DOI: 10.18727/0722-6691/5097

The ESO Digital Object Identifier Service

Dominic Bordelon[1]
Uta Grothkopf[1]
Silvia Meakins[1]

[1] ESO

Digital Object Identifiers (DOIs) are persistent identifiers in a global registry that assist in the citation, identification, and discoverability of research information. The ESO library has created a service to provide DOIs to departments within ESO. DOIs have been in use for articles in The Messenger since March 2017, and plans are underway to begin creating DOIs for datasets in the ESO Science Archive.

Digital Object Identifiers

Scholars have long cited the literature they used when defining a piece of work or study, both as a way of giving due credit, and as a way of bolstering the credibility of the study itself. In the hypertext era, such citations can be enhanced with direct links to the cited resources — including both literature and data. Data citation is especially important for the reproducibility of science. However, nearly three decades since the invention of the World Wide Web, we have seen that not all web addresses (or Uniform Resource Locators — URLs) are permanent. Digital Object Identifiers (DOIs) have entered the scene to address this problem.

DOIs are persistent, globally unique identifiers for publications, datasets, and other research products. They are used to refer to information resources in an unambiguous way, for example when citing a paper. They can also be used as URLs, preceded by "https://doi.org/", to conveniently retrieve the resource in question. Typically, this URL brings the user to a "landing page" that describes the resource, and provides a link via which it can be downloaded.

DOIs were conceived in the 1990s by associations of publishers, and the first DOIs were created in 2000[1]. At the time of writing, approximately 148 million DOIs had been created, and there are over five billion DOI lookups per year[2].

The global DOI registry is administered by the non-profit International DOI Foundation. DOIs are specified in a syntax defined by the standard ISO 26324:2012 format. They have become standard components of scholarly publishing, as major journals create a DOI for every article they publish.

A DOI consists of an identifier string (aka DOI name), and some information (metadata) about the resource to which it is assigned. An example of a DOI name is: "10.18727/docs/2". Every DOI name begins with "10.", followed by a number representing the data centre that registered the DOI, a forward slash, and then an arbitrary string ("docs/2", in the example shown in Figure 1). In this way, there is no risk of two data centres trying to register the same DOI name, and each data centre also has control over how it names its DOIs. The particular DOI given in Figure 1 resolves to the landing page for the 2017 ESO Annual Report[3].

The benefits of the DOI system include:
– unambiguous citation — traditional citations can contain confusing abbreviations and are sometimes incomplete;
– convenient and machine-readable retrieval — every DOI is easily turned into a URL, and the DOI name format is easily detected by computers;
– reducing (ideally entirely eliminating) the number of links that "break" after publication.

Publishers and data centres that register DOIs are making a pledge to maintain their records in the DOI registry, so that a requested DOI always resolves, even if the resource itself moves to a different URL. Furthermore, since DOIs give all stakeholders a simple and standard way to refer to the same resource, the DOI system enables the creation of a network of scholarly services and tools.

The ESO DOI service

The ESO Library began investigating the adoption of DOIs in 2015. The two use cases under consideration at the time were publications, such as The Messenger, and datasets in the Science Archive. Although DOIs have typically been used for text publications, they can also be applied to datasets, which should enable easier and higher-quality data citation and tracking. This use is not yet widespread, but data centres in various disciplines are increasingly interested in deploying DOIs for datasets. Furthermore, there is a trend in public funding agencies (for example, the European Commission and the German Research Foundation [Deutsche Forschungsgemeinschaft, DFG]) demanding that data be Open Access; it follows that they should also be citeable.

ESO negotiated an agreement with the Technische Informationsbibliothek (TIB) in Hanover, a national library that provides DOI registry access to academic institutions in Germany, in association with DataCite, one of the foremost DOI registration agencies. Once this agreement was in place, the ESO Library developed the software needed to manage interactions with DataCite. DataCite offers an HTTP[a] application programming interface (API) for registering DOIs.

The resulting software, called the ESO DOI Service, acts as a hub for ESO departments (or "clients") wishing to create DOIs (Figure 1). Several steps are required when a client wants to create a DOI:
1) The client sends a request to the DOI Service with the necessary resource metadata.
2) The DOI Service stores a record and attempts to register it with DataCite.
3) The DOI Service receives DataCite's response and notifies the client.
4) The DOI Service also renders a landing page for every DOI that it registers, so that when someone resolves the DOI, they arrive at a page provided by the DOI Service. Clients can also supply their own landing pages if desired. Figure 2 depicts the architecture of the ESO DOI Service.

In developing the ESO DOI Service, the Library had to take several factors into account. Since more than one type of information resource was involved,

Figure 1. The composition of a DOI name.

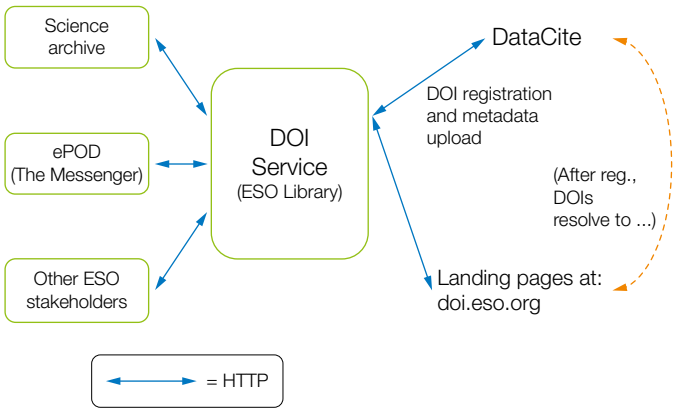

Figure 2. Architecture of the ESO DOI service, showing how the DOI service is used by the main stakeholders. Currently, the main users of the DOI service are the science archive and the education and Public Outreach Department (ePOD).

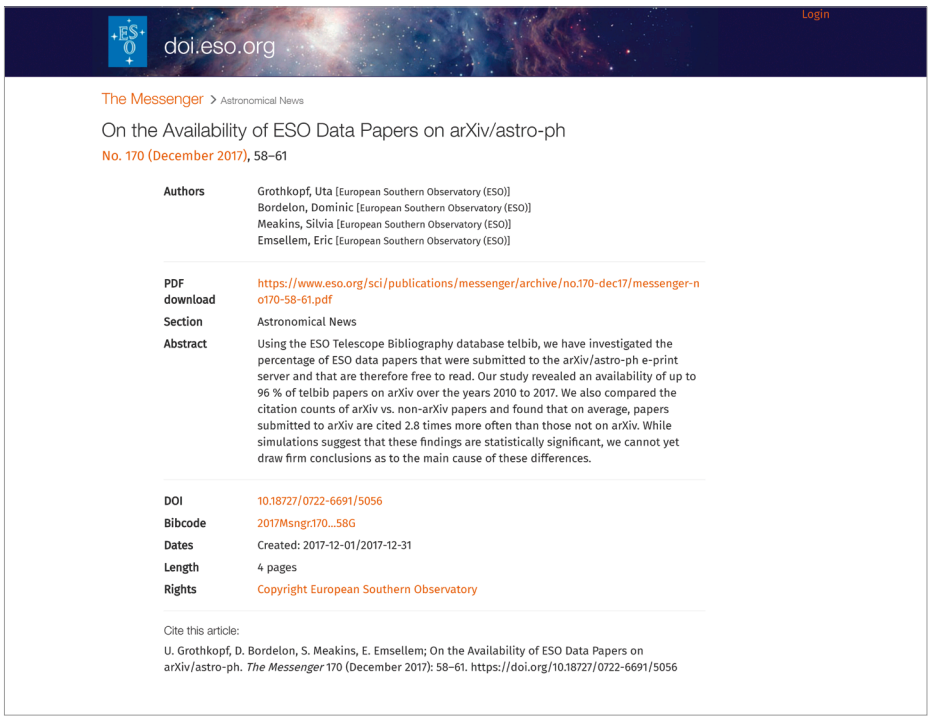

Figure 3. A screenshot of the landing page for an article in The Messenger.

a common data model which would accommodate multiple types of resources was needed. Since more than one department — potentially using different programming languages or database systems — would be served by the software, the solution needed to provide an "agnostic" and universal API to clients via HTTP. The software also needed to handle failure states in case DataCite is unavailable.

With these requirements in mind, development began in the autumn of 2016, and the service was launched in March 2017. The first DOIs minted were for articles of volume 167 of The Messenger (see Figure 3). In the meantime, ESO has created 67 DOIs, and feature developments for the ESO DOI Service have continued. The most important addition is a monitor that checks each record weekly to ensure that the link still works, greatly helping to keep the promise of persistence that is made when participating in the global DOI registry.

Future developments

Plans are under way for the Science Archive to use DOIs for raw datasets, data products, and ad hoc collections that have been assembled by Archive users. The eventual goal is for authors to cite DOIs of the datasets that they used in writing a paper, which would improve scientific reproducibility and greatly assist ESO in tracking the use of its data. DOIs are an excellent tool for continuing and extending the scholarly practice of citation in a way that is accessible, persistent, and situated in a network of resources. The rising trend of data citation not only complies with funding agency demands, but also facilitates further scientific analysis. With the ESO DOI Service, ESO is ready to fully participate in future developments.

Acknowledgements

We would like to thank ESO's Legal Services for their work on the agreement with the Technische Informationsbibliothek.

References

ISO 26324:2012: Information and documentation — Digital object identifier system. International Organization for Standardization, Geneva, Switzerland.

Links

[1] The DOI Handbook: https://doi.org/10.1000/182
[2] Key Facts on the Digital Object Identifier System: http://www.doi.org/factsheets/DOIKeyFacts.html
[3] Landing page for the 2017 ESO Annual Report: https://www.eso.org/public/products/annualreports/ar_2017/

Notes

[a] Hypertext Transfer Protocol (HTTP) is the protocol used for the World Wide Web.