# Platforms EOI: Sustainable Server-based Research Workspaces

**Project title**

Sustainable Server-based Research Workspaces

**Field of Research code(s)**

- 01 MATHEMATICAL SCIENCES
- 02 PHYSICAL SCIENCES
- 03 CHEMICAL SCIENCES
- 04 EARTH SCIENCES
- 05 ENVIRONMENTAL SCIENCES
- 06 BIOLOGICAL SCIENCES
- 07 AGRICULTURAL AND VETERINARY SCIENCES
- 08 INFORMATION AND COMPUTING SCIENCES
- 09 ENGINEERING
- 10 TECHNOLOGY
- 11 MEDICAL AND HEALTH SCIENCES
- 12 BUILT ENVIRONMENT AND DESIGN
- 13 EDUCATION
- 14 ECONOMICS
- 15 COMMERCE, MANAGEMENT, TOURISM AND SERVICES
- 16 STUDIES IN HUMAN SOCIETY
- 17 PSYCHOLOGY AND COGNITIVE SCIENCES
- 18 LAW AND LEGAL STUDIES
- 19 STUDIES IN CREATIVE ARTS AND WRITING
- 20 LANGUAGE, COMMUNICATION AND CULTURE
- 21 HISTORY AND ARCHAEOLOGY
- 22 PHILOSOPHY AND RELIGIOUS STUDIES

**EOI Lead Name**

Peter Sefton

**EOI lead Research Group**

eResearch Support

**EOI lead Organisation**

University of Technology Sydney

**Collaborator details**

<table>
<thead>
<tr>
<th>Name</th>
<th>Research Group</th>
<th>Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jason</td>
<td>Naimsey</td>
<td>University of Wollongong</td>
</tr>
<tr>
<td>Gavin Kennedy</td>
<td>Cloud Services</td>
<td>AARNet</td>
</tr>
<tr>
<td>Heidi</td>
<td>Perrett</td>
<td>Queensland Cyber Infrastructure Foundation</td>
</tr>
</tbody>
</table>

**Project description**
We have a sustainability problem with software from small CMS systems to virtual labs. Services are hard to keep up, often lack clear separation between services and data & can't be closed for fear of losing data.

We propose to start a program of "sustainabilizing and ephemeralizing" commonly used applications such as Omeka, Omero, Lab Archives, Alveo etc by exporting data from day 1 to a standards based archive and interchange format, to keep data safe, interchangeable and reusable.

The addition of new workspaces for research application services will allow transition from working project data / metadata and RDMPs through to archiving with the layout and machine/human readable format.

FAIR: Data will be Findable via search, Accessible w/ access restrictions Interoperable 'cos standard based, aiding Re-use in new systems.

Low-risk: OCFL[1] is based on decades-old file technologies and deep experience of data stewardship and aer quick to implement, which means if a better standard emerges migration is cheap. RO-Crate[2] aggregates widely used existing standards and current best practice for metadata as used by ARDC, Google and several Research Data Alliance groups.

### Existing technology

#### Adopt

The RedBOX RDM platform will form the basis for this work - it is sustained via a subscription model and widely used.

RO-Crate & OCFL standard

Provision: Cloudstor group drives, dropbox, OneDrive, Raw disk, S3: RClone.plugin, Github, Gitlab, Bitbucket, LabArchives, OneNote, other notebook/CMS services, Omeka, Mukurtu CMSs

PARADISEC Nabu: Example of a collection-specific application where data and app can be teased-apart and resources shared with other projects

Alveo: Potential re-engineering of Alveo Virtual Lab data for broad access

Teraform

**Compute / HPC:** Ronin, NCI, Nectar, AWS, GCP

#### Adapt

This project will generate techniques and code for data-export services from workspaces and/or adapt them to work with data stored in OCFL with RO-Crate data packages; as proven @UTS and UMelb via ARDC discovery grants: PoC with PARADISEC data adapted - complete with hooks for access control..

#### Build

Platform is light-weight “glue” code:

RO-CRate export adaptors so data can live on using generic storage/search with authorisation via plugins to do group/licence/cultural-protocol management as a appropriate.

Via RedBOX plugins for provisioning pre-configured instances of software with good metadata foundations. Builds on UOW/ARDC discovery: streamlining the onboarding of new ReDBox Workspaces, utilising Teraform.

### Anticipated requirements

<table>
<thead>
<tr>
<th><strong>Annual funding</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$100,000 - $199,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Proposed length</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>3 years</td>
</tr>
</tbody>
</table>

### Other information

**Other information you wish to provide**

See other EOI that build on GFA-182 : D&S/IR11 FAIR Simple Scalable Static Research data and the Melbourne U Modularising PARADISEC discovery grant.
CloudStor’s collaboration will involve the ability to use ReDBox as CloudStor’s metadata portal for Group Drives. This integration will be available to all users of CloudStor that have Group Drives and aligns with FAIR data principles which cover all sector users that have CloudStor Group Drives beyond the 25+% of the sector using ReDBox.

**Terms**

**I agree to the terms**

Yes