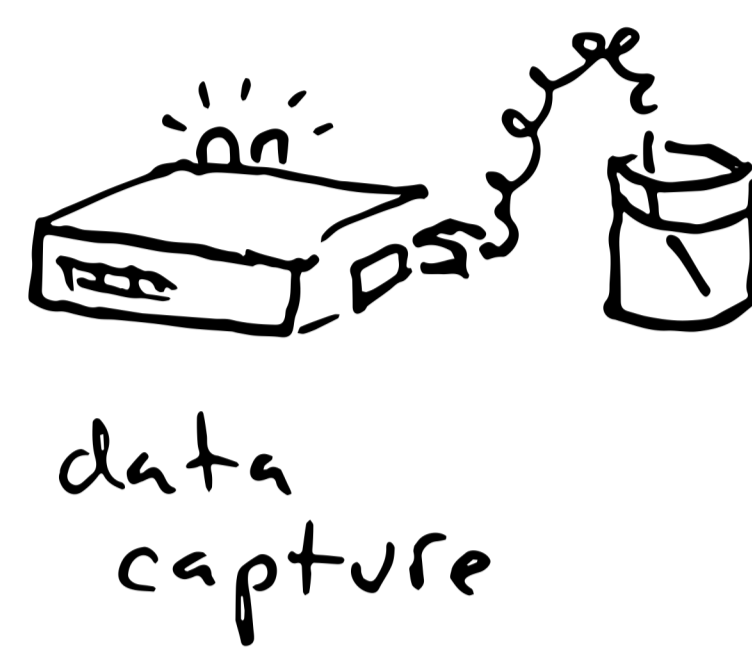


FAIR Simple Scalable Static Research Data Repository Demonstrator

Dr Peter Sefton and Michael Lynch, University of Technology Sydney

Data repositories need to:

- Scale
- Accept diverse inputs
- Safeguard access
- Outlive this year's web dev framework

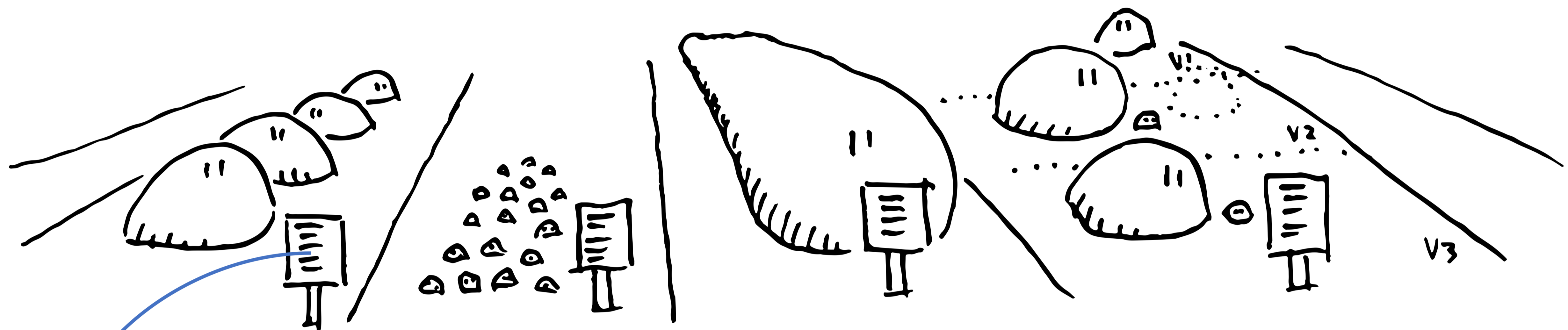


```
<?xml version="1"
<root>
  <data>
    <value>
    </value>
  </data>
</root>
```

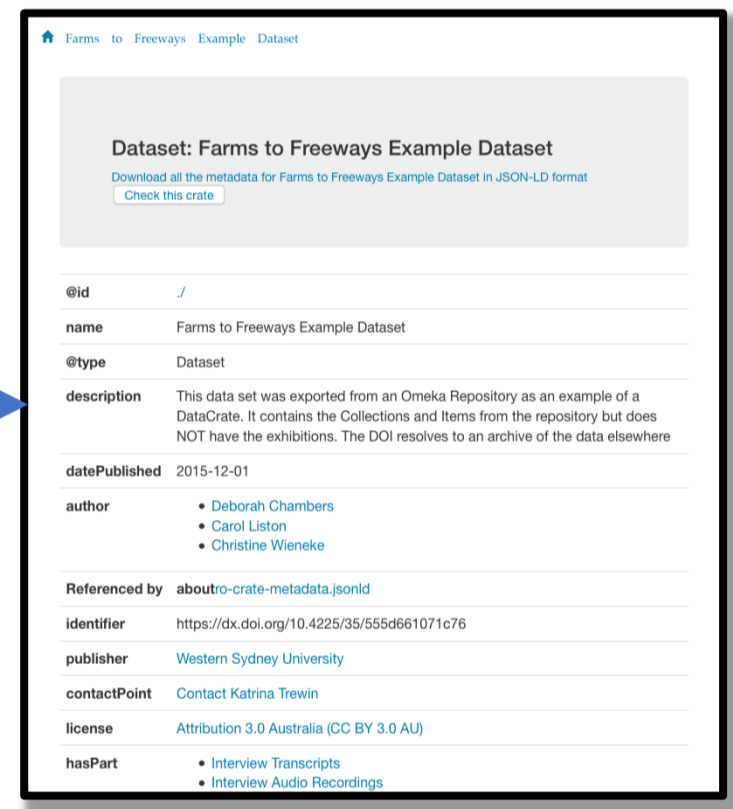
legacy metadata

OCFL: Oxford Common File Layout

- Simple standard for laying out a repository on file storage
- A Repository with Objects
- JSON inventories with content hashes and versioning



```
{
  "@context": "https://www.github.com/researchobject/ro-crate",
  "@graph": [
    {
      "@type": "CreativeWork",
      "@id": "https://doi.org/10.4225/151355664871c7c",
      "name": "Farms to Freeways Example Dataset",
      "description": "This data set was exported from an Oneka Report."
    },
    {
      "@type": "Dataset",
      "@id": "https://www.github.com/researchobject/ro-crate",
      "name": "Farms to Freeways Example Dataset",
      "description": "This data set was exported from an Oneka Report."
    },
    {
      "@type": "Dataset",
      "@id": "https://www.github.com/researchobject/ro-crate",
      "name": "Farms to Freeways Example Dataset",
      "description": "This data set was exported from an Oneka Report."
    }
  ]
}
```

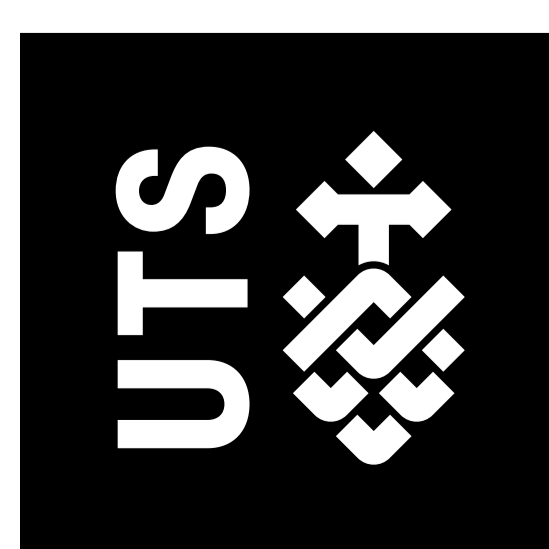
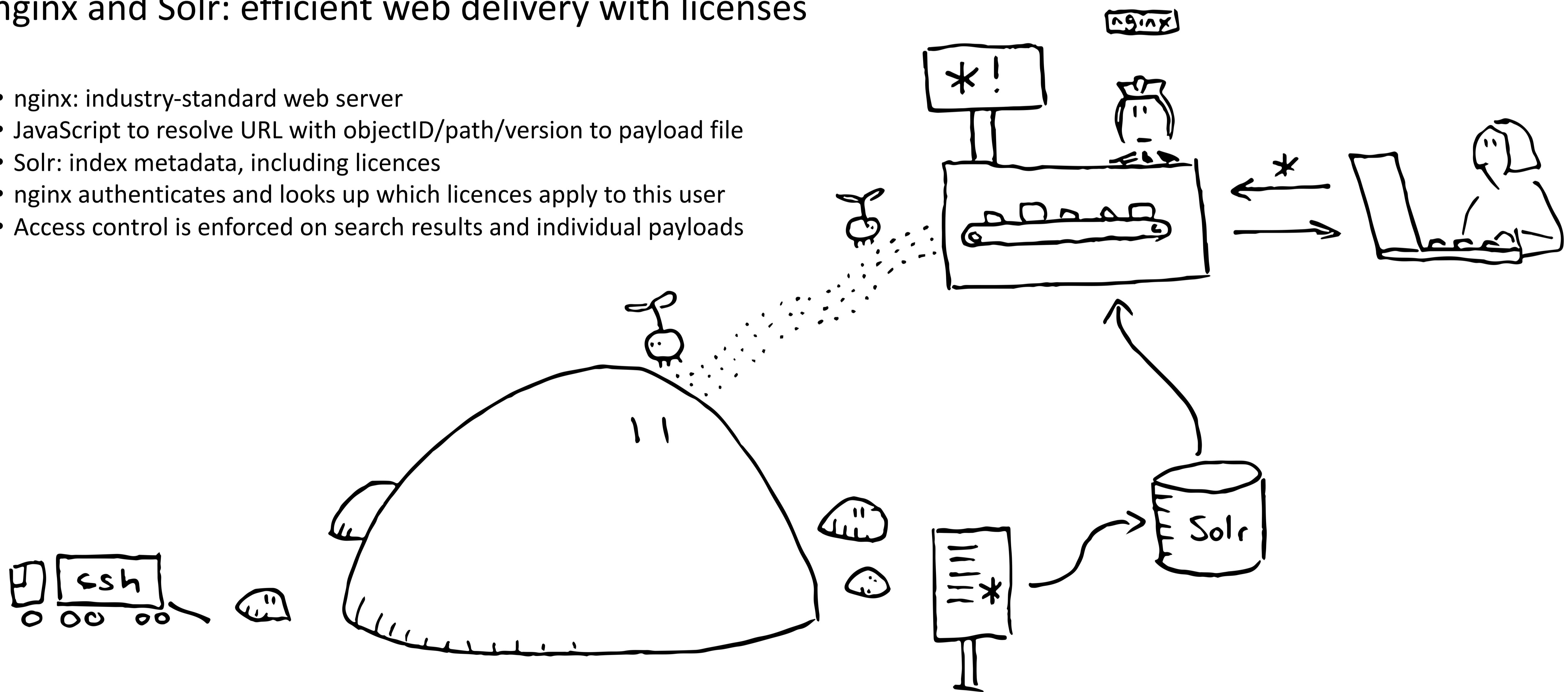


RO-Crate: JSON-LD Dataset Descriptions

- Based on earlier work: DataCrate and Research Object
- JSON-LD metadata using Schema.org – aligns with Google's Dataset Search, can be extended to include specialist vocabularies
- HTML preview: datasets are self-describing
- HTML previews can embed images: we're working on content players, maps, and more

nginx and Solr: efficient web delivery with licenses

- nginx: industry-standard web server
- JavaScript to resolve URL with objectID/path/version to payload file
- Solr: index metadata, including licences
- nginx authenticates and looks up which licences apply to this user
- Access control is enforced on search results and individual payloads



<https://ocfl.io/>
<https://researchobject.github.io/ro-crate/>
<https://github.com/UTS-eResearch/ocfl-nginx>
 Docker: mikelync/nginx-ocfl

This research/project is supported by the Australian Research Data Commons (ARDC). The ARDC is enabled by NCRIS.



Australian Research Data Commons