

# Data Management Plan for the Study 'Do Reputable Open Access Journals Require Open Data Sharing'?

Gabriela Tenorio, BUAP

jueves, agosto 09, 2018

```
# Add the dataset of DOAJ Seal journals
```

```
doaj_seal <- read_csv("data/DOAJ_Seal.csv")
```

## Administrative Data

### ID

*Not Applicable*

### Funder

This research is being submitted for funding to the J. Bohannon Foundation <http://www.johnbohannon.org/>.##

### Grant Reference Number

\*Not Applicable (Proposal in preparation)####

### Project Name

Do Reputable Open Access Journals Require Open Data Sharing?

### Project Description

This study analyzes the submission requirements of the most reputable open access journals to determine the prevalence and

## Data Collection

### Existing Data Being Reused

This study relies on the DOAJ Journal Metadata available as a csv file download from the DOAJ website

<https://doaj.org/faq#metadata>. The csv file is updated every 30 minutes.

This data will be read into RStudio, using the tidyverse package (need to cite Tidyverse somehow) to filter it for those journals awarded the DOAJ Seal, and to remove unneeded columns containing the web addresses for journal policies around plagiarism, submission fees, and other urls not related to this study.

The filtered version of the data set will be exported as a new file named `doaj_seal.csv` for importing into analytical software.

A sample of the `doaj_seal.csv` data set is shown below. The complete data set is available in searchable and browsable format in the Annex near the end of this document.

```
knitr::kable(head(doaj_seal, 4), caption = 'A Table of the
```

## Documentation and Metadata

The journal metadata contained in the project's data sets comes directly from the Directory of Open Access Journals.

The final outputs from the project will be documented in metadata files according to the DataCite DOI registration agency – see the DataCite Metadata Schema 4.1 <https://schema.datacite.org/> for specific details. By following this standard metadata format, other researchers (and computers) will be able to find, access, and reuse the outputs from this project by searching the DataCite metadatabase.

## Ethics and Legal Compliance

No additional ethical or privacy issues arise in this study because both the DOAJ data, and the information about data policies for any published journal, are publicly posted online.

The data provided about journals awarded the Directory of Open Access Journals Seal of Approval is distributed under a CC BY-SA license. This license requires that reusers of the data share their derivative data set under the same license. Therefore, the output of this research will be disseminated under the CC BY-SA license. This license adheres to the *Principles and Guidelines of the Research Data Alliance Legal Interoperability Group*, which recommends the use of Creative Commons Attribution licenses to allow the broadest sharing of data while guaranteeing attribution to the data provider. (FN: <https://www.rd-alliance.org/rda-codata-legal-interoperability-research-data-principles-and-implementation>)

## Storage and Backup

During the active phase of the project data will be stored on and backed up to the Research Data Storage Facility (RDSF) at My Institution. This facility represents 2 million pounds of digital resilient storage, with ongoing capital investment. The RDSF is overseen by a steering group of senior research and support staff, which includes the PVC Research. Backup procedures are robust (overnight backup, copies held remotely on tape) and secured access is in place

Selection and long-term preservation

$||>$  To be completed by the participants!  $<||$



## Data Sharing

The data and metadata will be made openly accessible under a CC-BY SA license in the CERN-maintained open access repository Zenodo <https://zenodo.org/>, an open dependable home for the long-tail of science, enabling researchers to share and preserve any research outputs in any size, any format and from any science. (FN: Zenodo policies are available online at <http://about.zenodo.org/policies/>)

## Responsibilities and Resources

The Principal Investigator is responsible for implementing the Data Management Plan and ensuring it is reviewed and revised as necessary. (S)he will be responsible for all data collection and recording; for data analysis and visualization; and for maintain all files under version control using git and GitHub.

The Data Management Specialist assigned to the project as an in-kind contribution from the My Institution Library will be responsible for creating the DataCite metadata documentation for all outputs and ensuring timely DOI registration of each final output. (S)he will also deposit all final outputs to the Zenodo repository and update metadata associated with the DOI as necessary.

## Annexes

### Complete dataset doaj\_seal.csv

*This data table was compiled using the DT package by Xie (2018)*

`||> Turn this into a code chunk`

```
{r data-table} doaj_seal %>%   datatable(rownames =  
FALSE,                        colnames = c("Title",  
"Publisher", "Country", "Fees", "Waivers",  
"Identifiers", "Start Year", "Language(s)", "Review  
Process", "Plagiarism check", "Time to Press",  
"License", "Author owns Copyright", "DOAJ Seal"),  
class = "cell-border stripe",      caption =  
"Journals with DOAJ Seal",          filter =  
list(position = "bottom"),          extensions =  
'Buttons', options = list(dom = 'Bfrrtip',  
buttons = c('colvis', 'csv', 'pdf'))  )  
<||
```

### Principal Investigator's BioSketch

This is auto-populated from your ORCID profile using the rorcid  
package. (cite the package)

# References

Castro et al, 2017, blahblahblah

Rmarkdown package citation of some kind, blahblahblah

Vasilevsky et al, 2017, blahblahblah

Tidyverse package citation of some kind, blahblahblah

Rorcid package citation of some kind, blahblahblah

Xie (2018), DT Table package, CRAN, blahblahblah

Xie et al (2018) Markdown book published by CRC