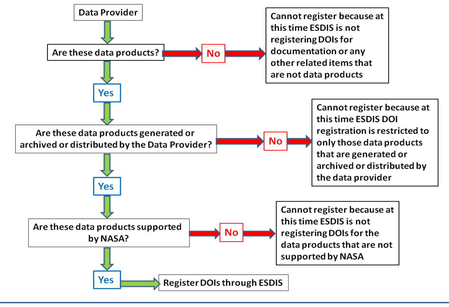
- **Under Review:**
  DOIs are under review when data provider has provided information to ESDIS and DOI suffix and the metadata attributes are being finalized .
- **Reserved:**
  DOIs are reserved by ESDIS when the data provider has not started distributing the data products or does not have accessible landing pages or all the mandatory data products metadata are not available. Purpose will be to reserve the DOIs with ESDIS to ensure unique identifiers. This information will be only visible from the ESDIS wiki website.
- **Registered:**
  DOIs are registered with the EZID once the data provider has provided all the mandatory data products metadata to the ESDIS, have accessible landing pages and have started distributing the data products to the public users. Registered DOIs are accessible by the general public using

DOIs in Product Metadata

a) Attribute Name Identifier
The attribute name: "identifier_product_doi"
The attribute value: "10.5067/MEASURES/GSSTF/DATA302"

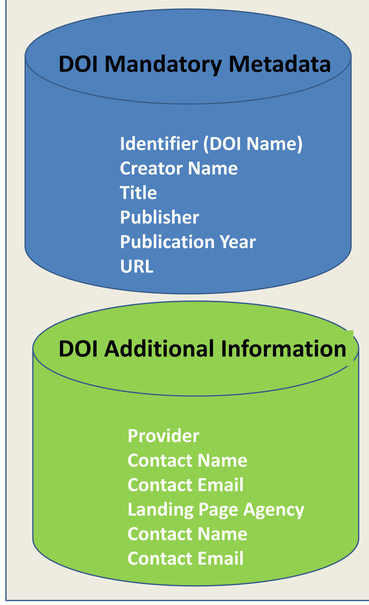
b) Attribute Name Authority
Provides the permanent service for resolving DOIs to their URL
The attribute name: "identifier_product_doi_authority"
The attribute value: "http://dx.doi.org/"

Identification of the Products that Get DOIs

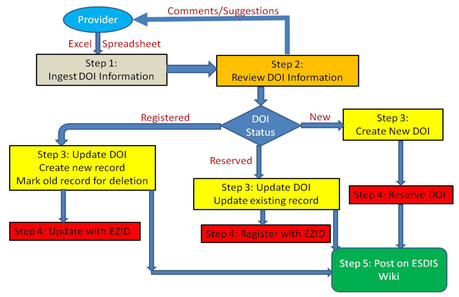


ESDIS DOI WIKI URL: https://wiki.earthdata.nasa.gov/display/DOIsforEOSDIS

DOI Registration Metadata Requirements



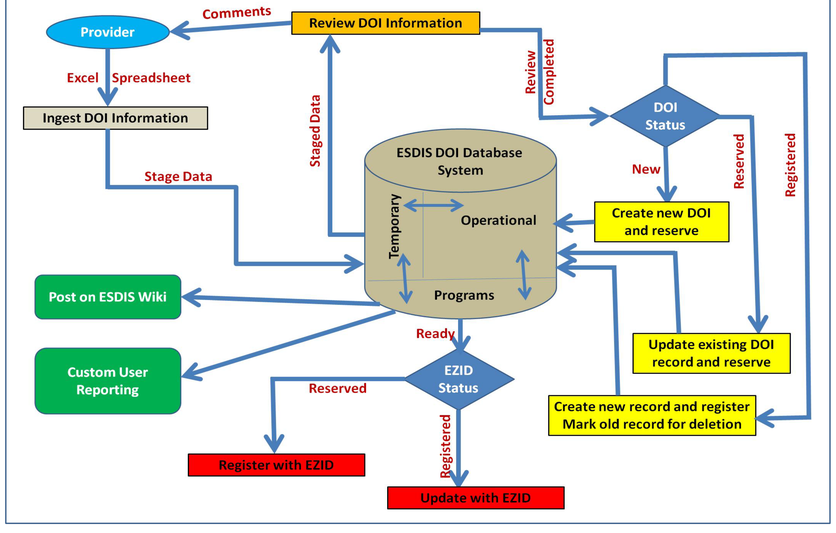
ESDIS DOI Registration Process



ESDIS DOI Registration Status

Number of Data Providers (Provided DOIs)	= 9
Number of Missions/Projects	= 22
Number of DOIs Reserved	= 139
Number of DOIs Registered	= 193
Number of DOIs Under Review	= 26
Total Number of DOI Requests	= 358

Future ESDIS DOI System



Acknowledgments

Nate James's contributions to this study were made as a part of his employment by NASA and Lalit Wanchoo's contributions to this study were funded through the Science and Exploration Data Analysis (SESDA III) GSFC NASA Contract No: NNG12PL17C.