

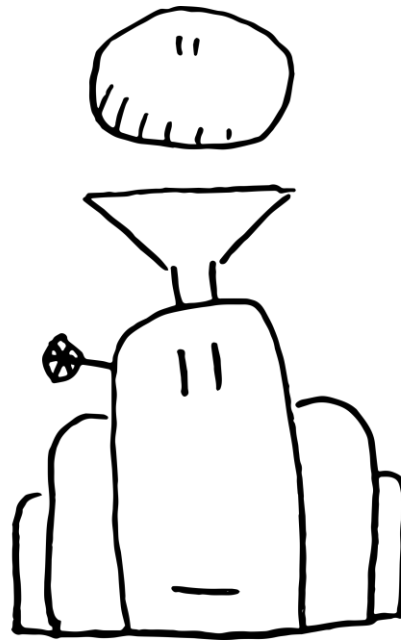
Publishing versioned datasets using OCFL and nginx

Mike Lynch – University of Technology Sydney



Australian Research Data Commons

Data repositories

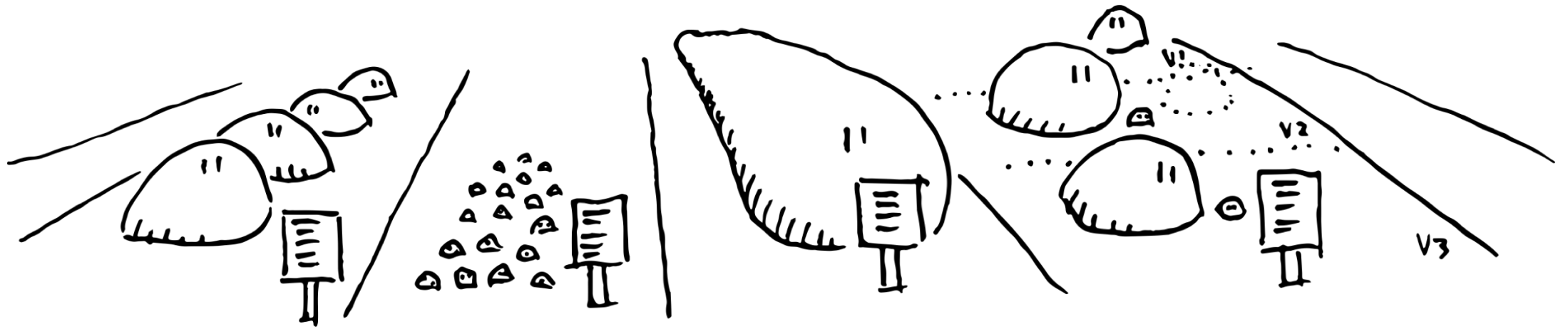


Data repositories



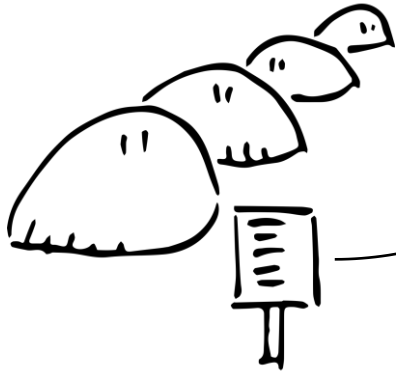
Oxford Common File Layout

- A standard layout for static, file-based repositories
- Human and machine readable (JSON)
- Lightweight file-level versioning



RO-Crate

- Research Objects + DataCrate
- JSON-LD metadata

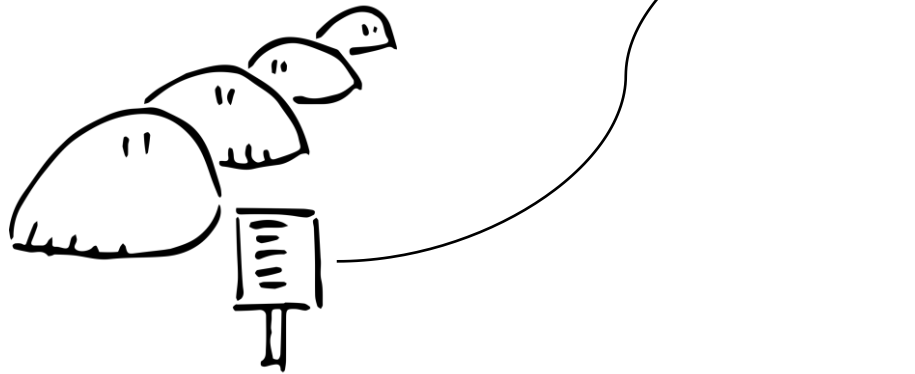


```
{
  "@context": "https://raw.githubusercontent.com/ResearchObject/ro-crate/master/docs/0.2-DRAFT",
  "@graph": [
    {
      "@type": "CreativeWork",
      "@id": "ro-crate-metadata.jsonld",
      "identifier": "ro-crate-metadata.jsonld",
      "about": {
        "@id": "./"
      }
    },
    {
      "identifier": [
        "https://dx.doi.org/10.4225/35/555d661071c76",
        {
          "@id": "_:local-id:ocfl_demo:57fc60e0-d7d2-437b-95d9-da557e809f50"
        }
      ],
      "name": "Farms to Freeways Example Dataset",
      "description": "This data set was exported from an Omeka Repository as an example of a",
      "publisher": {
        "@id": "http://westernsydney.edu.au"
      },
      "datePublished": "2015-12-01",
      "contactPoint": {
        "@id": "K.Trewin@westernsydney.edu.au"
      },
      "@type": "Dataset",
      "license": {
        "@id": "https://creativecommons.org/licenses/by/3.0/au/"
      },
      "@id": "./",
      "author": [
        {
          "@id": "http://nla.gov.au/anbd.aut-an35048773"
        },
        {

```

RO-Crate

- Research Objects + DataCrate
- JSON-LD metadata
- HTML preview



[Farms to Freeways Example Dataset](#)

Dataset: Farms to Freeways Example Dataset

[Download all the metadata for Farms to Freeways Example Dataset in JSON-LD format](#)

[Check this crate](#)

@id	/
name	Farms to Freeways Example Dataset
@type	Dataset
description	This data set was exported from an Omeka Repository as an example of a DataCrate. It contains the Collections and Items from the repository but does NOT have the exhibitions. The DOI resolves to an archive of the data elsewhere
datePublished	2015-12-01
author	<ul style="list-style-type: none">• Deborah Chambers• Carol Liston• Christine Wieneke

nginx and Solr

- fast, versioned web endpoints
- Solr indexes metadata and licences
- licences enforced on both search results and payloads



Examples

- In production for UTS data publication repository
- PARADISEC – crosswalking to OCFL and RO-Crate
- State Library of NSW – Mitchell collection of public domain books



web
upload



repository
app



data
capture

```
<?xml version="1"
<nw>
  <wa>
    <shw>
      <mw>
</mw>
</shw>
</wa>
</nw>
```

legacy metadata

Links and acknowledgements

- <https://ocfl.io/>
 - <https://researchobject.github.io/ro-crate/>
 - <https://github.com/UTS-eResearch/ocfl-nginx>
 - Docker: mikelynch/nginx-ocfl
-
- This research/project is supported by the Australian Research Data Commons (ARDC). The ARDC is enabled by NCRIS.



Australian Research Data Commons