



# DOIs and Persistent Identifiers for PDS4 Products

Action Item: 2015-07-13/A

August 2015

# Action Items

- 2015-04-21/A (Crichton, 2015-05-11):  
Investigate DOI registration costs presented by Tom Statler.
- 2015-07-13/A (Crichton, 2015-08-10):  
Investigate conditions under which PDS could join DataCite

# Who Keeps Track of DOIs?\*

## Primary Entities in the US:



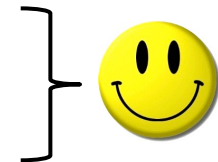
A well-established global consortium of publishers, libraries, private foundations, and (a few) government agencies that assigns and serves as the registry for publication DOIs. CrossRef provides this service free of charge (i.e., generally the cost is borne by publishers).



A similar global consortium for data, led primarily by research university libraries (esp. California Digital Library & Purdue University Library). DataCite provides DOI registry services for a fee.

DataCite and CrossRef are collaborating, as of Nov 2014, in order to:

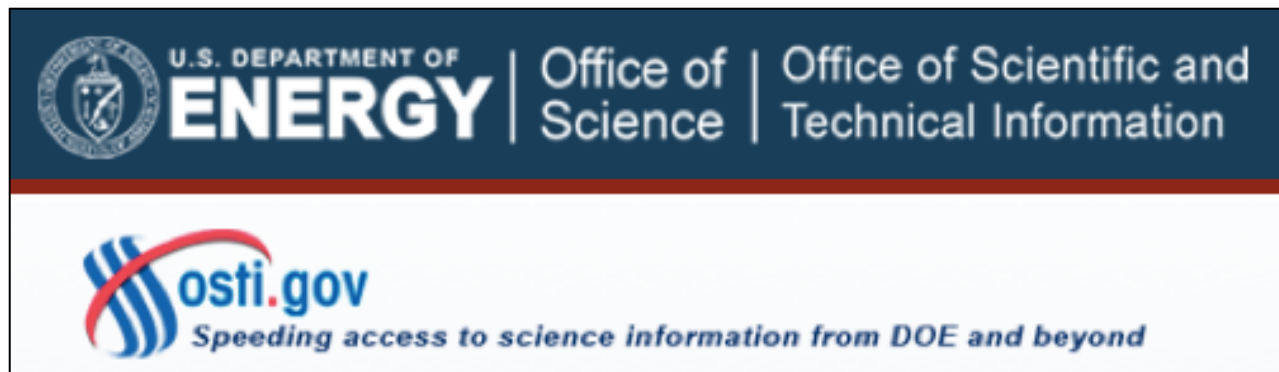
- Enhance interoperability;
- Provide support for interlinking between articles and data;
- Integrate with other initiatives such as ORCID and FundRef.



\* Credit: Tom Statler's PDS MC presentation

## Who Can Provide Registry Services for US Govt?\*

There is one (1) USG entity that is a member of DataCite:



OSTI's Interagency Data (IAD) Web Service allows U.S. federal agencies to obtain DOIs for datasets and ensures that the DOIs are properly registered with DataCite.

Suggestion: PDS management work with DOE/OSTI/IAD to introduce and maintain data DOIs in PDS.

\* Credit: Tom Statler's PDS MC presentation

# DataCite Investigation

- For 2015-07-13/A (Crichton, 2015-08-10): Investigate conditions under which PDS could join DataCite
- Reviewed DataCite Rules and governance structure
- Discussed DOI activities with NASA ESDIS

# DataCite

DataCite does not allocate persistent identifiers itself; this is done by its members, who act as allocating agents.

In the U.S., there are three full members

- California Digital Library (through UC Irvine)
  - OSTI/DOE
  - Pursue University Libraries
- 
- Membership seems to only be open to countries without representation

# DataCite

- Membership Types:
  - Member
    - Participate in WG; have voting rights; be an allocation agency
    - Fee: 8,500€
  - Affiliated membership
    - Can serve as an observer in DataCite
    - Does not allocate DOIs
    - Fee: 1,500€
- Groups only interested in assigning DOIs work with member agents.
- Membership is governed by a set of bylaws which identify the duties of a Full and Affiliated Member  
<https://www.datacite.org/sites/default/files/document/DataCite%20Statutes%202014.pdf>

# NASA ESDIS

- Has developed a set of procedures to acquire DOI assignments for DAACs
  - <https://wiki.earthdata.nasa.gov/display/DOIsforEOSDIS/ESDIS+DOI+Process>
  - Centralized assignments of DOI for NASA DAACs
- Uses EZID provided through the California Digital Library for DOI allocations
  - One of the DataCite Members, similar to DOE/OSTI
  - \$2500/year for the allocation
  - GSFC manages contract for NASA DAACs
- Working to automate the process



## ESDIS DOI Registration Process

Step 1: Data Provider interested in registering DOIs, please visit ESDIS DOI wiki for DOI related information: (<https://wiki.earthdata.nasa.gov/display/DOIsforEOSDIS/>). Access to wiki is restricted to registered users only. Users need to register with Users Registration System. To register, please visit URS at <https://urs.eosdis.nasa.gov/> and sign up for an account.

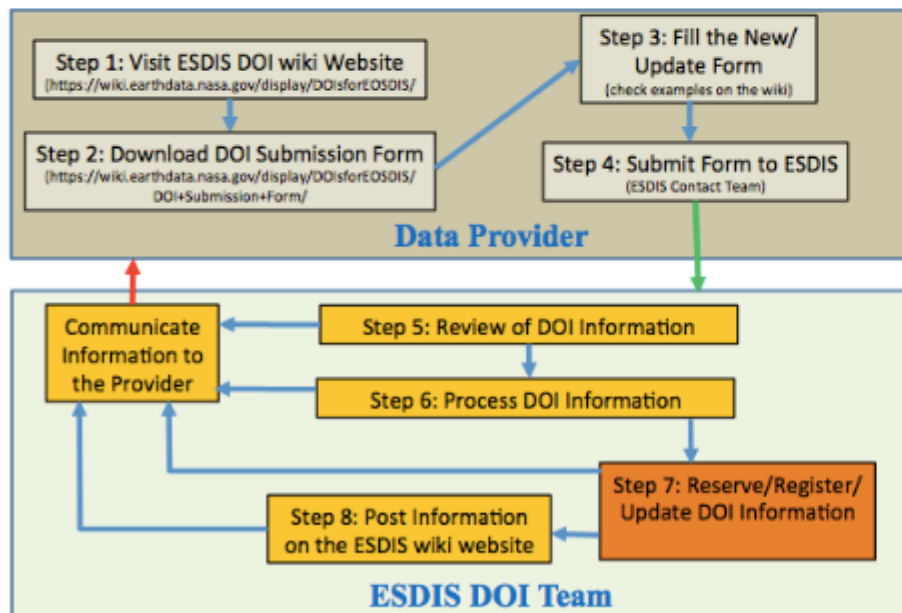
Step 2: Contact ESDIS DOI Staff to review the procedure for developing DOIs, review a DOI model that will suit data provider's data products, review the EZID registration process and required metadata, go over any questions or clarification. ESDIS can setup a teleconference with the data provider staff for this step.

Step 3: Data Provider to make final selection of a DOI model and generating DOIs with the metadata as required by EZID (refer Table 1). If Data Provider would like ESDIS to develop DOIs then Data provider should provide list of data products with metadata information to ESDIS and ESDIS will generate DOI list.

Step 4: Data Provider to send the DOI list to ESDIS for feedback.

Step 5: Once the DOI and related information is finalized, ESDIS will reserve or register DOIs with the EZID based on the completeness of the DOI's metadata information.

The ESDIS DOI Registration Process is explained in more detail in the diagram below.

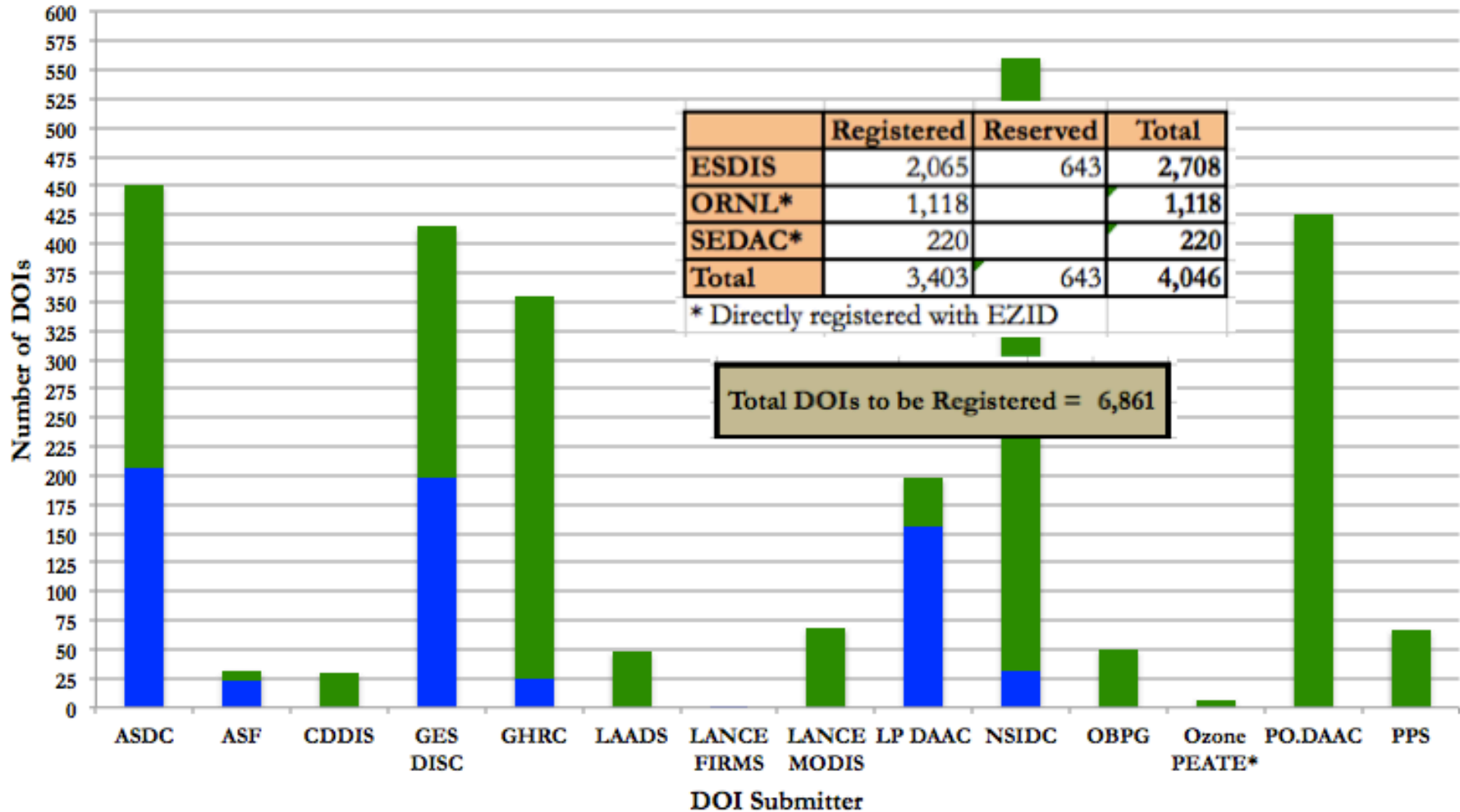


## ESDIS DOI Workflow

Chart Area

## Status of Digital Object Identifiers (DOI) Created with the ESDIS by the Data Providers as of 31 July 2015

■ Registered: Created with the ESDIS and also registered with the EZID   ■ Reserved: Created with the ESDIS but not registered with the EZID



DOIs assigned at mission, instrument, dataset level

## **“EZID”**

- Managed by the California Digital Library @ UC Irvine
- \$2500/year for DOI Assignments
- Many science groups are using them (NCAR, NOAA, ESIP, etc)

## EZID Partners and Clients

New in September

- Facebase Consortium <https://www.facebase.org/>
- Landscape Conservation Cooperative <http://lccnetwork.org/>
- University of Southern California's ISI Informatics Systems Research Division  
<http://www.isi.edu/research/informatics/>

Complete list

- UC Berkeley Library (on behalf of the UC Berkeley campus) <http://www.lib.berkeley.edu/>
- UC Davis Library (on behalf of the UC Davis campus) <http://lib.ucdavis.edu/>
- UC Irvine Libraries (on behalf of the UC Irvine campus) <http://www.lib.uci.edu/>
- UC Los Angeles Library (on behalf of the UC LA campus) <http://www.library.ucla.edu/>
- UC Merced Library (on behalf of the UC Merced campus) <http://library.ucmerced.edu/>
- UC Riverside Library (on behalf of the UC Riverside campus) <http://library.ucr.edu/>
- UC San Diego Library (on behalf of the UC San Diego campus) <http://libraries.ucsd.edu/>
- UC San Francisco Library (on behalf of the UC San Francisco campus) <http://www.library.ucsf.edu/>
- UC Santa Barbara Library (on behalf of the UC Santa Barbara campus) <http://www.library.ucsb.edu/>
- UC Santa Cruz Library (on behalf of the UC Santa Cruz campus) <http://library.ucsc.edu/>
- American Astronomical Society <http://www.aas.org>
- American University Library (on behalf of the AU campus) <http://www.american.edu/library/>
- Bay Area Environmental Research Institute <http://baeri.org/>
- Benchfly <http://www.benchfly.com>
- Broad Institute of MIT and Harvard <http://www.broadinstitute.org>
- Brown University Library <http://library.brown.edu/>
- California Institute of Technology Library (on behalf of the CalTech campus) <http://library.caltech.edu/>
- Cancer Imaging Archive <http://www.cancerimagingarchive.net>
- Carnegie Mellon University Library (on behalf of the CMU campus) <http://search.library.cmu.edu/>
- Center for Open Science <http://centerforopenscience.org/>
- Center for Research in Computational Linguistics (CRCL) <http://sealang.net>
- Centre national de documentation pédagogique <http://www.cndp.fr/>
- Columbia University Libraries/Information Services on behalf of Columbia University  
<http://library.columbia.edu/>

# In Summary

- DataCite limits the number of institutions that can be full members to assign DOIs. Members are expected to contribute substantially to the working groups (<https://www.datacite.org/sites/default/files/document/DataCite%20Statutes%202014.pdf>)
- We could become an associate member of DataCite, but that allows us to participate in working groups. It's not for assignment of DOIs. <https://www.datacite.org/join-datacite>
- Both DOE/OSTI and UCI/EZID act in the capacity of allocating DOIs and registering them on one's behalf.
- Several science repositories use EZID (NCAR, NOAA, ESIP, etc) use EZID
- ESDIS submits DOIs to EZID for registration
- ESDIS does the registration, etc, on behalf of the DAACs (<https://wiki.earthdata.nasa.gov/display/DOIsforEOSDIS/EOSDIS+DOIs+Status+and+Listing>)
- For ESDIS, registration seems to be at an instrument level and not a granule so approximately 4000 registered objects.