

DOI Pilot Briefing

April 2017

Background

- At the November F2F, PDS launched a pilot project for creating and registering PDS4 data using Digital Object Identifiers (DOI)

Participants

- Mike A'Hearn, SBN
- Nancy Chanover, Atmos
- Dan Crichton, EN
- Steve Hughes, EN
- Ron Joyner, EN
- Todd King, PPI
- Ludmilla Kolokolova, SBN
- Carole Neese, PSI
- Eric Palmer, PSI
- Anne Raugh, SBN
- Ray Walker, PPI

Approach

Three telecons were held to discuss

1. Options for registering DOIs
2. Mappings to PDS4
3. Testing and results for the DOI service using DN PDS4 products

DOI Registration Options

- DOIs are registered in DataCite
- To obtain and register a DOI, PDS needs to work through a provider. Three providers in the US are recognized by DataCite. PDS explored two of the three.
 - **Option 1: EZID through the UC System**
 - **Option 2: OSTI**
 - Option 3: Purdue Library System

DOI Assignment Organization

- While EZID is used by NASA ESDIS, it is no longer accepting new clients due a renegotiation of their contract.
- The WG then worked with OSTI to test their IAD service.

OSTI's Interagency Data ID Service (IAD) allows U.S. federal agencies to obtain persistent identifiers known as Digital Object Identifiers (DOIs) for datasets and ensures that the DOIs are properly registered with DataCite. The Office of Scientific and Technical Information (OSTI) became a member and an allocating agency for DataCite in 2011. OSTI can assign DOIs to the Department of Energy through its DOE Data ID Service

OSTI IAD Service

- Supports registration of DOIs with DataCite
- Ensures unique DOI assignments are provided
- Provides web-based and APIs for registering data
- Will enter into an IAA to provide support to groups for using their service @ \$2500/year

Structuring DOIs

- A DOI consists of multiple parts
 - <http://doi.org/DOI-prefix/DOI-infix/DOI-suffix>
 - **DOI-prefix** (fixed by DataCite) – the unique number assigned to PDS: doi.org/**10.19597**
 - **DOI-infix (optional)** – product specific “intelligent” string (for example, : doi.org/10.19597/**pds_product**)
 - **DOI-suffix** (fixed by DataCite) – unique DOI number: doi.org/10.19597/pds_product/**1105143**
- DataCite ensures that no two data clients will ever have the same DOI prefix.
- ***OSTI IAD ensures that suffixes are unique.***

DOI to LID/LIDVID Mapping

- Created DOI to LID/LIDVID mapping
 - Mapped PDS4 to OSTI XML structure
 - Tested with a set of Bundle & Collection products provided from the DNs:
 - ATMOS_LADEE_NMS_BUNDLE
 - ATMOS_MPF_MPIM_BUNDLE
 - PPI_MAVEN_MAG_CALIBRATED_DATA_BUNDLE
 - PPI_MAVEN_SWEA_BUNDLE
 - PPI_MAVEN_SWEA_3D_DATA_COLLECTION
 - SBN_BOPPS_2014_BUNDLE
 - SBN_BOPPS_2014_BIRC_CAL_IMAGES_COLLECTION
 - SBN_LADEE_LDEX_BUNDLE

DOI OSTI Registration

- Registered DN Bundles & Collections with OSTI service
 - They were registered with OSTI but not all the way to DataCite this was a test
- Linked landing pages for PDS4 products
- A “landing page” or introductory page describes the registered data.
 - The DOI must point to a landing page, which in turn, “links” to the data products.
 - Landing pages are linked to the node resources that are registered with EN:
 - Bundles
 - Collections
 - Documents (being deployed with build 7b)

Sample DOI XML Label (DOI request)

```
<?xml version="1.0" encoding="UTF-8"?>
<records>
  <record>
    <title>MAVEN MAG Calibrated Data Bundle</title>
    <creators>J.Mafi</creators>
    <creators_emails>jmafi@igpp.ucla.edu</creators_emails>
    <publisher>PDS Planetary Plasma Interactions Node (PPI)</publisher>
    <publication_date>1/24/2017</publication_date>
    <site_url>https://pds.jpl.nasa.gov/dsview/pds/viewBundle.jsp?identifier...</site_url>
    <product_type>Collection</product_type>
    <product_type_specific>PDS4 Bundle</product_type_specific>
    <product_nos>urn:nasa:pds:maven.mag.calibrated</product_nos>
    <description>This bundle contains...</description>
    <keywords>PDS; Rainforest; Carbon Dating; Remote-Sensing</keywords>
    <related_resource></related_resource>
    <contributor_organizations></contributor_organizations>
    <sponsor_org>National Aeronautics and Space Administration (NASA)</sponsor_org>
    <contract_nos></contract_nos>
    <other_identifying_nos></other_identifying_nos>
    <file_extension></file_extension>
    <availability>NSSDCA</availability>
    <contact_name>PDS OPERATOR</contact_name>
    <contact_org>PDS</contact_org>
    <contact_email>pds-operator@jpl.nasa.gov</contact_email>
    <contact_phone>555-555-5555</contact_phone>
  </record>
</records>
```

<blue> -- values are static

<red> -- values derived from registry contents

<black> -- values inferred from Node affiliation; captured in Tracking Service

Email Response from Registration

- Date: 01/26/2017
- Site Code NASA-PDS
- Number of Records Submitted: 1
- Number of Records Successfully Loaded into IAD: 1
Number of Records Failed: 0

- Record Number: 1
- OSTI ID: 1327241
- Report Number: urn:nasa:pds:maven.mag.calibrated
- Title: MAVEN MAG Calibrated Data Bundle
- Contract Number:
- DOI Assigned: 10.5072/1327241
- Status: SUCCESS Record added and processed by DataCite.

OSTI Fall Meeting

The Department of Energy's Office of Scientific and Technical Information (OSTI) will hold a 2-day research data focused workshop September 13-14th, 2017 at OSTI in Oak Ridge, TN. The goal of the workshop is to share information about OSTI's Data Services, gain a deeper understanding of researchers' data needs, and determine how to better support those needs.

Below is a working list of workshop sessions. We'd welcome any feedback about session topics, presenters, panel participants, etc.

- **Updates since 2016 Data ID Workshop**
- **Obtaining data DOIs via the DOE Data ID Service and Interagency Data (IAD) Service** – (a break-out informational session for potential data clients)
- **Defining data "types"** – (controlled vocabulary for data types: what should change, how, and why)
- **Discovering data** – (How do researchers/program managers/public search for data? How should these needs translate into enhancements to OSTI's data products?)
- **Data publishing** – (Data citation, Public Access for digital data, Data papers and Data journals, etc.)
- **DataCite Update** – (Event data, Standardization of organizational identifiers, ORCID/CrossRef/ DataCite integrations)
- **Software** – (Software as data, OSTI's new product DOE Code {}, registering software with DOIs)
- **Data in Today's Science** – (Open floor for talking through common issues and questions. Success stories too)

We are happy to offer lodging suggestions near OSTI for those who are interested.

So that we can plan appropriately, please let us know by **May 1, 2017** via email to DOEDataID@osti.gov if you will try to attend.

Status and Recommendations

- The working group met on March 16 and reviewed results of the pilot.
 - Results of tests were successful and OSTI provides a useful service
 - The WG recommends that PDS move forward with OSTI
- OSTI has agreed to turn on the registration service as soon as we are ready with DataCite
 - IAA can be put in place in parallel
- Once registrations are complete with DataCite, the working group recommends demonstrating how citations will work in papers.

Backup