AGU Paper Number: IN13C-1854

Developing Data Citations from Digital Object Identifier Metadata

Nathan James (nathan.l.james@nasa.gov) Lalit Wanchoo (lalit.wanchoo@nasa.gov)
Earth Science Data and Information System Project (Code 586), NASA Goddard Space Flight Center, Greenbelt, MD 20771
ADNET Systems, Inc., 7515 Mission Drive, Suite A100, Lanham, MD 20706

Introduction

Data citations are gaining significant attention from the scientific community and data centers because they provide identification of the specific data used in the article. This makes it easier for a scientists to discover the data and validate the findings as reported in the publication. In addition, the data creator and the data distributors receive credit for the data. This procedure also addresses the reproducibility and transparency of the data, which has become increasingly more critical for the advancement of the scientific analyses. Creating a data citation may pose small challenges, but it provides benefits to the both scientific community and the data center. Furthermore, by providing proper data citations, it is likely that the data products being easily discoverable there can be an increase in the data distribution.

One of the benefit of data citation with identifier would be to determine the data products usage for various research activities by searching the published data. NASA's Earth Science and Data Information System (ESDIS) Project has taken the first step towards facilitating the use of data citations for those products by registering identifiers and establishing guidelines for displaying a sample data citation on each identifier landing page.

Digital Object identifier

Digital Object Identifiers (DOIs) are being assigned to datasets for providing access to the data and assessment of the data usage by tracking the citations. DOIs have been implemented as a part of the metadata standards and various applicable international standards by various agencies. ESDIS has been processing information for the registration of DOIs for the past five years, and an automated system for assigning DOIs has been in operation for two years. The ESDIS DOI registration system registered over 4,200 DOIs thus far, and over 490 DOIs are reserved for later use until all required information has been collected or the data product is made available to the end user.

ESDIS Metadata Attributes Requirements

DOI Registration

Identifier
Creator Name
Title
Publisher
Publication Year
URL

DOI Landing Page

Identifier
Data Set Long Name
Version No.
Data Set Description
Data Access
Data Citation
Associated Landing Pages

Acknowledgments

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

Data Citation Approaches

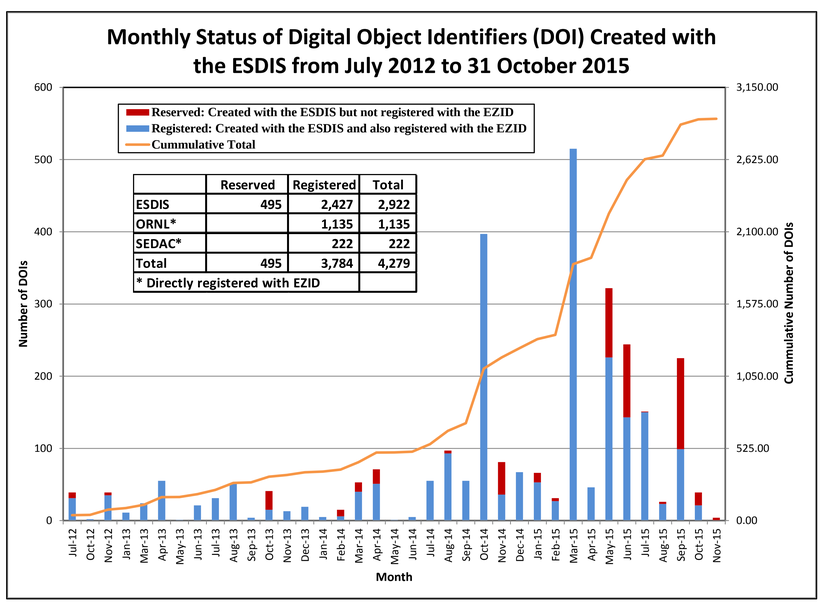
a) ESIP recommendations:
Creator, Year, Product Title, Version, Distributor, Date of Data Access, DOI Name
Example: AIRS Science Team/Joao Texeira; (2013): Aqua AIRS Level 3 Standard Daily Product using AIRS and AMSU with HSB V6; NASA Goddard Earth Sciences Data and Information Services Center. May 2014
http://dx.doi.org/10.5067/AQUA/AIRS/DATA302

b) DataCite Citation tool results:
Creator, Year, Product Title, Distributor, DOI Name
Example: AIRS Science Team/Joao Texeira; (2013): Aqua AIRS Level 3 Standard Daily Product using AIRS and AMSU with HSB V6; NASA Goddard Earth Sciences Data and Information Services Center.
http://dx.doi.org/10.5067/AQUA/AIRS/DATA302

c) American Meteorological Society:
Creator, Year, Product Title, Version, Distributor, Date of Data Access, DOI Name
Example: AIRS Science Team/Joao Texeira; (2013): Aqua AIRS Level 3 Standard Daily Product using AIRS and AMSU with HSB V6; NASA Goddard Earth Sciences Data and Information Services Center. May 2014
http://dx.doi.org/10.5067/AQUA/AIRS/DATA302

d) The Economic and Social Research Council (ESRC) - United Kingdom Data Services:
Identifier, Creator, Title, Publisher, and Publication year, Resource Type, Version
Example: University of Essex. Institute for Social and Economic Research and National Centre for Social Research, Understanding Society: Wave 1, 2009-2010 and Wave 2, Year 1 (Interim Release), 2010 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], February 2012. SN: 6614, http://dx.doi.org/10.5255/UKDA-SN-6614-3.

Version – not currently in ESDIS DOI model but most of the data providers include it in the Title.
Date of Data Access: to be provided by the authors
Resources Type, Version: These are optional



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

Data Citation of Services

To give the Earth science data users the capability of creating data products with user-specified parameters, data centers are developing tools and technologies, such as Geospatial Interactive Online Visualization ANd aNalysis Infrastructure (GIOVANNI), Global Imagery Browse Services (GIBS), and the Land and Atmospheric Near Real-Time Capability for EOS (LANCE). Some of these products are archived only for a very short time period; e.g., LANCE products are archived for two weeks, and GIOVANNI and GIBS products are both created real time without being archived. This provides a challenge in assigning DOIs and developing data citations for the data products that are created through such tools.

For these special service tools, a DOI can be assigned to the service, and the user can copy the parameters selected for the generation of the product. The data citation can include the parameters used in the creation of such data products, which will ensure reproduction of the exactly same data that was used in the research. For example, for Worldview, an interactively global image browse tool that interfaces with GIBS for data, a DOI can be assigned to this service and when a product is created, Worldview generates a URL that can be quoted in the citation with the DOI, making it a complete citation. Example:.

https://worldview.earthdata.nasa.gov/?p=geographic&l=MODIS_Terra_SurfaceReflectance_Bands143,MODIS_Aqua_CorrectedReflectance_TrueColor%28hidden%29,MODIS_Terra_CorrectedReflectance_TrueColor,MODIS_Terra_Brightness_Temp_Band31_Day,MODIS_Terra_Aerosol,Reference_Labels%28hidden%29,Reference_Features%28hidden%29,Coastlines&t=2015-11-20&v=-177.890625,-48.796875,47.109375,55.265625
GIBS (Creator), Earthdata (distributor), 2015(year), and DOI name.

DOI Metadata Recommended for Data Citation

The metadata attributes that are required for the data citation and are common among most of the agencies are.

Identifier
Creator Name
Title
Publisher
Publication Year
URL

These elements are required for the registration of the DOI and such citations can often be generated automatically using DOI attributes. Having these citations available, data users can easily copy and paste them into their publications