

# Grey research data

*Les Kneebone 1*

**Les Kneebone, Michelle Zwagerman**

<sup>1</sup>Analysis & Policy Observatory, Melbourne, Australia, [admin@apo.org.au](mailto:admin@apo.org.au)

## GENERAL INFORMATION

To help conference organisers with scheduling, please answer the following questions:

- Is this workshop half-day or full-day? Half-day workshop
- Does the workshop include a hands-on component? Compiling lists on butchers' paper
- Are there any constraints on the number of attendees? 25
- Are there any special seating or table requirements (e.g. for breakouts, teams)? Island-style for teams of 5-6
- Are there any technical requirements beyond AV and access to wireless network? Flip charts, butchers paper, markers

## DESCRIPTION

Research data management has matured over the last 20 years. Evolving standards, platforms and practices have resulted in the identification, description, publication and preservation of datasets as part of the research lifecycle. As a result, data citation practices in scholarly publications have been enabled by uniquely identified and persistently stored datasets. The practice of minting persistent IDs for publications, datasets, researchers and research concepts enables exploitation of complex and powerful graphs of research activity and outputs.

However, progress into making datasets find-able, accessible, interoperable and re-usable (FAIR) has not been uniform, with much of the progress made in traditional research institutions, and less so in public, commercial and not-for-profit sectors. Furthermore, familiar identifier systems such as ORCID, GRID and DOI have had lesser uptake in research practice outside of universities.

APO provides a repository solution for research outputs from all sectors. APO address the failure of the market to manage grey literature, or direct-publishing literature, with the same FAIR principles. However, the best practice management of research objects seen in universities is not being matched in the area of grey research data. Research datasets produced outside of the academy are seldom available as discrete research objects. When grey datasets are available, the citation practices in grey literature is applied sporadically, inconsistently or not at all.

In this workshop, we give an overview of the key features of APO website and demonstrate how anyone can add a resource, including adding metadata. We then recall challenging scenarios where we have not been able to access datasets or easily reference datasets that our policy literature is based on. We then open the discussion to participant to

talk about their own struggles with datasets, whether it be publishing them through local portals, institutional repositories or cross-institutional platforms; minting Digital Object Identifiers; finding and accessing datasets; citing datasets in reports; versioning and provenance. The workshop discussion will produce a table of stories and requirements centering on grey dataset management that APO will use in future research proposals and papers.

Outline of the workshop content:

1. Introduction and Presentation. APO will give an overview  
**45 minutes**
2. Floor discussion  
**30 minutes**
3. Activity – small groups draft requirements  
**30 minutes**
4. Synthesis – requirements collection and tabling  
**30 minutes**
4. Summing up  
**15 minutes**

#### **WHO SHOULD ATTEND**

Researchers; Research data curators; Data librarians; Publishers

#### **WHAT TO BRING**

BYOD optional – no special equipment or software required.