

# Frequently Asked Questions about Digital Object Identifiers (DOIs)

*for USGS Researchers*

July 2015

*Document created by the Community for Data Integration, a team representing all Mission Areas in USGS<sup>1</sup>*

[What is a Digital Object Identifier \(DOI\)?](#)

[What does a Digital Object Identifier look like?](#)

[How do I use a Digital Object Identifier to locate an online resource?](#)

[What's the difference between a URL and a Digital Object Identifier?](#)

[What does a Digital Object Identifier let me do?](#)

[Which USGS products need a Digital Object Identifier assigned to them for release?](#)

[How do I designate the particular parts of a dataset that I used in my publication for the Digital Object Identifier?](#)

[How do I cite a dataset that is continuously updated?](#)

[What about versioning of datasets?](#)

[How will a Digital Object Identifier be assigned to my publication?](#)

[How will a Digital Object Identifier be assigned to my data?](#)

[What information is needed to register my publication and/or my data for a Digital Object Identifier?](#)

[My article is going to be published in an external, peer-reviewed journal. Do I need to register for a Digital Object Identifier for my article before I submit it for publication?](#)

[How do I include a Digital Object Identifier in my data citation for inclusion in a publication?](#)

[Can Digital Object Identifiers be deleted?](#)

**Further Information about Digital Object Identifiers:**

- [What are the components of a Digital Object Identifier?](#)
- [Are Digital Object Identifiers extensible?](#)

## **What is a Digital Object Identifier (DOI)?**

A Digital Object Identifier (DOI) is one type of unique, persistent identifier that is permanently assigned to a specific electronic resource. It is sometimes likened to a Social Security Number (SSN) for a person. Just as a person may be assigned a unique number that identifies him or her to various parties – employers, creditors, the Internal Revenue Service – and remains the same no matter where he or she moves, so an electronic object receives a unique sequence of alphanumeric characters that remain tied to that object, no matter how many times the object moves to different servers or property rights owners. A Digital Object Identifier can link an end user to the unique digital object. A Digital Object Identifier supports scientific integrity in that it provides access to the data, workflow, or software version used in a research project from which results can be reproduced.

Two official Digital Object Identifier registration organizations used by the USGS and sanctioned by the Digital Object Identifier Consortium are CrossRef and DataCite. For more information refer to [www.doi.org](http://www.doi.org).

CrossRef was developed within the academic and professional publishing community. Academic, scholarly, and professional publications, including journal articles, books, reports, proceedings, and dissertations, comprise the majority of objects referenced in the CrossRef Digital Object Identifier system.

DataCite emerged from the academic library community with the express goal of providing persistent access to data. While DataCite Digital Object Identifiers can be assigned to other objects, they are used primarily to identify datasets.

CrossRef and DataCite have an agreement to cross-index their registries; this will enable end users to discover publications and datasets through a search of either registry.

### **What does a Digital Object Identifier look like?**

A Digital Object Identifier can consist of numbers, or a combination of letters and numbers. Some real examples of Digital Object Identifiers:

10.5061/DRYAD.9077  
10.3133/sir20155010  
10.5066/F7BR8Q75

### **How do I use a Digital Object Identifier to locate an online resource?**

A Digital Object Identifier looks like this:

**10.3133/pp1814A**

To hyperlink a Digital Object Identifier to resolve to an online location, you append <http://dx.doi.org/> to the front of the Digital Object Identifier:

<http://dx.doi.org/10.3133/pp1814A>

### **What's the difference between a URL and a Digital Object Identifier?**

Uniform resource locators (URLs) are basic Web addresses and like Digital Object Identifiers, are used to connect to resources on the Internet. URLs are often used in documentation and as a component of citations. The challenge with URLs is that they change: sites are reorganized and restructured, domain names change, and intellectual property rights shift to new content owners. Citations using URLs give no indication as to the new location, so the user sometimes cannot find a resource. Even when the object appears to be found under a new URL, there's no guarantee that its content is exactly the same. A URL can't be considered a permanent, fixed reference for an object, because they often change over time.

In contrast, a Digital Object Identifier is permanently attached to a digital object, and is managed apart from the object's physical location. A Digital Object Identifier is registered through a Digital Object Identifier registration agency, which records metadata about that object, including title, publisher, URL, author, and date. When a digital object moves to a different folder, domain, or organizational owner, the party responsible for registering the Digital Object Identifier should update the Digital Object Identifier

metadata in the registry to reference the new URL, as well as any other changes in information (such as a new publisher). The Digital Object Identifier itself doesn't change. This provides enormous advantage, because a *properly managed* Digital Object Identifier will always point to the current online location of that object, and can therefore be used reliably in all references and citations. The reliability of Digital Object Identifiers provides a strong advantage over a cited URL, which does little good for end users if the desired object is no longer available from that electronic location.

### **What does a Digital Object Identifier let me do?**

As USGS scientists, we sometimes need to identify a specific online resource, such as the final version of a data set or publication that was approved for release, a data set that was improved and annotated to meet a requirement, or the version of a numerical model that was used in a scientific publication.

Using a Digital Object Identifier to identify a specific digital object lets us:

- ensure that the product (publication, data, or dataset for example) is one that was approved, in compliance with Fundamental Science Practices;
- formally reference a resource (for example data used in a publication) in a citation;
- make USGS products more discoverable because Digital Object Identifiers are indexed in citation databases along with external products;
- provide the foundational materials that enable our scientific peers to replicate our work and validate our conclusions;
- avoid an extra step in updating references to a resource if it is moved to a new URL;
- efficiently update citations by changing the metadata at the registry only, when products are moved to new URLs, to keep linked references to that product current;
- make the linking to a product more persistent.

### **Which USGS products need a Digital Object Identifier assigned to them for release?**

There are certain products that should have a Digital Object Identifier at the time of release according to new requirements from Office of Science and Technology Policy (OSTP)<sup>2</sup> and Office of Management and Budget (OMB)<sup>3</sup> in 2013. For USGS, these products include:

- USGS Series Publications: Currently one Digital Object Identifier from CrossRef is assigned in the publishing process during dissemination.
- Data released to the public and/or used to produce conclusions for a publication (includes Scientific Data for Release and USGS Data Series Publications)
- Web Services (optional)
  - A Digital Object Identifier applied to a web service allows a persistent view into a data resource.

### **How do I designate the particular parts of a dataset that I used in my publication for the Digital Object Identifier?**

Digital Object Identifiers can be applied at various levels of granularity. For USGS at this time, a best practice is to apply a Digital Object Identifier to the highest level, such as the collection level dataset. However, in some cases it may be important to assign a Digital Object Identifier to a subset of data that we wish to distinguish as defined information or data parcel for a particular audience or customer. That will not always be defined as the collection level, but the Digital Object Identifier will be applied to the

data parcel most useful to achieve our goals in identifying the information or data used in our research. To fully support reproducibility of science, researchers should consider applying a Digital Object Identifier to a specific subset of a dataset used in analysis.

### **How do I cite a dataset that is continuously updated?**

The following are possible ways to cite datasets that are continuously and rapidly updated:

- Cite a specific time slice:  
Moody, J.A., and Meade, R.H., 2013, Powder River: Data for cross-channel profiles at 22 sites in southeastern Montana from 1975 through 2013: U.S. Geological Survey Data Release, <http://dx.doi.org/10.5066/F70Z719C>.
- Cite a specific snap shot:  
Curran, C.A., Magirl, C.S., Duda, J.J., 2013, Suspended-sediment concentration during dam decommissioning in the Elwha River, Washington, September 2011 to February 2013: U.S. Geological Survey Data Release, <http://dx.doi.org/10.5066/F7G73BQ2>.
- Cite a dataset with an access date:  
Zwally, H.J., Schutz, R., Bentley, C., Bufton, J., Herring, T., Minster, J., Spinhirne, J., and Thomas, R., 2003, GLAS/ICESat L1A Global Altimetry Data V018, 15 October to 18 November 2003: National Snow and Ice Data Center, accessed July 21, 2011, at <http://dx.doi.org/10.3334/NSIDC/gla01>.

### **What about versioning of datasets?**

In the future, USGS should consider application of Digital Object Identifier to versions of our datasets and databases. Curators of USGS datasets are asked to consider applying Digital Object Identifiers to versions of datasets as they change at appropriate intervals to their database.

### **How will a Digital Object Identifier be assigned to my publication?**

A Digital Object Identifier for USGS series publications (including Data Series Publications) will be obtained for you when they are released through the USGS approval process specified in the "Fundamental Science Practices" (<http://www.usgs.gov/usgs-manual/500/502-4.html>). These Digital Object Identifiers will originate from the CrossRef system. The DOI prefix for USGS scholarly series publications is 10.3133.

*Example USGS Publications (DOI automatically obtained through CrossRef):*

Engott, J.A., Johnson, A.G., Bassiouni, Maoya, and Izuka, S.K., 2015, Spatially distributed groundwater recharge for 2010 land cover estimated using a water-budget model for the Island of O`ahu, Hawai`i: U.S. Geological Survey Scientific Investigations Report 2015–5010, 49 p., <http://dx.doi.org/10.3133/sir20155010>.

Diffendorfer, J.E., Compton, Roger, Kramer, Louisa, Ancona, Zach, and Norton, Donna, 2014, Onshore industrial wind turbine locations for the United States through July 2013: U.S. Geological Survey Data Series 817, <http://dx.doi.org/10.3133/ds817>.

## How will a Digital Object Identifier be assigned to my data?

Digital Object Identifiers are assigned after the data product is approved and prior to release and dissemination on any official USGS website. The Digital Object Identifier must be included in the metadata record describing the dataset. Within the USGS data approval and release process, assignment of a DOI is the responsibility of a Center's Metadata Coordinator. Programs and Offices will have the responsibility for determining how this will be accomplished in their Center. However it is also possible for USGS scientists to obtain a DOI for citing their approved and released data using the USGS Digital Object Identifier Creation Tool which accesses DataCite.org. The DOI prefix for USGS scientific data release is 10.5066. Hence when a DOI has a 10.3133 it is a USGS scholarly series publication that is referenced and when a DOI has a 10.5066 it is USGS data that is referenced. Ideally a USGS series publication should include a DOI reference to the data that supported it.

*Data associated with the scholarly publication (DOI obtained by data owner through DataCite):*

Engott, J.A., 2015, Mean annual water-budget components for the Island of Oahu, Hawaii, for average climate conditions, 1978-2007 rainfall and 2010 land cover: U.S. Geological Survey Data Release, <http://dx.doi.org/10.5066/F7XP72ZX>.

Catchings, R.D. Strayer, L.M. Goldman, M.R. Criley, C.J. Garcia, S.H. Sickler, R.R. Catchings, M.K. Chan, J.H. Gordon, L. Haefner, S. Blair, L. Gandhok, G. and Johnson, M., 2015, 2013 East Bay Seismic Experiment (EBSE)--implosion data, Hayward, Calif.: U.S. Geological Survey Data Release, <http://dx.doi.org/10.5066/F7BR8Q75>.

Information about the process can be found here:

<http://www.usgs.gov/datamanagement/preserve/persistentIDs.php>

For more information about how your Center's coordinator can obtain a DOI for your data, see:

[http://www.usgs.gov/core\\_science\\_systems/csas/](http://www.usgs.gov/core_science_systems/csas/)

## What information is needed to register my publication and/or my data for a Digital Object Identifier?

As stated above, a Digital Object Identifier will be assigned to your USGS series publication or your data in the approval and dissemination process. However, in order for a Digital Object Identifier to be obtained from a registration system, some information about your product will need to be provided. Since your Digital Object Identifier will be indexed into a registry, and each Digital Object Identifier registry is searchable, the metadata about your information product allows people to discover it through a registry search. USGS uses information from IPDs and other systems to populate Digital Object Identifier registries. When you go through your approval and release process through IPDS, it is very important that you provide as many details as possible so that the product is properly described in the Digital Object Identifier registry.

**My article is going to be published in an external, peer-reviewed journal. Do I need to register for a Digital Object Identifier for my article before I submit it for publication?**

No. Any items that will be published by an external agent, such as a scholarly journal or a publisher, will be assigned a Digital Object Identifier from their publisher. Do not register these objects yourself in a Digital Object Identifier registry. USGS will register Digital Object Identifiers only for those items published by the Bureau itself. The DOI prefix will be entirely different depending on the scholarly journal or publisher.

*Example outside journal publication (No action needed to create a DOI):*

Estimates of Natural Salinity and Hydrology in a Subtropical Estuarine Ecosystem: Implications for Greater Everglades Restoration Frank E. Marshall, G. Lynn Wingard, Patrick A. Pitts.  
<http://dx.doi.org/10.1007/s12237-014-9783-8>

**How do I include a Digital Object Identifier in my data citation for inclusion in a publication?**

Data citations follow a general USGS citation format, with the USGS Data Release type and the Digital Object Identifier at the end of the citation.

For example:

Shah, A., 2014, Airborne Geophysical Surveys over the 2011 Mineral, Virginia, Earthquake Area: U.S. Geological Survey Data Release, <http://dx.doi.org/10.5066/F78K773V>.

**Can Digital Object Identifiers be deleted?**

No. Once registered, Digital Object Identifiers cannot be removed from the registry or reassigned to a different object. For this reason, Digital Object Identifiers are to be assigned close to the publication release, when the data are at or near their final form, and a location for the product is identified.

## Further Information about Digital Object Identifiers:

### What are the components of a Digital Object Identifier?

10.5061/DRYAD.9077  
prefix                      suffix

Prefix: All Digital Object Identifier handles start with “10.” The four digits that follow are unique to a consortium member. A member may have just one four-digit code (e.g. an organization) or multiple four-digit codes (for example, a publisher may have a unique four-digit code for each journal it publishes). Four digit prefixes are assigned to content publishers by one of the designated Digital Object Identifier Registration Agencies (e.g. CrossRef, DataCite). The USGS four digit code for data is 5066 and the four digit code for USGS series publications is 3133.

The prefix is always separated from the suffix by a “/” .

Suffix: The suffix is a unique alpha-numeric string assigned by the publisher of the content. It is case-insensitive, and can consist of one or more “nodes.” In the example above, “DRYAD” is a node, and “9077” is also a node; if more than one node is used, successive nodes must be separated by delimiters such as “.” “:” or “|” (periods, colons, or pipes). Nodes can consist of bibliographic metadata (e.g. journal name, date, issue, page range), ISBNs, project codes, or any other scheme that is relevant to the content publisher. The key point is that the suffix must be unique.

### Are Digital Object Identifiers extensible?

Yes, Digital Object Identifiers can be made extensible through the suffix nodes. Extensibility allows publishers to build relationships among different components or expressions of an object. Hypothetical examples include:

Dataset:

/USGS.7539

Web feature service for this dataset:

/USGS.7539.wfs

Metadata record for this dataset:

/USGS.7539.meta

---

<sup>1</sup> Viv Hutchison, Michelle Chang, John Faundeen, Jean Freney, Dorrie Gellenbeck, Greg Gunther, Keith Kirk, James Kreft, Fran Lightsom, JC Nelson, Helen Tong, Roland Viger, Lisa Zolly

<sup>2</sup> [http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf)

<sup>3</sup> <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>