

Monash University's Content Migration: A case study

Key Points

- **Migrating content using the API**
How to organize the collections so they can be managed in a useful way by the API
- **Monash University Theses**
A collection that consists of about 6,200 manuscripts
- **Impact**
The theses have been viewed over half a million times and downloaded over 45,000 times
- **Next steps**
They'll encourage postgraduate candidates to link their research with their thesis

Introduction

Monash University had been running a Fedora based institutional repository (IR) since 2004. Since then, they've published 190,000 records, of which 30,000 are open access.

The IR included publication records from the Research Office, open access working papers and journals from faculty, theses, datasets, and special collections of primary and historical research material.

In 2016, they decided to split the institutional repository into three strands: Pure for publications, Omeka for special library collections, and Figshare for data, theses, working papers, and other grey literature. They were particularly interested in using Figshare because of:

- the self-submitting workflow that deviated from the traditional library-driven submission systems.
- improved Google search exposure compared to their previous IR.
- the ability to mint a Digital Object Identifier (DOI).
- the Application Programming Interface (API).
- the review workflow, useful in particular for the theses collection.

They started the migration into Figshare in late 2016 by manually uploading working papers and image collections into their repository. In early 2017, they began the process of uploading the theses series collection into Figshare using the API. In August 2017, they completed the migration of working papers and archived journals into Figshare using the API.

Migrating content using the API

One of the first steps was deciding how to organize the collections so they could be managed in a useful way by the API. They wanted to create groups in order to better manage separate collections; decisions on how to organize these groups were determined based on Figshare functionality with subgroups and projects.

Subgroups allowed a browsing option which was very important for the theses, particularly browsing by faculty. They also wanted to be able to set custom metadata fields at the subgroup level. Theses had the most complex metadata of all the migrated content, while others mostly had Dublin Core, which translated well into Figshare.

They found that the API worked best with projects; uploading content into projects allowed them to draw quota from the group and not from the account uploading the items. Projects provided a way to usefully organize hundreds of records into logical groups that could be managed by the uploader and kept separate depending on whether the item was a thesis or a working paper, for example. Projects could also remain private but the content within them published publicly. They're also useful for usage reporting on the project level.

“I'd emphasize the fact that thinking about how you use projects if you want to do large-level migration is quite important. It certainly allows you to manage a large bulk-load into manageable chunks.”

Andrew Harrison,
Research Infrastructure Librarian at Monash University

Impact

When the theses were hosted in their Fedora repository - from 2008-2016 - they accumulated about 1.3 million views and 180,000 downloads. Since the migration of this content to Figshare in late 2016, they've been viewed over half a million times and downloaded over 45,000 times. This is on track to surpass the statistics while the content was hosted in Fedora.

Next steps

The university will be encouraging their postgraduate candidates to publish their research as they're generating it and encouraging them to link that research with their thesis to help them build a profile. This is to help improve the impact of research outcomes that the wider research data management team are implementing in 2018.

You can see Monash University's theses repository at <https://monash.figshare.com/theses>.

For more information on Monash University's content migration, including the coding required to migrate the content and work done in-house versus commissioned by the university, please reach out to Andrew: andrew.harrison@monash.edu.

Find out more:

monash.figshare.com

info@figshare.com