Using *ReDBox* 2 for Research Data Management Plans (RDMPs), Connecting to cloud services, and Archiving

Jason Nairnsey – Nairnsey@uow.edu.au
Research Data, Storage & Management (RDSM)

5 year program at UOW

- **Data Storage**
  - On Campus
  - High Performance Computing
  - Collaboration
  - Public Cloud

- **Data Management**
  - Metadata management
  - RDMPs
  - Workspaces / Links to data locations
  - Plan, Create, Publish & Archive within the research project lifecycle
Research Data Workflows

Embedding research data management into the research lifecycle

- Development of research proposal
- Source Funding
- Compliance (Ethics, RDMPs etc)
- Data Collection
- Data preparation / Analysis
- Publishing
- Research Translation / Commercialisation
Research Data Workflows and the Supporting Applications

Development of research proposal
Source Funding
Compliance (Ethics, RDMPs etc)
Data Collection
Data preparation / Analysis
Publishing
Translation / Commercial-isation

Faculties, Institutes & Schools
Grants Systems MGMT Application
Education / ReDBox / Ethics Apps
Instruments / Surveys / Existing datasets / eNotebooks
Tools: HPC, R, RedCap, XNAT, SQL, GitLab etc
Institutional Repository / Academic Profiles / MGMT
Commercial Research department
- **Requirements** split into:
  - **Essential** (policies and audit requirements)
  - **Recommended** (funder requirements, statutory guidelines, strategic alignments)
  - **Nice to have** (advanced of above)

- **Integrations** with:
  - Inductions / on-boarding (**PODS**)
  - Ethics and Integrity training (**RSO**)
  - How to use tools, inc guidelines for Storage, Backups, Archiving, Versioning etc (**IMTS**)
  - RDM Compliance and discipline specific best practice (**Faculties, Schools & Institutes**)
  - Digital Literacy / Research Students (**Library**)

- **Other**
  - All teams to be aligned with the same message and know where to send people to
  - Multiple Training locations
  - Delivery timeframes - coincide with ReDBox application release
Taking a look at ReDBox

Welcome to ReDBox

Plan a Research Project
- Create RDMP
- View & Update RDMP
- Get advice

Organise Workspaces
- View workspaces
- List of available services

Manage data
- Create archival Data Record
- View & Update archival Data Record
- Transfer responsibility

Publish and share
- Create Data Publication
- View & Update Data Publication

Data Storage
- Estimate the size of the data to be collected (*)
  - 100GB - 2TB
- Please detail arrangements for secure storage and transfer during the project (*)
  - eResearch storage platforms (e.g. CloudStor, StorageGrid)
- If other, please provide further details:

Ethics
- Human Ethics Protocol Number (if available)
  - HE2018/999
- Information about ethical considerations relevant to this project and how they will be managed
  - Minors - parental approval

Will the data provided to the researchers be:
- Identifiable
- Re-identifiable
- De-identified

Please specify the format which the data will be provided
- Survey results
RDMPs connecting to Workspaces

ReDBox an RDMP tool – providing Management for Research Data Storage
ARDC Discovery Project

Deliverables from Grant

- **Discovered** how best to Streamline the on-boarding of new workspaces
  - ✔️

- **Discovered** how to provision a Google Cloud Platform (GCP) compute instance from ReDBox
  - ✔️

- **Implemented** “template” using Teraform to streamline the on-boarding of new workspaces
  - ✔️

- **Implemented** the provisioning of a GCP compute instance from within ReDBox
  - ✔️

- **Shared** all code with the ReDBox community
  - ✔️

Terraform Plugin: [https://github.com/redbox-mint/redbox-hook-terraform](https://github.com/redbox-mint/redbox-hook-terraform)
# Workspaces

## Associated Workspaces

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Cloud - Omeka</td>
<td></td>
</tr>
<tr>
<td>GitLab</td>
<td></td>
</tr>
<tr>
<td>OMERO</td>
<td></td>
</tr>
<tr>
<td>Google Cloud - CeNTOS</td>
<td></td>
</tr>
</tbody>
</table>

## Create or Link a Workspace

Select from the list the desired app.

- Google Cloud - Omeka
- GitLab
- OMERO
- Google Cloud - CeNTOS

---

## Workspaces

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Location</th>
<th>Plan</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>workspace vm 1</td>
<td>workspace vm 1 description</td>
<td>IP Address: 35.189.6.2.18, User: shilo2</td>
<td>test-shilo</td>
<td>Google Cloud - CeNTOS 7</td>
</tr>
<tr>
<td>jason test GCP1</td>
<td>A test VM</td>
<td>IP Address: 35.189.6.2.18, User: naimsey</td>
<td>Test Project</td>
<td>Google Cloud - CeNTOS 7</td>
</tr>
<tr>
<td>omeka 1</td>
<td>test omeka instance</td>
<td>Omeka URL: <a href="http://35.244.76.68">http://35.244.76.68</a>, SSH User: shilo2</td>
<td>test-shilo</td>
<td>Google Cloud - Omeka</td>
</tr>
<tr>
<td>Test0omekainstancencGCP</td>
<td>An Omeka test from Redbox</td>
<td>Omeka URL: <a href="http://35.189.6.2.18">http://35.189.6.2.18</a>, SSH User: OmekaTest1</td>
<td>Test Project</td>
<td>Google Cloud - Omeka</td>
</tr>
<tr>
<td>TestGCPCentos2</td>
<td>Testing GCP Instance V2</td>
<td>IP Address: 35.244.76.68, User: jasonetest</td>
<td>Test Project</td>
<td>Google Cloud - CeNTOS 7</td>
</tr>
</tbody>
</table>

---

[University of Wollongong Australia]
What is GCP?
Google (GCP) Workspaces

The Process

**Workspaces**

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jason test GCP1</td>
<td>A test VM</td>
</tr>
<tr>
<td>TestGCPCentos2</td>
<td>Testing GCP Instance v.</td>
</tr>
<tr>
<td>TestOmekaInstanceGCP</td>
<td>An Omeka test from Ri</td>
</tr>
</tbody>
</table>

**Create or link a workspace**

Select from the list the desired app

- Google Cloud - Omeka
- GitLab
- OMERO

---

**Google Cloud - CeNTOS 7**

Provisions a CeNTOS 7 instance in a VM within your institution's Google Cloud account.

**Workspace Name (*)**

**Description**

**SSH Username (*)**

**SSH Public Key (*)**
# GCP Workspace continued

## Workspaces
### Associated workspaces

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jason test GCP1</td>
<td>A test VM</td>
<td>IP Address: 35.189.6.218, User: nairnsey</td>
</tr>
<tr>
<td>TestOmekainstanceInGCP</td>
<td>An Omeka test from Redbox</td>
<td>Omeka URL: <a href="http://35.189.6.218">http://35.189.6.218</a>, SSH User: OmekaTest1</td>
</tr>
<tr>
<td>eResearch Test'1</td>
<td>A virtual machine for presentation purposes</td>
<td><strong>Workspace still provisioning, please wait.</strong></td>
</tr>
<tr>
<td>TestGCPCentos2</td>
<td>Testing GCP Instance V2</td>
<td>IP Address: 35.244.76.68, User: jasontest</td>
</tr>
</tbody>
</table>

## Workspaces
### Associated workspaces

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>eResearch Test'1</td>
<td>A virtual machine for presentation purposes</td>
<td><strong>IP Address: 35.244.76.68, User: JasonNairnsey</strong></td>
</tr>
<tr>
<td>Jason test GCP1</td>
<td>A test VM</td>
<td>IP Address: 35.189.6.218, User: nairnsey</td>
</tr>
<tr>
<td>TestOmekainstanceInGCP</td>
<td>An Omeka test from Redbox</td>
<td>Omeka URL: <a href="http://35.189.6.218">http://35.189.6.218</a>, SSH User: OmekaTest1</td>
</tr>
<tr>
<td>TestGCPCentos2</td>
<td>Testing GCP Instance V2</td>
<td>IP Address: 35.244.76.68, User: jasontest</td>
</tr>
</tbody>
</table>
GCP Compute Engine

VM instances

- terraform-gcp-linux-12c1f2e704bef1bbe22f9b665ef81b
  - Zone: australia-southeast1-b
  - Internal IP: 10.152.0.48
  - External IP: 35.244.76.68
  - Connect options:
    - Open in browser window
    - Open in browser window on custom port
    - Open in browser window using provided private SSH key
    - View gcloud command
    - Use another SSH client

- terraform-gcp-linux-e691ce57a23655347ef30ae472fb0308
  - Zone: australia-southeast1-b

- terraform-gcp-omero-156976e82ef2f2e360c19b599eced90fe
  - Zone: australia-southeast1-b
What is omeka desktop
omeka desktop

Install Omeka S

Create the first user
- Email: nairmsey@uow.edu.au
- Confirm Email: nairmsey@uow.edu.au
- Password: ************
- Confirm Password: ************
- Display Name: Omeka eResearch Australasia 2019

Settings
- Installation Title: Omeka eResearch Australasia 2019
- Time Zone: Australia/Brisbane

Submit
omeka desktop contd
Acknowledgements – (without whom:)

Users

- Deakin University
- Flinders University
- James Cook University
- La Trobe University
- University of South Australia
- Macquarie University
- The University of Newcastle
- UNE University of New England
- UTS University of Technology Sydney
- Western Sydney University
- Data LifeCycle
- QCIF
- University of Technology Sydney
- Google Cloud
- Australian Research Data Commons
- University of Wollongong