

Assigning Persistent Identifiers in Action

The next case study by Susan M. Braxton, Bethany Anderson, Margaret H. Burnette, Thomas G. Habing, William H. Mischo, Sarah L. Shreeves, Sarah C. Williams, and Heidi J. Imker describes how the University of Illinois is using a DOI-minting service in a novel way.

A Participant Agreement for Minting DOIs for Data Not in a Repository

*Susan M. Braxton, Bethany Anderson, Margaret H. Burnette, Thomas G. Habing, William H. Mischo, Sarah L. Shreeves, Sarah C. Williams, and Heidi J. Imker**

In 2014, the eResearch Implementation Committee[†] at the University Library, University of Illinois at Urbana-Champaign carried out a pilot study using EZID via subscription from Purdue University (<http://ezid.lib.purdue.edu>) to mint DataCite (<https://www.datacite.org>) DOIs for data upon request. The goals of the pilot were to explore researcher interest in DOIs for data on our campus and to explore the delivery of such a service by the University Library using EZID. The library-mediated service was announced via the all-faculty and staff weekly e-mail newsletter (figure 6.5) and the corresponding graduate student newsletter. The full announcement on our website specified the types of eligible resources and terms of participation (figure 6.5).

Ten individuals representing three colleges and two research institutes applied to participate between February and October 2014. Although the response was underwhelming, the unique challenge of providing DOIs for “external” resources not archived in a library repository raised questions that helped guide the development of our pilot service to production.

* This study is licensed under a Creative Commons Attribution 4.0 License, CC BY (<https://creativecommons.org/licenses/by/4.0/>).

† Now the Research Data Service, or RDS Committee.



February 16, 2014

[Use EZID to Obtain DOIs for Your Data](#)

Researchers at the University of Illinois at Urbana-Champaign will soon have an opportunity to obtain Digital Object Identifiers (DOIs) for their data as part of a pilot service focused on data sharing and management. The EZID pilot is spearheaded by the University Library and will run during spring and summer 2014. It is open to faculty and staff members and students.

researchdata@library.illinois.edu • University Library

Using EZID to obtain DOIs for your data – a pilot research data service

Who may participate?

Individual students, faculty and staff as well as research groups/programs at the University of Illinois at Urbana-Champaign that have created data resources are eligible to participate in the pilot.

What types of resources are eligible for inclusion in the pilot?

Data resources such as simple data sets (e.g., spreadsheets, CSV files), visualizations, software/code, or collections are eligible. We will exclude resources which already have DOIs, and resources that are not data resources (e.g., article preprints or reports). As a requirement of the pilot, the resource must be accessible via the World Wide Web, and the creator must supply basic metadata about the resource to include in the DOI registry. Where appropriate, participants will be strongly encouraged to deposit their data resources into our [institutional repository, IDEALS](#), to ensure their long-term preservation. If IDEALS deposit is not an option (e.g., for data sets that are actively growing), participants will be responsible for keeping the DOI registry up to date in the event the resource is moved to a new URL.

FIGURE 6.5

Campus e-mail announcement and excerpt from the full description of the pilot DOI for data service.

We had, perhaps naively, expected requests to mint DOIs primarily for repository-ready data from completed projects. Applicants presented instead a range of resources, including several continuously updating, interactive, online databases (e.g., ongoing climate monitoring). These resources were shared by their creators via live, interactive web interfaces that would not be replicated in our institutional repository. Static files (“snapshots”) deposited in the repository would be less functional for users than data in the native interface. We recognized the need to make these resources discoverable and citable, which is at the heart of DataCite’s purpose.³¹ It can be argued that it is precisely these types of resources that are most in need of identifiers—after all, data sets archived in

repositories typically have persistent URLs and descriptive metadata already. However, we were also acutely aware of the social contract that the University Library was making to ensure the persistence of the resources, and we were reluctant to promise long-term access for resources over which we lacked curatorial control. We ultimately decided that the benefits of registering these resources with DataCite (i.e., improved discoverability through DataCite’s broad dissemination of the metadata, and facilitation of citation) outweighed the risk that the DOI may fail to resolve to the resource in the future. Also, we speculated that researchers who made the effort to request DOIs and help craft descriptive metadata for the DataCite registry would be inclined to notify us if the resource moved, or if they later needed help finding a repository to preserve and create access to the resource (e.g., if the program was defunded, and they were no longer able to maintain the active instance of the resource). Finally, the ability within EZID to change the resource status to “unavailable” and redirect the DOI to a “tombstone” page displaying the metadata and reason for unavailability gave us a way to manage the DOI if the resource subsequently disappeared.³²

In an attempt to curb dead links and subsequent accumulation of tombstones in DataCite, we developed a Participant Agreement to outline the responsibilities of the University Library and of the data creator in the creation and management of the DOI.³³ The agreement has five components: the agreement purpose, description of agreement use, definitions, general expectations, and responsibilities. Special attention was given to Creator (or designee) responsibility for resources that would live outside a trusted repository; this section is excerpted below.

Participant Agreement Creator (or designee) responsibilities:

- As the creator (or designee) I certify that the descriptive metadata which I have provided to the University Library are accurate.
- As the creator (or designee), I certify that I am authorized to allow the University Library to share this descriptive metadata via the DataCite registry.
- If the resource does not reside in a trusted digital repository, then:
 - I as the creator (or designee) take full responsibility for maintaining stable access to the resource.
 - If the location of the resource changes (including any and all URL changes), I as the creator or designee agree to notify the University Library of the new location so that the DataCite metadata record can be updated.
 - If contact information changes, I as the creator (or designee) agree to provide the University Library with the updated information.

- If I, as the creator or designee, am no longer able to provide access to the resource, I agree to inform the University Library and to work with the University Library to develop a plan for long-term preservation of the resource as appropriate.
- I understand that if the University Library becomes aware that access to the data has been disrupted and efforts to address the issue with the creator or designee are not resolved within a timely manner, the University Library reserves the right to edit the DataCite record to indicate unavailability (e.g. generate a “tombstone” record).

During our pilot, the agreement was integrated as a modal popup window into our DOI minting web interface* at the metadata review and approval stage, and referenced in the e-mail notice of the newly minted DOI to the requestor. The integration allowed us to retain a record of the agreement as well as an acknowledgement by the creator or designee that the metadata registered with DataCite accurately described the resource and could be made public. Although there are no repercussions for failure to comply, the agreement serves as an opportunity to emphasize the importance of DOI persistence to the requestor. Thus far we have minted thirty DataCite DOIs and 40 percent of them resolve at the request of the data creators to URLs outside of library-run repositories. We have already been notified to update DOI link resolving locations for resources in one external location when the host site URL changed. It remains to be seen how cumbersome or problematic maintaining DOIs for resources outside repositories will be. However, at this early stage in data sharing and citation, it has been a worthwhile experiment.

ACKNOWLEDGMENTS

Thanks to Harriett Green, Karen Hogenboom, Laila Hussein, Beth Sandore Namachchivaya, Kyle Rimkus, Mary Schlembach, Ayla Stein, and Christie Wiley for input on development of the DOI Service Participant Agreement as well as work on the pilot.

* A password-protected web interface to our EZID account with Purdue was developed locally to facilitate entry and management of metadata for our DOIs beyond what was possible with the standard EZID interface.