

# Clever Crosswalking

## Starting point

- Repository: mainly theses
- Research management system

Integrate repository & RMS  
(once-off sync, then ongoing)

## Challenges

- Different metadata granularity
- Good repo workflows but RMS now primary data source

**Clever Crosswalking - what do you take from one system to another?**  
Yanan Zhao, Kim Shepherd, Leone Hayes & Andrea Schmeer  
The University of Auckland, New Zealand, y.zhao@auckland.ac.nz  
Library Consortium of New Zealand (LCNZ), schmeer@waikato.ac.nz

**Our story started here ...**

We needed to provide a solution for a complex project that involved making the existing repository content which contains mainly theses content to becoming the institutional collection for research outputs.

- We had a schema that was based on Dublin Core which works well for simple and generic resource description, but it isn't sufficient for describing publications
- We had well defined workflows in the repository, but we needed to make the Research Management System (RMS) the primary place for data input and then push content from the RMS to our repository

This presented a number of questions:

- How could we take a very granular level of metadata from our RMS and expose it in our repository which uses less granular level of metadata schema
- Without losing anything and 2) with the least amount of work, highest amount of accuracy and facilitate the maximum reuse possible?
- Do we put the publication citation data in one single field in the repository? Or do we store the parts of a bibliographic citation separately, and how?

**Our solutions**

**Our experience:** It's better to store the parts of a bibliographic citation separately and then employ programmatic ways to render these as one field. The separate parts would still be available for other functions e.g., we used an API to feed data from the repository back to the RMS where it is needed, and thus enable greater use and reuse of the content in the digital repository.

1. Create local schemas (PROSP, data, prefix with pubs.)
2. Crosswalking each data element via repository software based on SWORD v2
3. Some data mapping using the DSpace batch metadata editing function
4. Using API services in our RMS to backwords enhance the source to prevent crosswalk overwriting data

**Workflow: A simple deposit process @ UoA**

**Our story continues ...**

- Why do we feed citation information in the repository items' metadata? Is it about locating the item/record?
- OR is it about knowing how to cite it?
- Flat metadata schemas - is it the easiest way to crosswalk?
- Complex relational models are in most cases too challenging to implement for re-use, and how do you crosswalk non metadata elements?

We would love to hear what you think of these questions; contact us at y.zhao@auckland.ac.nz or schmeer@waikato.ac.nz