1   Referencing GLIS DOIs in publications and datasets

Version 1.2 - 28/07/2023

Introduction

This document describes the connection between GLIS DOIs and publications and datasets (that also have a DOI assigned), explaining how this is done and what stakeholders need to do to obtain this important function.

Data citation

Connecting GLIS DOIs to DOIs assigned to publications and/or datasets is an example of *data citation*, i.e. the practice of providing a reference to data in the same way as researchers routinely provide a bibliographic reference to outputs such as journal articles, reports and conference papers. Citing data is now recognised as one of the key practices leading to recognition of data as a primary research output.

In publications and datasets about PGRFA, data citation mostly coincides with citing the materials that have been studied. Accurate citation is best done by citing the DOIs assigned by GLIS to the material.

Accurate data citation also gives proper credit to the conservation, development and research work done on the PGRFA.

Introducing EventData

GLIS DOIs are the most appropriate way of citing the PGRFA subject of the study resulting in a publication or a dataset. But how can GLIS find those publications and datasets?

Recognising the importance of data citation, Crossref and DataCite[1] jointly developed EventData, a repository of connections among DOIs established by publications or datasets.

How does EventData work?

Figure 1 shows how EventData works. When a researcher prepares a publication, she sends the PDF of the article to the Publisher (usually through an Editor). The Publisher's system processes the PDF and any associated material, including tables, diagrams and bibliographic references, and sends an XML document to Crossref to obtain the DOI that will be assigned to the publication. The XML document contains information such as the title, date of publication, authors and the list of bibliographic references. Note that many publications are now cited through their own DOI. So, what Crossref does is to enter a record in the EventData repository essentially stating that the

---

[1] Crossref (https://www.crossref.org) and DataCite (https://www.datacite.org) are the two leading DOI Registration Agencies. A DOI Registration Agency is an organization that assigns DOIs on behalf of the International DOI Foundation (https://www.doi.org), the body coordinating and regulating the DOI System. Crossref focuses on assigning DOIs to publications while DataCite focuses on datasets. Please note that EventData is currently limited to these two Registration Agencies (RA), i.e. DOIs assigned by other RAs will not be collected by EventData.

publication's DOI "references" all the DOIs in the list of bibliographic references attached to the article.

Likewise, when a researcher performs a study, e.g. on Characterization and Evaluation of a set of PGRFA, the resulting dataset is deposited in a repository and a DOI is assigned to it. In the process it is possible, actually recommended, to list the DOIs of the PGRFA to which the dataset refers. In this case, DataCite enters a record in the EventData repository associating the dataset's DOI to the DOIs of the related PGRFA.

The EventData repository is queried periodically by GLIS to obtain the list of publications and datasets referencing one or more GLIS DOIs.
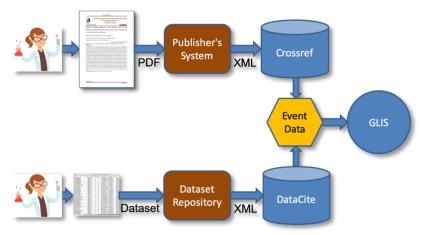

Figure 1: The EventData workflow

Preparing publications for EventData

In practice, what do authors need to do to make sure their publications are properly formatted for EventData so that the connection between GLIS DOIs and the publication's DOI is established? The current state of the art of data citation through EventData is described in a blog post by Crossref and DataCite[2] presenting available options; essentially, there are two ways:

- using bibliographic references
- using specific relation-setting sections of the Crossref metadata

At this time, the recommended option is to cite GLIS DOIs among other bibliographic references. The reason is simple: publisher's systems are designed to extract the bibliographic references and submit them to Crossref in the correct way to establish the citation relationship. The second option is currently offered by a limited number of publisher systems. An effort is under way to update publisher systems to accept a list of DOIs submitted, for example, as a separate document, and forward the proper XML to Crossref.

Editors and publishers ae familiarizing with the new technology

When researchers prepare the PDF, they list the (potentially many) GLIS DOIs among the bibliographic references; an example is provided in Figure 2 below.

---

[2] See https://doi.org/10.5438/9456-sb49

Figure 2: GLIS DOIs among other bibliographic references

Please note that you can enter any other text within the GLIS DOI citation (in the example above, the IRRI Accession Number was also listed); the publisher's system will look for the "https://doi.org" string to identify a DOI.

Sometimes, when the Editor receives the PDF from the Author, he/she is surprised by the "unusual look" of the DOI citations because the traditional elements of the citation (author, title, publisher, publication year and so on) are missing.

Editors need to familiarize with data citation and need to understand that listing GLIS DOIs among other bibliographic references, while odd-looking, is a necessity currently imposed by publisher systems; authors would be quite happy to submit the list of GLIS DOIs as a separate file and not inflate the list of bibliographic references, if only the publisher's system was able to accept it!

Crossref and DataCite are going great lengths talking to publishers about data citation and the Secretariat is supporting authors trying to get their papers through recalcitrant Editors.

If you experience difficulties, please forward the following text in an email to the Editor and copy PGRFA-Treaty@fao.org so that we can assist in the resolution of the issue.

> The DOIs I am including in the list of bibliographic references are associated to the Plant Genetic Resources (PGRFA) subject of the research presented in the paper. The system assigning such DOIs is the Global Information System (GLIS, https://glis.fao.org/glis) of the International Treaty on Plant Genetic Resources for Food and Agriculture (http://www.fao.org/plant-treaty).
> The inclusion of the PGRFA DOIs in the list of bibliographic references is recommended by Crossref and DataCite (see https://doi.org/10.5438/9456-sb49) as a way of establishing a relationship between the publication's DOI and the PGRFA DOIs that GLIS can use to make the publication more easily findable by its users.
> I am copying the Treaty Secretariat that will be happy to provide any additional clarification you or your Publisher may require on this issue.

This is a transitional phase where Editors, Publishers as well as Authors need to become familiar with data citation and that will hopefully lead to publisher systems being updated to accept GLIS DOIs as a separate list without cluttering the list of bibliographic references.

Depositing a dataset

For datasets, the process is usually less problematic because all modern public repositories (e.g. Figshare) allow users to list the DOIs referenced by the dataset.

When the dataset is deposited to a custom-built repository, instead, it is necessary to verify that the proper XML is forwarded to DataCite or Crossref when acquiring a new DOI to be assigned to the dataset.

When an institution makes datasets available to users through its own website, rather than using a repository, in order for Event Data to capture the citation of GLIS DOIs, it is necessary that a DOI is assigned to the dataset. This may be an issue if the institution is not routinely assigning DOIs to datasets. Again, the Secretariat would be happy to provide advice on how to address this potential problem.

What it looks like

When the EventData query returns publications and/or datasets associated to a DOI, the GLIS page shows at the bottom the list of publications and datasets citing the current PGRFA, see Figure 3 below.
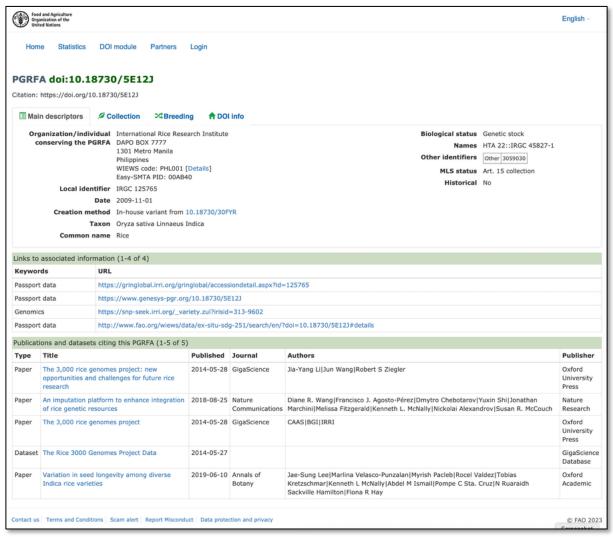

Figure 3: Publications and datasets citing the GLIS DOI listed at the bottom of the page

Please note that the publications and datasets listed in the GLIS DOI page are accessible under the conditions set by the publisher or repository: a subscription or a payment may be required to access the full publication or dataset.

I have a publication that I cannot get through, what do I do?

If your Editor or your Publisher refuse to publish your paper unless you remove the GLIS DOIs from the bibliographic references, we can offer a workaround:

- publish the paper without the GLIS DOIs in the bibliographic references, then
- send us the DOI assigned to the publication and the list of GLIS DOIs it should be connected to

We will enter the relation in GLIS manually. Please note that, in the future, you may be able to update the publication's DOI and add the GLIS DOIs in a way that makes the relationship appear in EventData. GLIS will then replace the manually entered record with this "official" one.

Of course, this is a stopgap measure that cannot be sustained in the long run. However, we can still use it for important papers that have been published in the past.

I still need support!

If you need any clarification, please do not hesitate to send an email to PGRFA-Treaty@fao.org and we will be delighted to assist!