

# Arkisto: a repository based platform for managing all kinds of research data

Peter Sefton<sup>1</sup>, Marco La Rosa<sup>2</sup>, Michael Lynch<sup>1</sup>

1. University of Technology Sydney
2. The University of Melbourne

This presentation is

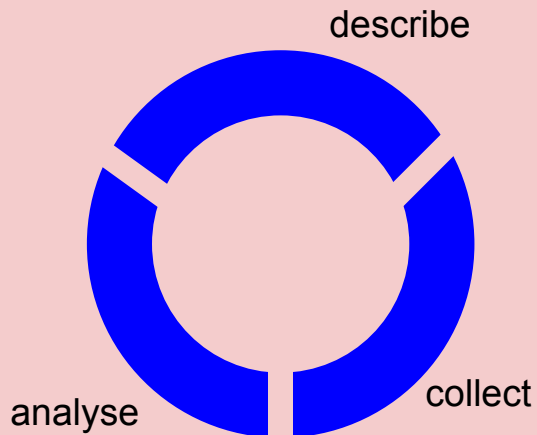
**FAIR**

driven

## Workspaces:

- working storage
- domain specific tools
- domain specific services

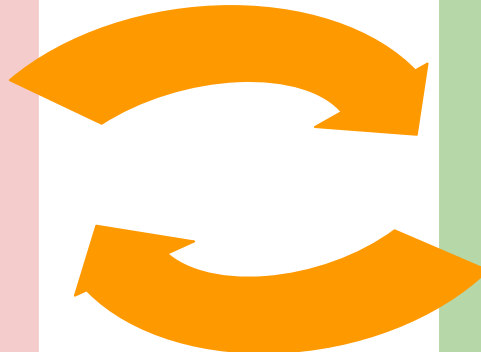
Research  
Data  
Management  
Plan



Active cleanup processes  
workspaces considered ephemeral

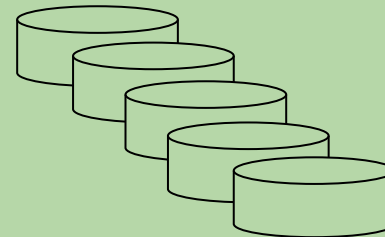
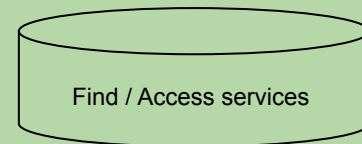
## Reusable, Interoperable data objects

- deposit early
- deposit often



reuse data objects

## Repositories: institutional, domain or both



Findable, Accessible, Reusable  
data objects



Policy based data management

Q. How can we “FIRST LOOK AFTER THE DATA”



# Oxford Common File Layout Specification

Recommendation 07 July 2020

## This version:

<https://ocfl.io/1.0/spec/>

## Latest published version:

<https://ocfl.io/latest/spec/>

## Editors:

[Andrew Hanking](#) (University of Oxford)

[Neil Jeffery](#) (University of Oxford)

[Rosaleen](#) (University)

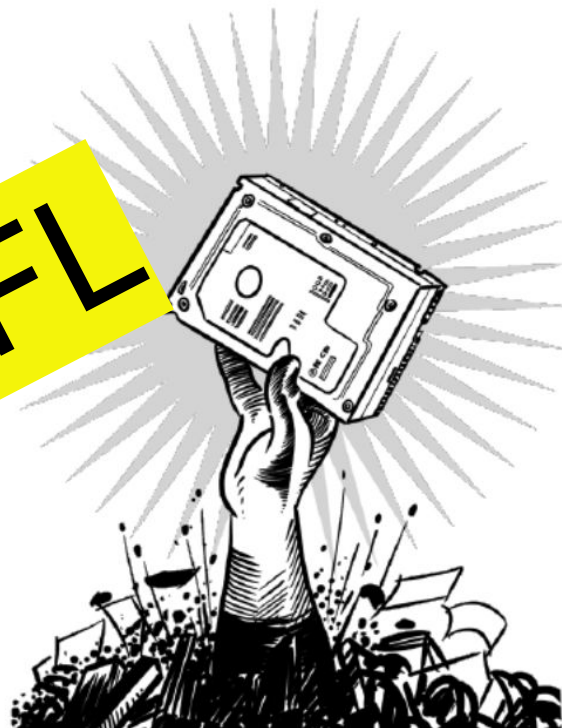
[Julian](#) (University)

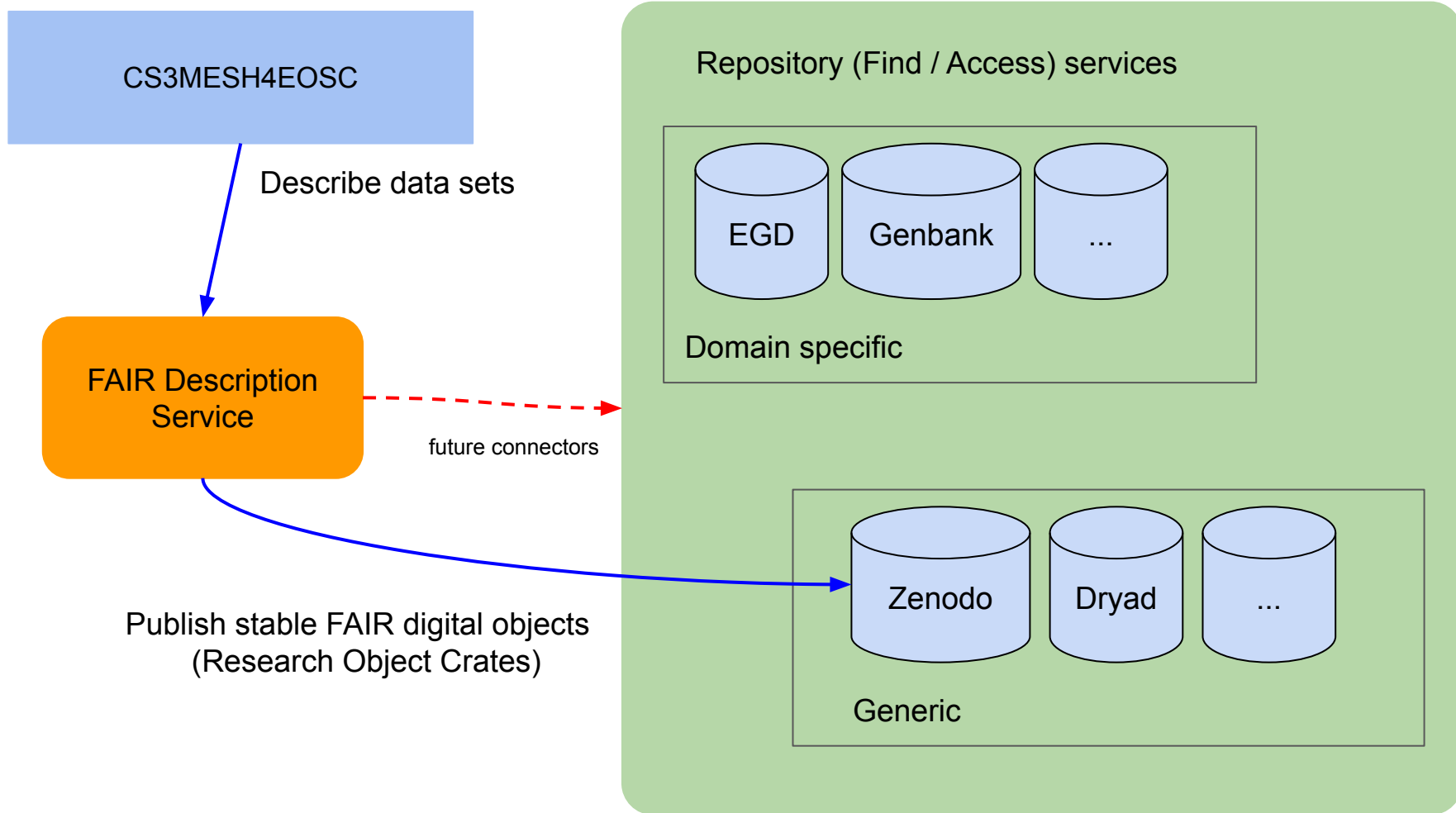
[Simeon](#) (Cornell University)

[Andrew Woods](#) (LYRASIS)

Additional Documents:

ANSWER: OCFL





Addressable  
resources



Local Data



<https://orcid.org/0000-0001-2345-6789>



ID? Title? Description?



Who created this data?



What parts does it have?



When?



What is it about?



How can it be reused?



As part of which project?



Who funded it?



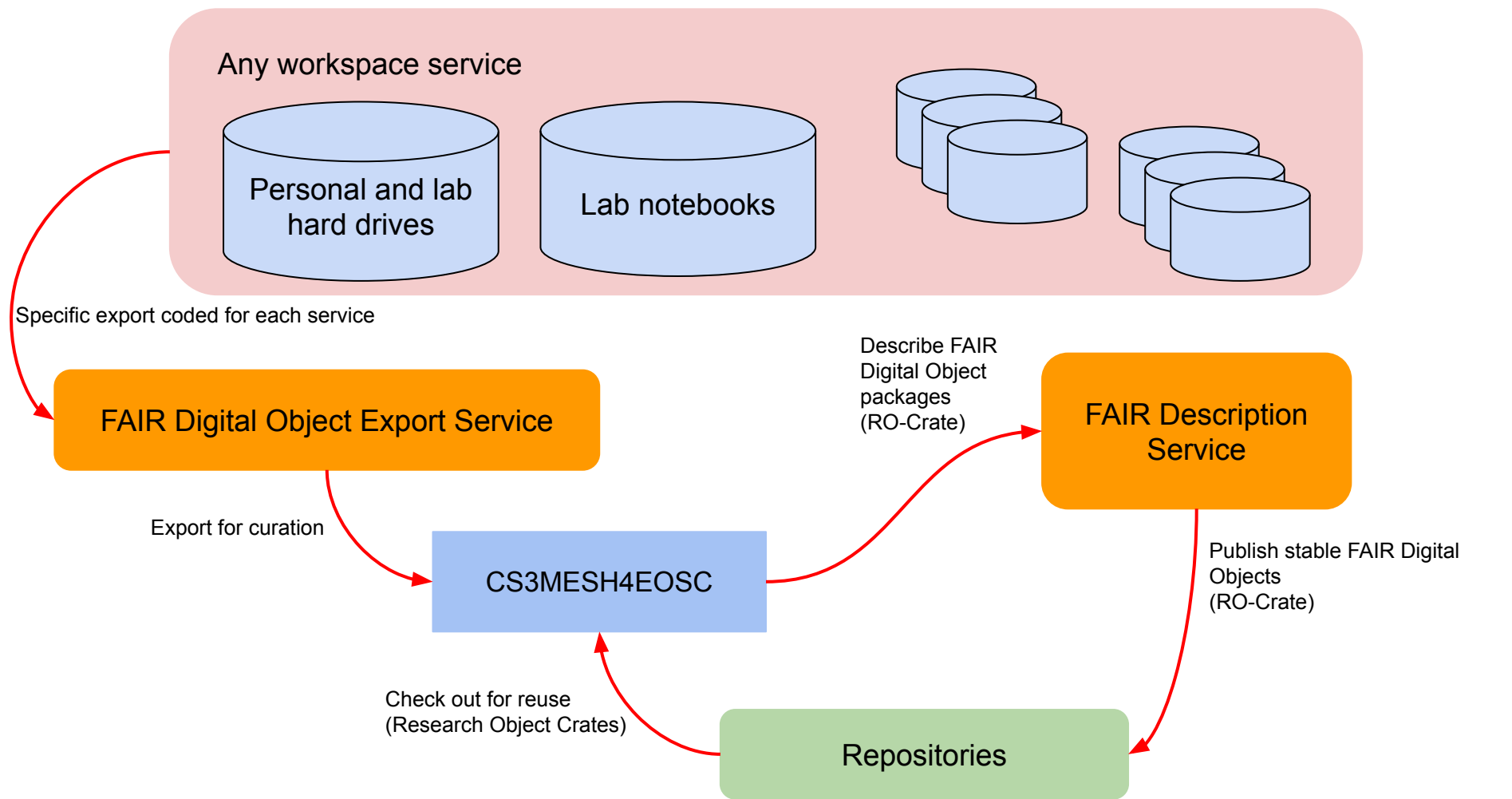
How was it made?



[https://en.wikipedia.org/wiki/Scanning\\_electron\\_microscope](https://en.wikipedia.org/wiki/Scanning_electron_microscope)

**METADATA**







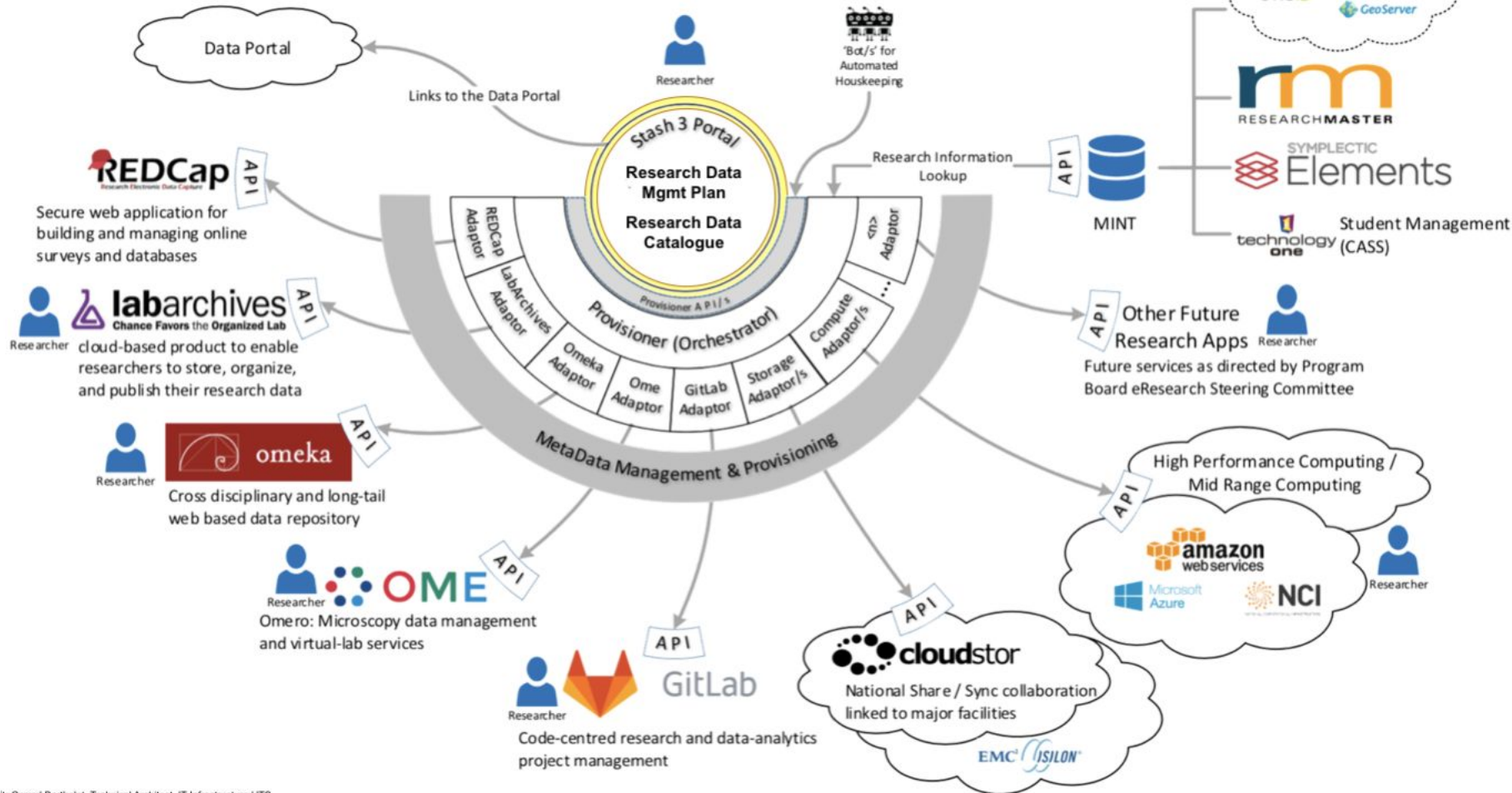
## Identity: authentication, authorisation and group services

Provisioning

workspaces

repositories

# UTS Provisioner





# TOOLS and SPECS

OCFL Spec: <https://ocfl.io/>

Research Object Crate (RO-Crate) Spec: <http://www.researchobject.org/ro-crate>

UTS: <https://github.com/UTS-eResearch/ro-crate-js>

OCFL JS

- UTS OCFL JS Implementation: <https://github.com/uts-eresearch/ocfl-js>
- CoEDL OCFL JS implementation: <https://github.com/CoEDL/ocfl-js>

UTS RO Crate / SOLR portal: <https://github.com/uts-eresearch/oni-express>

Describo:

- <https://github.com/Arkisto-Platform/describo>
- <https://github.com/Arkisto-Platform/describo-online>
- <https://github.com/Arkisto-Platform/describo-data-packs>

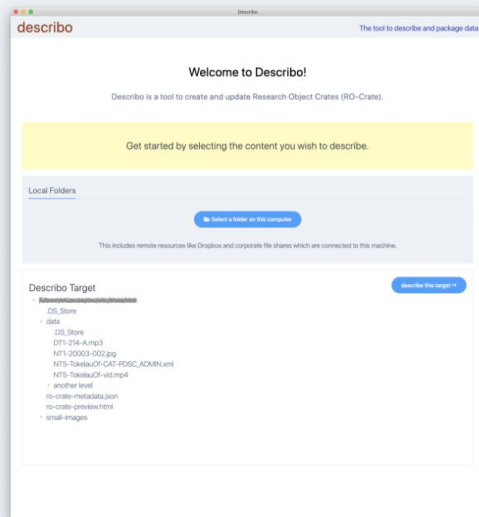
CoEDL Modern PARADISEC: <https://github.com/CoEDL/modpdsc>

CoEDL OCFL tools: <https://github.com/CoEDL/ocfl-tools>

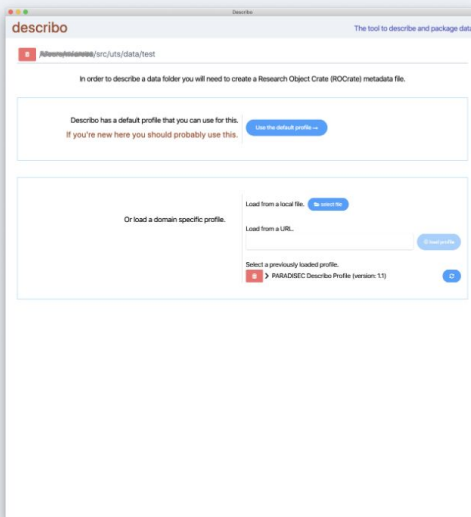
# Describo

The tool to describe and package data.

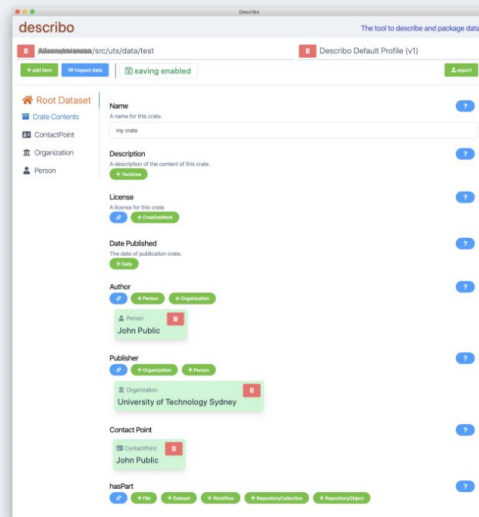
Describo is a tool to create and update **Research Object Crates (RO-Crate)**



Select a folder to describe.



Select a profile.



Describe your data.

# arkisto

## Why Arkisto

## About

## Standards

- Storage
- Packaging
- Identifiers

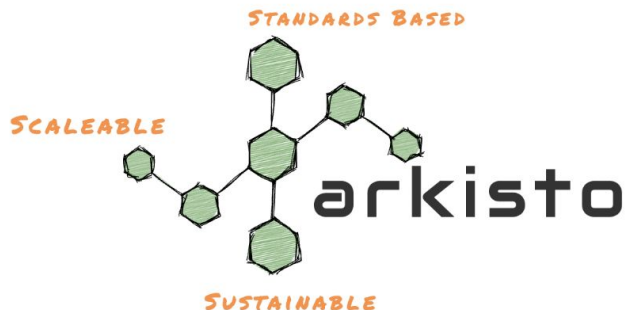
## Case Studies

- PARADISEC
- UTS Data
- Grants
- UTS Cultural Data

## Use Cases

## Tools

- Data Description
- Data Discovery
- Data Import



A scaleable, standards based platform for sustainable data.

The basis of Arkisto is that the long-term preservability of well-described data is *a/ways* the first consideration.

Data on an Arkisto deployment is always available on disc (or object storage) with a complete description *independently* of any services such as websites or APIs. Once the data is safe and well described, Arkisto has a flexible model for how data can be accessed using a variety of services.

Arkisto is built on top of [Research Object Crate \(RO-Crate\)](#) and the [Oxford Common File System](#)