IDIA Science Gateway For Radio Astronomy ACP 2023

Rob Simmonds – rob@idia.ac.za

Professor UCT Computer Science

Acting Director IDIA

CTO ilifu

Talk Overview

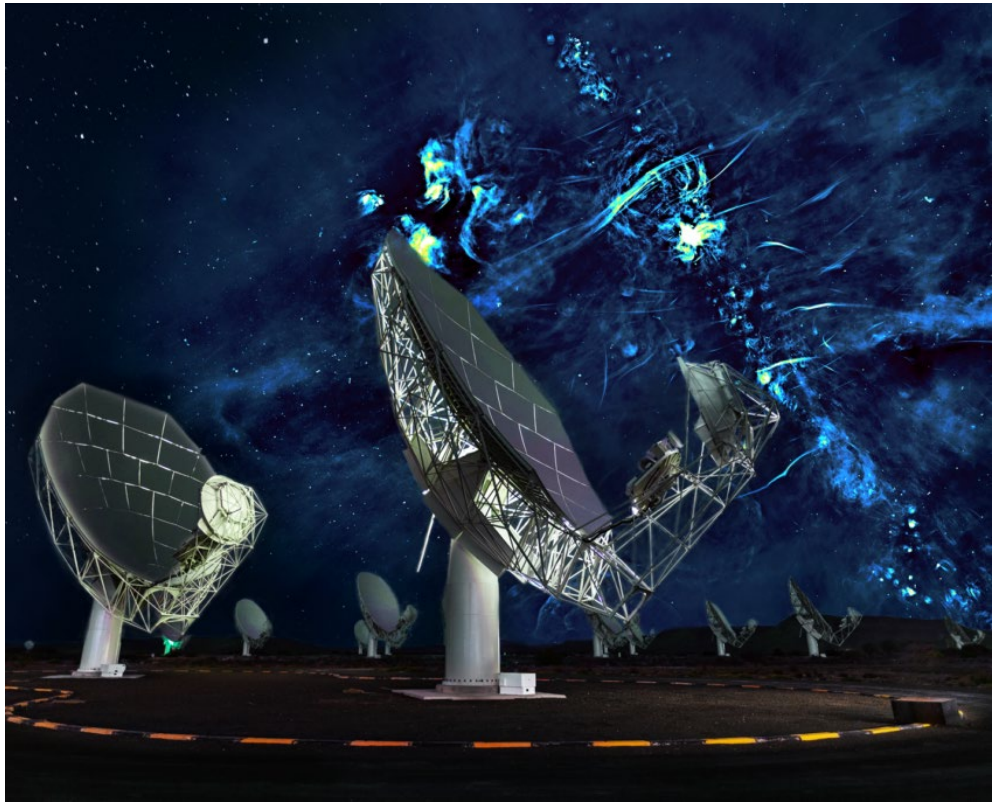
- Science Gateways in general
- MeerKAT and the SKA
- Federation
- The IDIA Science Gateway
- African shared computing initiatives

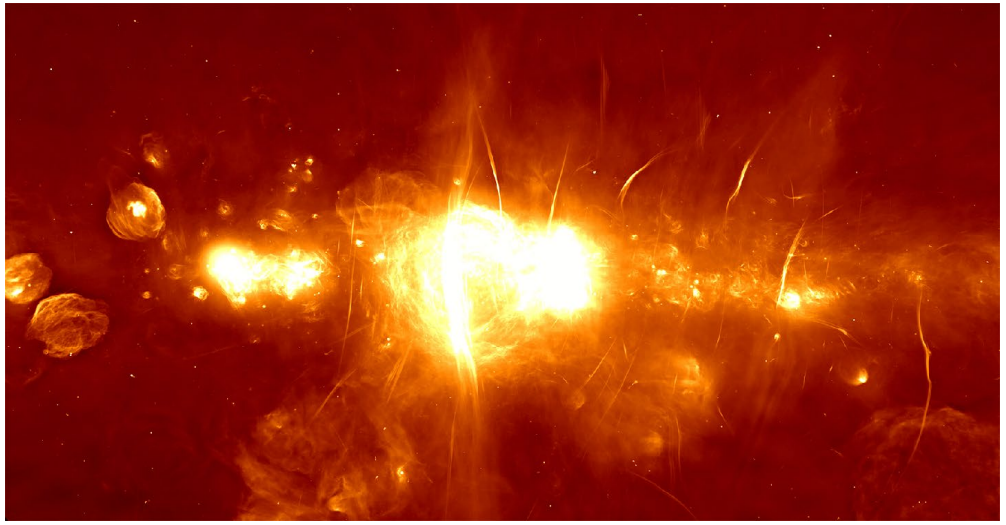
IDIA

- Partnership of three South African Universities
  - University of Cape Town
  - University of the Western Cape
  - University of Pretoria
- Aims to build the capacity and expertise of the South African university community in data intensive research to support:
  - MeerKAT large survey science projects
  - Large projects on other SKA pathfinder telescopes
- Runs hackathons and provides resources to other organisations running hackathons
- Working on the path to the SKA telescope

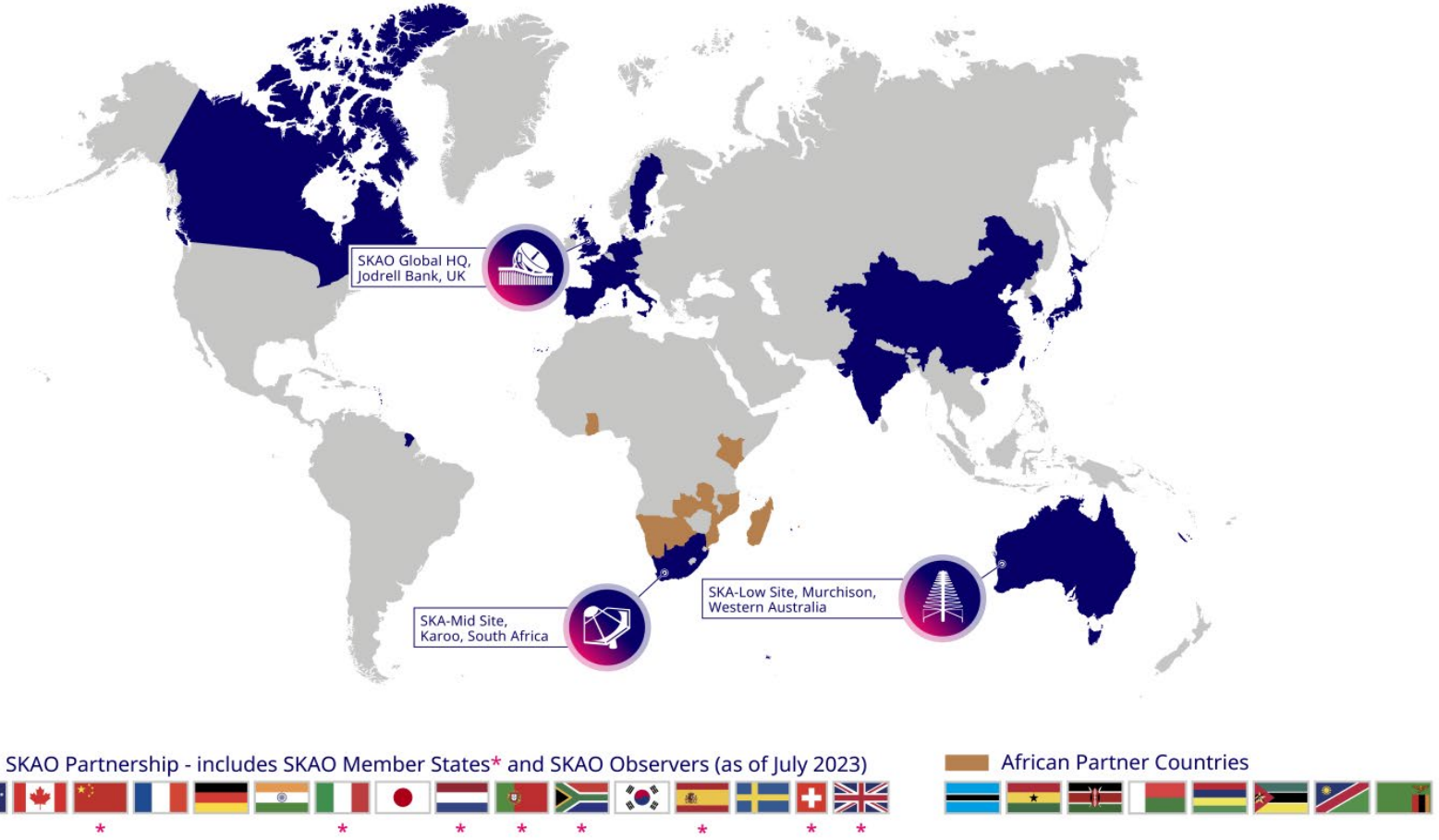
MeerKAT

- MeerKAT
  - South Africa's SKA precursor
  - Will form the core of the SKA-mid telescope
- Completed (on schedule and within budget) in Mid-2018
- Delivering Transformational Science from Day One
- Owned and operated by South Africa until SKA is built
  - See http://sarao.ac.za





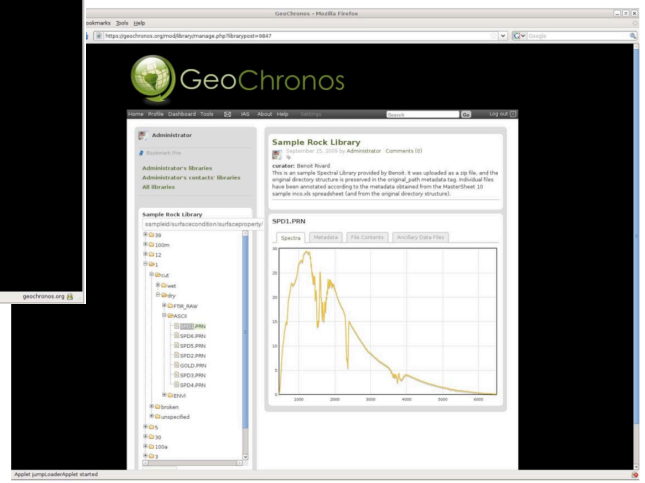
Square Kilometre Array (SKA)

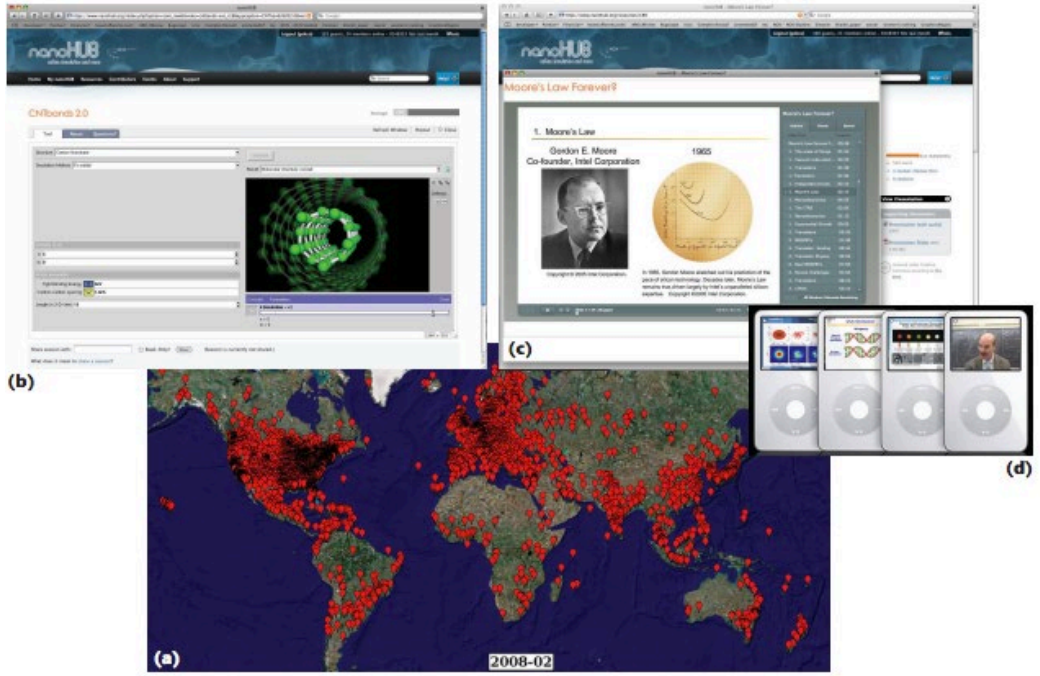


- See http://skao.int

Science Gateways

- SGs bring together a set of web-based tools that are useful to some community
- Communities are generally domain based, with tools added to support their work and collaborations



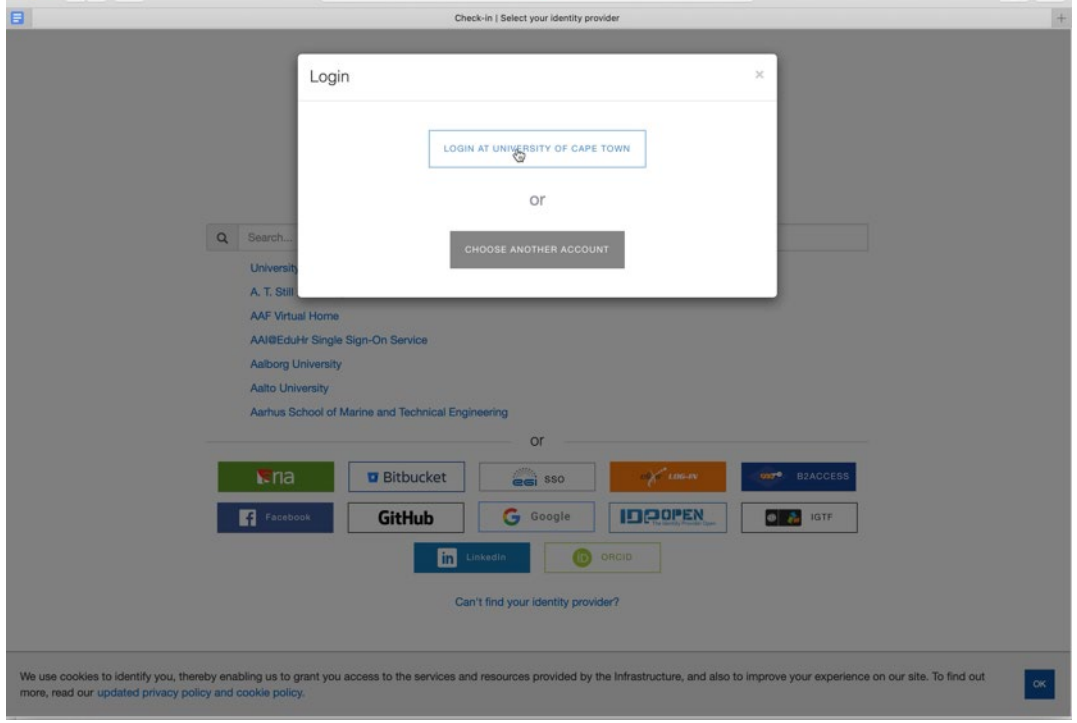


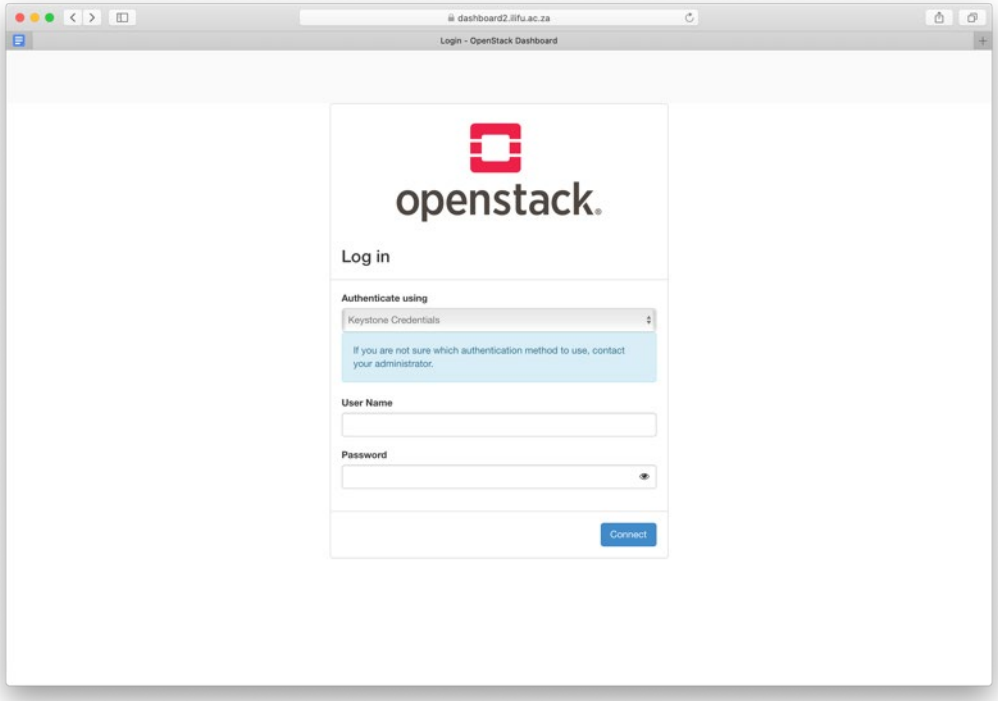
Federation

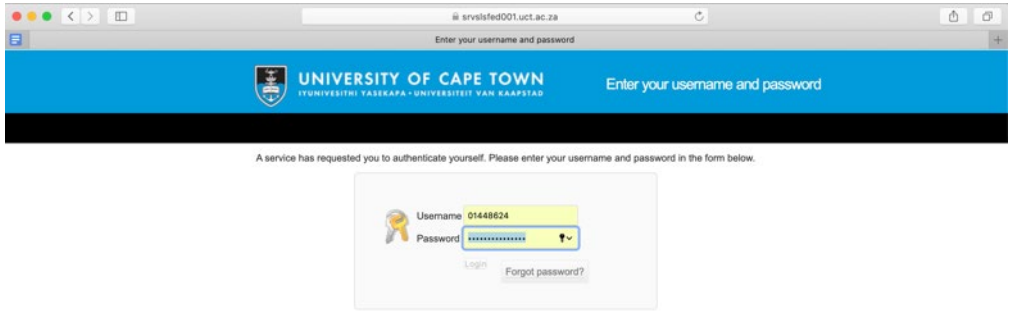
- There are many forms of federation
- Now, most often refers to using common Authentication tools
- More advanced systems also use common Authorization tools
- Grid systems provided these, plus common APIs to many services, but were not user friendly
  - Most grid frameworks designed to sit out of sight of users, controlled by automation systems, or by Science Gateways
- Improvement in web-based federation tools is making them user friendly
- Authentication provided by IdPs which could be from an institution or a company (e.g., Google)

Working with EGI

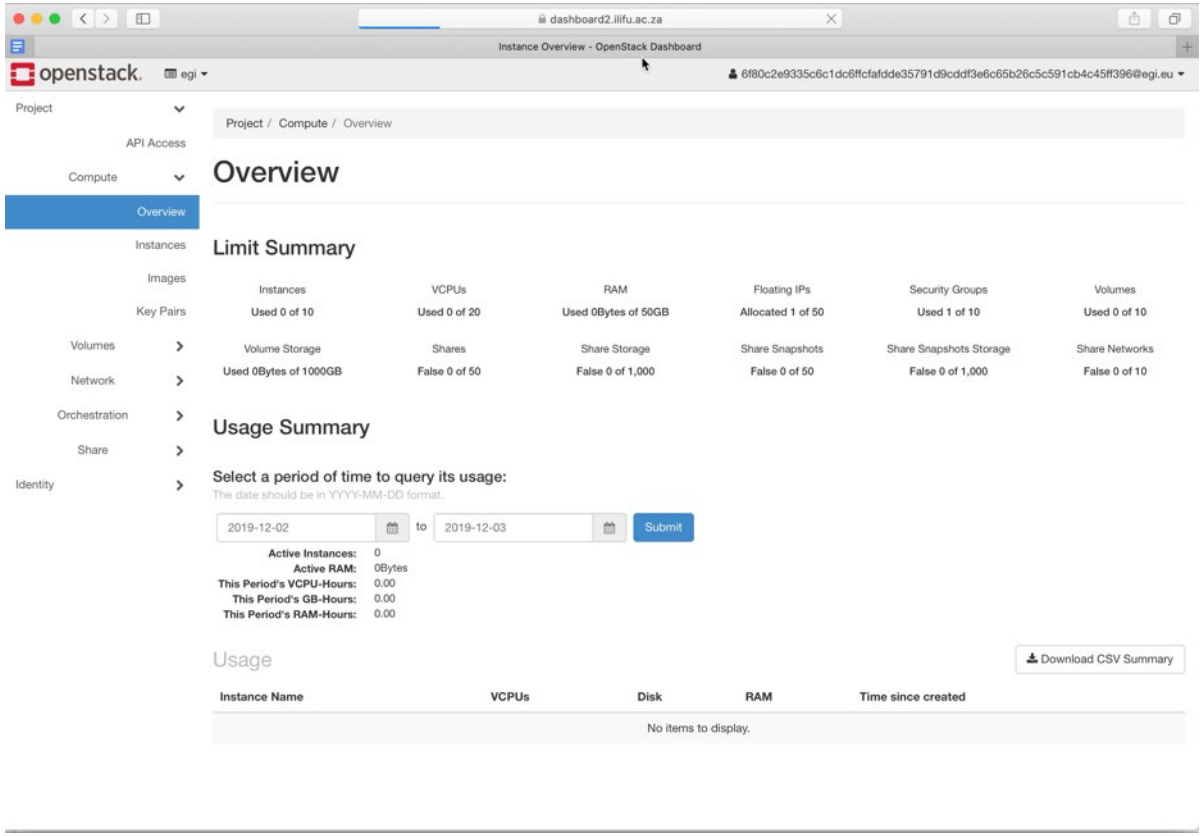
- IDIA's ilifu cluster accessible via EGI Checkin for about 4 years
- Became part of the production EGI Cloud Federation in 2022
- Most (all?) EOSC systems accessed via EGI Cloud Federation
- EGI Checkin provides access to many Identity Providers (IdPs)
  - These include Edugain / SAFIRE as well as commercial IdPs (e.g., Google)
- The EGI CoManage instance hosts groups that can be managed by individual research teams
- Groups can be linked OpenStack tenants on ilifu so members of those groups can access the tenants from the ilifu dashboard

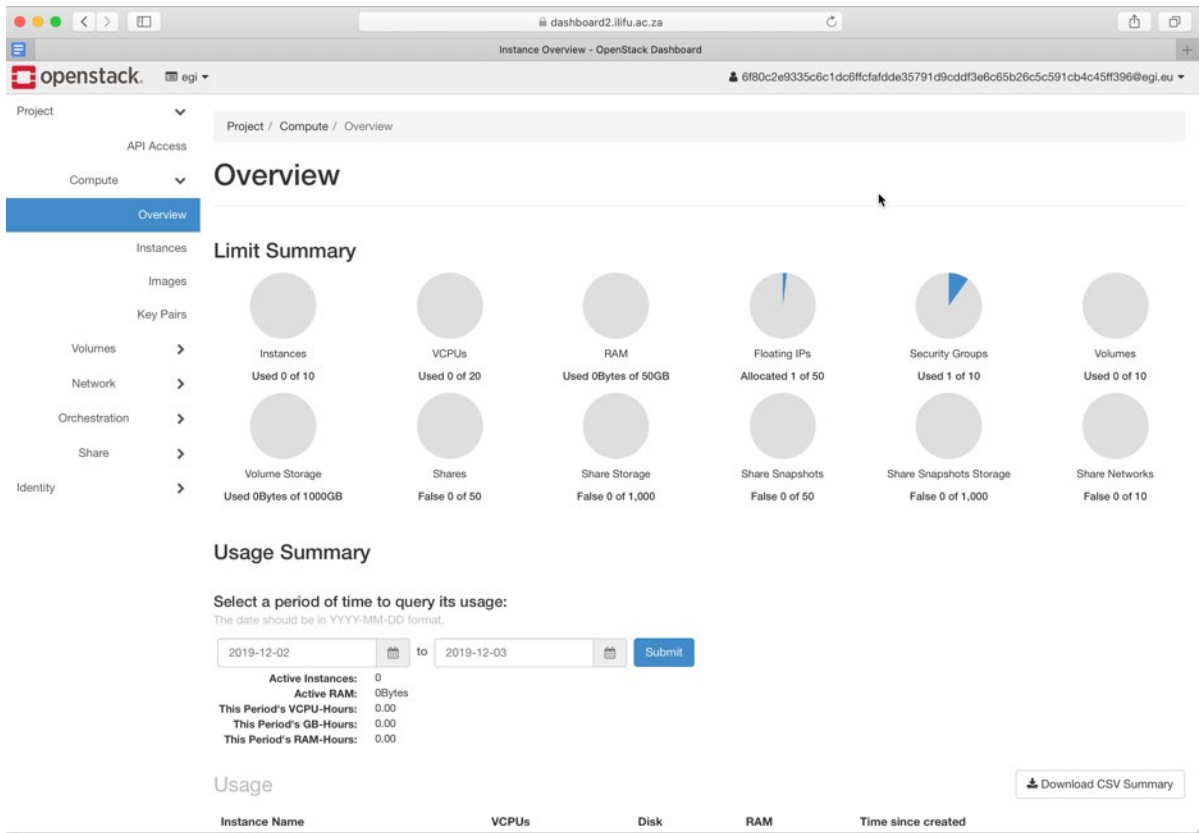

Check-in | Select your identity provider
Login
LOGIN AT UNIVERSITY OF CAPE TOWN
or
CHOOSE ANOTHER ACCOUNT
Search...
University
A. T. Still
AAF Virtual Home
AAI@EduHr Single Sign-On Service
Aalborg University
Aalto University
Aarhus School of Marine and Technical Engineering
or
Bitbucket
sso
B2ACCESS
Facebook
GitHub
Google
IGTF
LinkedIn
ORCID
Can't find your identity provider?
We use cookies to identify you, thereby enabling us to grant you access to the services and resources provided by the Infrastructure, and also to improve your experience on our site. To find out more, read our updated privacy policy and cookie policy.
OK

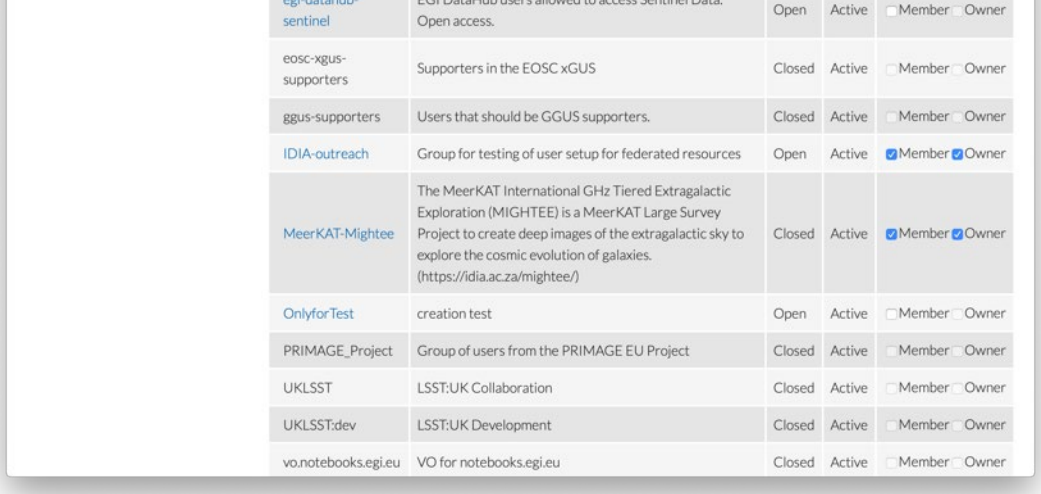

dashboard2.ilifu.ac.za
Login - OpenStack Dashboard
openstack.
Log in
Authenticate using
Keystone Credentials
If you are not sure which authentication method to use, contact your administrator.
User Name
Password
Connect


srvslsfed001.uct.ac.za
Enter your username and password
UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD
Enter your username and password
A service has requested you to authenticate yourself. Please enter your username and password in the form below.
Username 01448624
Password
Login
Forgot password?

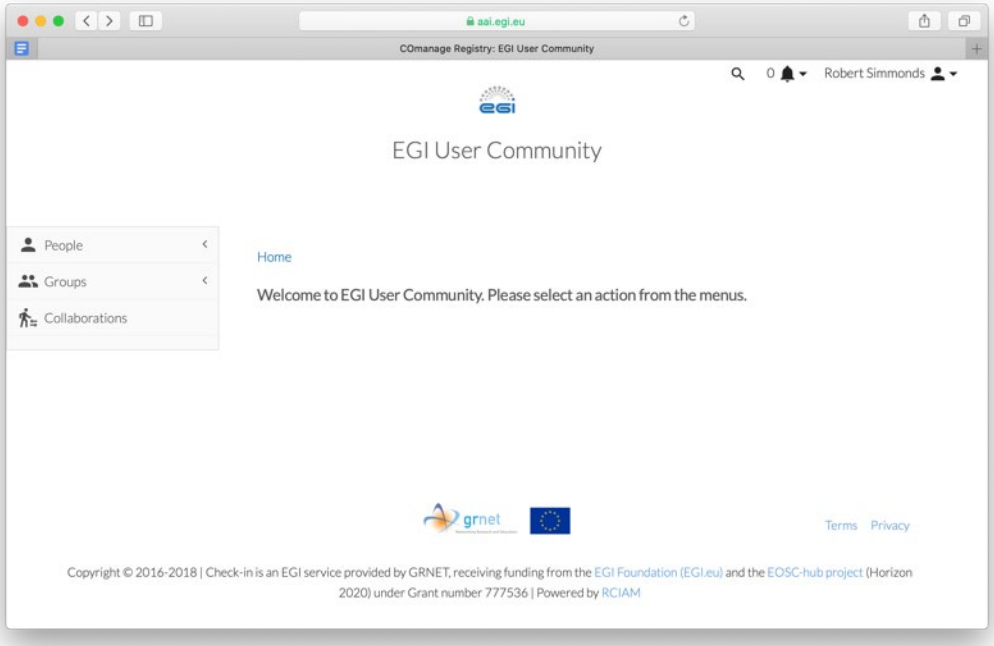
Copyright © 2018 University of Cape Town

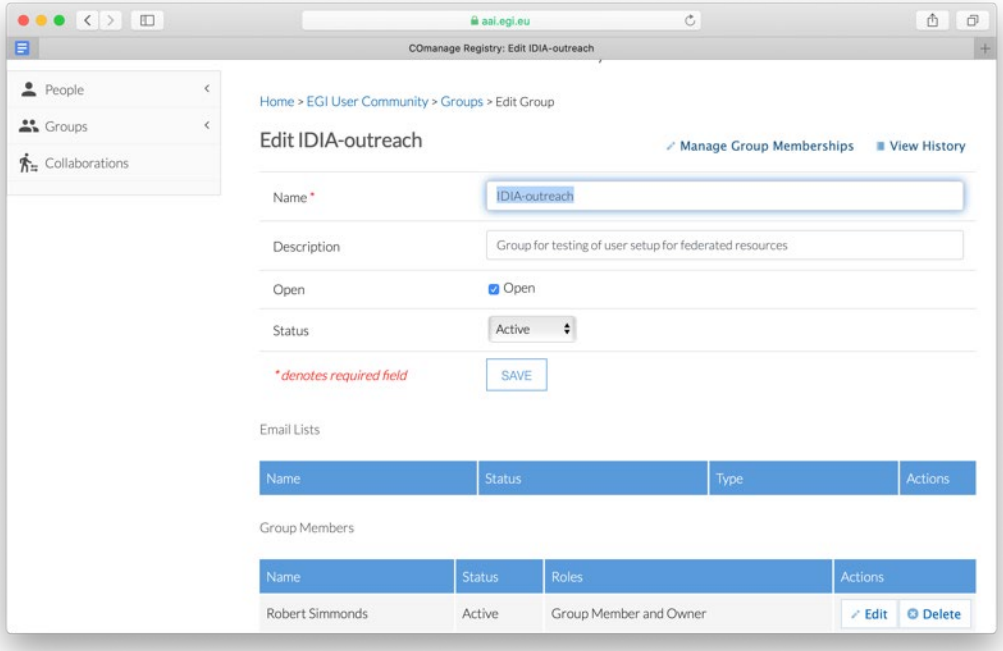

dashboard2.ilifu.ac.za
Instance Overview - OpenStack Dashboard
openstack.
egi
6f80c2e9335c6c1dc6ffcfafdde35791d9cddf3e6c65b26c5c591cb4c45ff396@egi.eu
Project
API Access
Compute
Overview
Instances
Images
Key Pairs
Volumes
Network
Orchestration
Share
Identity
Project / Compute / Overview
Overview
Limit Summary
Instances
Used 0 of 10
VCPUs
Used 0 of 20
RAM
Used 0Bytes of 50GB
Floating IPs
Allocated 1 of 50
Security Groups
Used 1 of 10
Volumes
Used 0 of 10
Volume Storage
Used 0Bytes of 1000GB
Shares
False 0 of 50
Share Storage
False 0 of 1,000
Share Snapshots
False 0 of 50
Share Snapshots Storage
False 0 of 1,000
Share Networks
False 0 of 10
Usage Summary
Select a period of time to query its usage:
The date should be in YYYY-MM-DD format.
2019-12-02
to
2019-12-03
Submit
Active Instances: 0
Active RAM: 0Bytes
This Period's VCPU-Hours: 0.00
This Period's GB-Hours: 0.00
This Period's RAM-Hours: 0.00
Usage
Download CSV Summary
Instance Name
VCPUs
Disk
RAM
Time since created
No items to display.


dashboard2.ilifu.ac.za
Instance Overview - OpenStack Dashboard
openstack.
egi
6f80c2e9335c6c1dc6ffcfafdde35791d9cddf3e6c65b26c5c591cb4c45ff396@egi.eu
Project
API Access
Compute
Overview
Instances
Images
Key Pairs
Volumes
Network
Orchestration
Share
Identity
Project / Compute / Overview
Overview
Limit Summary
Instances
Used 0 of 10
VCPUs
Used 0 of 20
RAM
Used 0Bytes of 50GB
Floating IPs
Allocated 1 of 50
Security Groups
Used 1 of 10
Volumes
Used 0 of 10
Volume Storage
Used 0Bytes of 1000GB
Shares
False 0 of 50
Share Storage
False 0 of 1,000
Share Snapshots
False 0 of 50
Share Snapshots Storage
False 0 of 1,000
Share Networks
False 0 of 10
Usage Summary
Select a period of time to query its usage:
The date should be in YYYY-MM-DD format.
2019-12-02
to
2019-12-03
Submit
Active Instances: 0
Active RAM: 0Bytes
This Period's VCPU-Hours: 0.00
This Period's GB-Hours: 0.00
This Period's RAM-Hours: 0.00
Usage
Download CSV Summary
Instance Name
VCPUs
Disk
RAM
Time since created



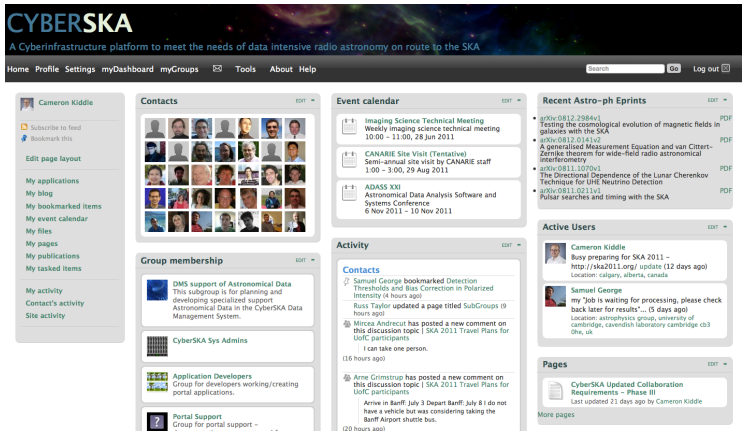
egi-datahub-admins	EGI DataHub admins	Closed	Active	Member Owner
egi-datahub-members	EGI DataHub users having access to EGI Foundation space - Closed group.	Closed	Active	Member Owner
egi-datahub-sentinel	EGI DataHub users allowed to access Sentinel Data. Open access.	Open	Active	Member Owner
eosc-xgus-supporters	Supporters in the EOSC xGUS	Closed	Active	Member Owner
ggus-supporters	Users that should be GGUS supporters.	Closed	Active	Member Owner
IDIA-outreach	Group for testing of user setup for federated resources	Open	Active	Member Owner
MeerKAT-Mightee	The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE) is a MeerKAT Large Survey Project to create deep images of the extragalactic sky to explore the cosmic evolution of galaxies. (https://idia.ac.za/mightee/)	Closed	Active	Member Owner
OnlyforTest	creation test	Open	Active	Member Owner
PRIMAGE_Project	Group of users from the PRIMAGE EU Project	Closed	Active	Member Owner
UKLSST	LSST:UK Collaboration	Closed	Active	Member Owner
UKLSST:dev	LSST:UK Development	Closed	Active	Member Owner
vo.notebooks.egi.eu	VO for notebooks.egi.eu	Closed	Active	Member Owner

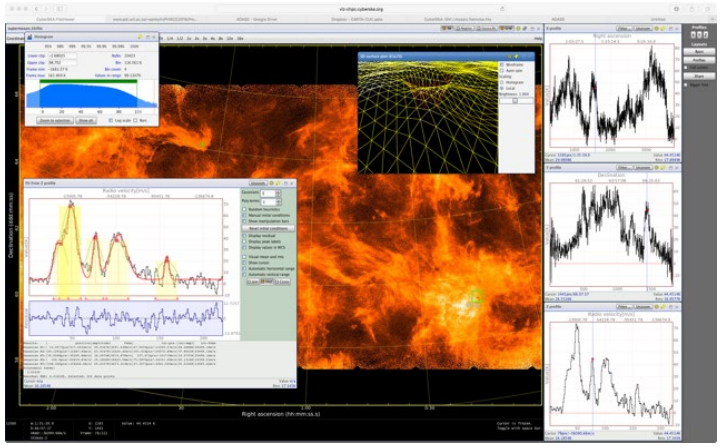




IDIA Gateway Planning

- Directors of IDIA have experience with implementing and running SGs
  - CyberSKA Scence Gateway ~15 years ago

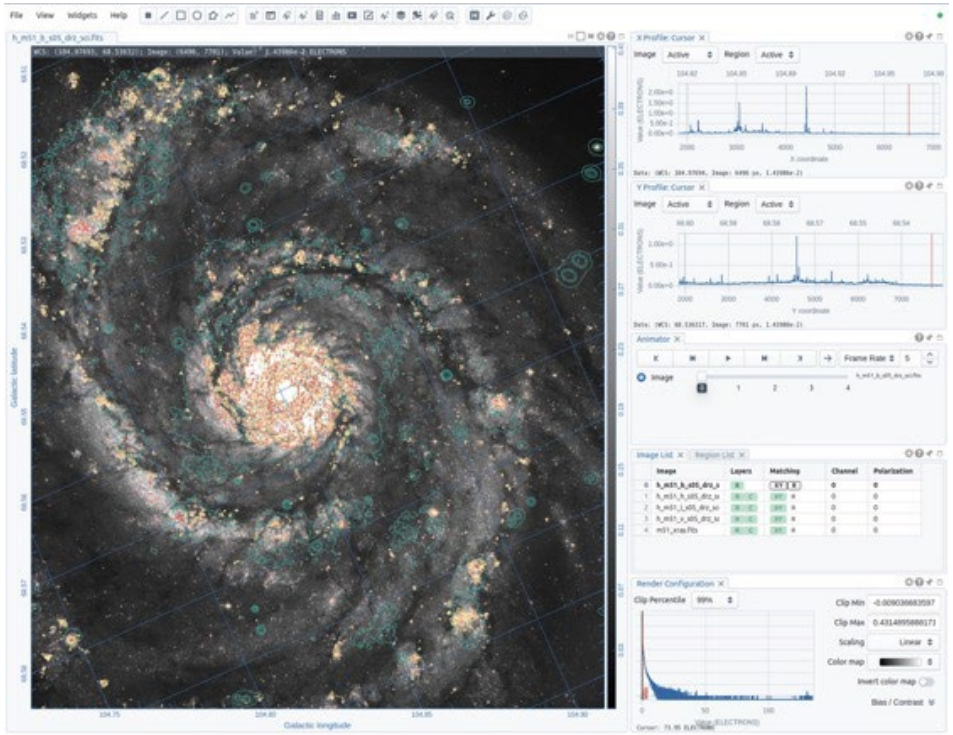




- Simmonds involved in other SGs for a range of Science and Engineering domains over last 20 years

IDIA Gateway Planning : Users

- IDIA Astronomers
- Hackathon / data-science workshop participants
- Accessors of Public Data Releases



- Also need to keep supporting non-Astronomy users of the ilifu resources so want it to be easy to customize views to different communities

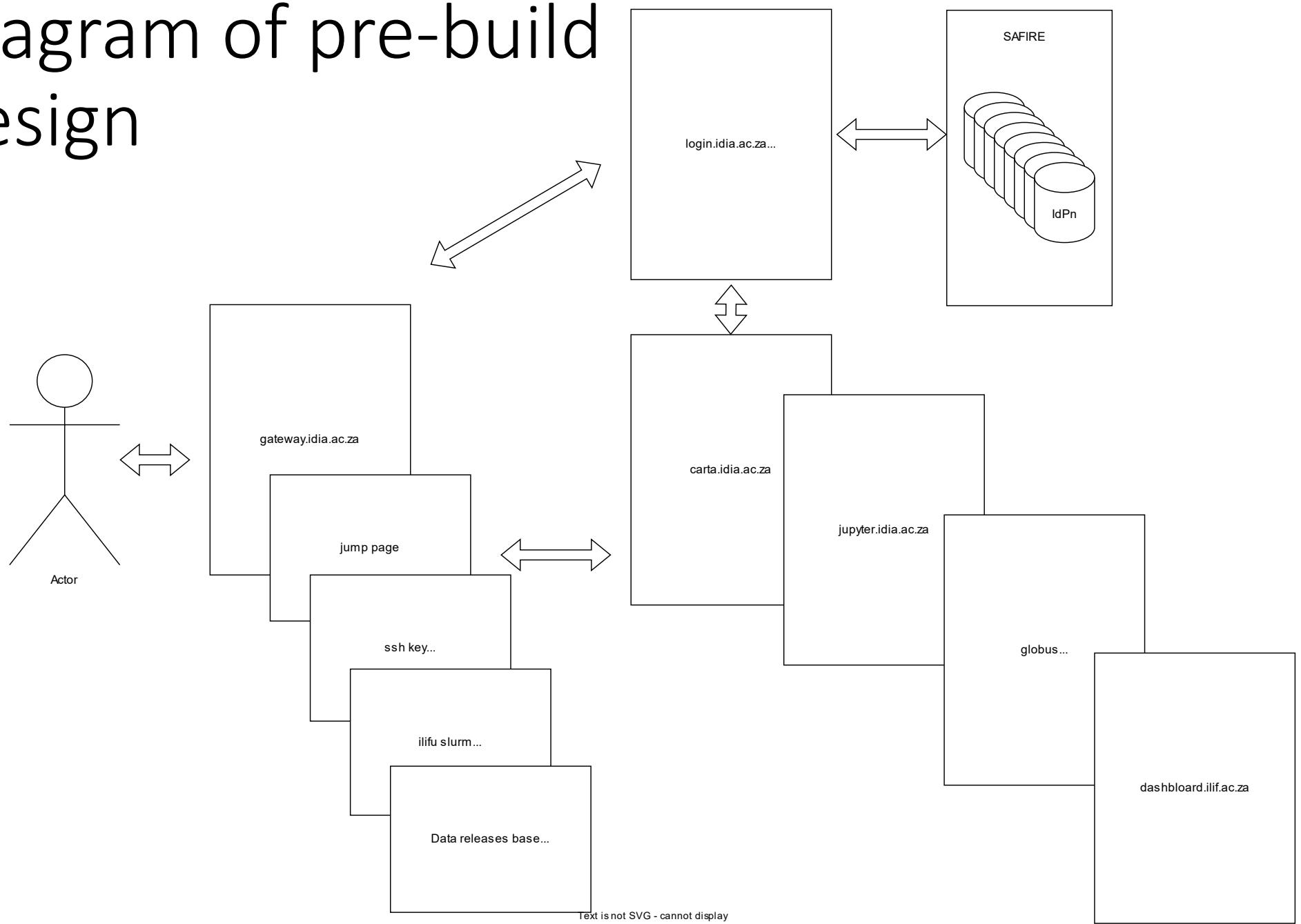
IDIA Gateway High Level Requirements

- System level
  - Limit that amount of code IDIA needs to write and maintain
  - Minimize the amount of effort required to maintain user accounts
  - Facilitate federation with other resources

- Platform level
  - Provide SSO access to best-in-class user applications

- Application level
  - Notebooks, Data Visualization, File Access, WAN Data Transfer, Data Delivery, Account Management

Diagram of pre-build design

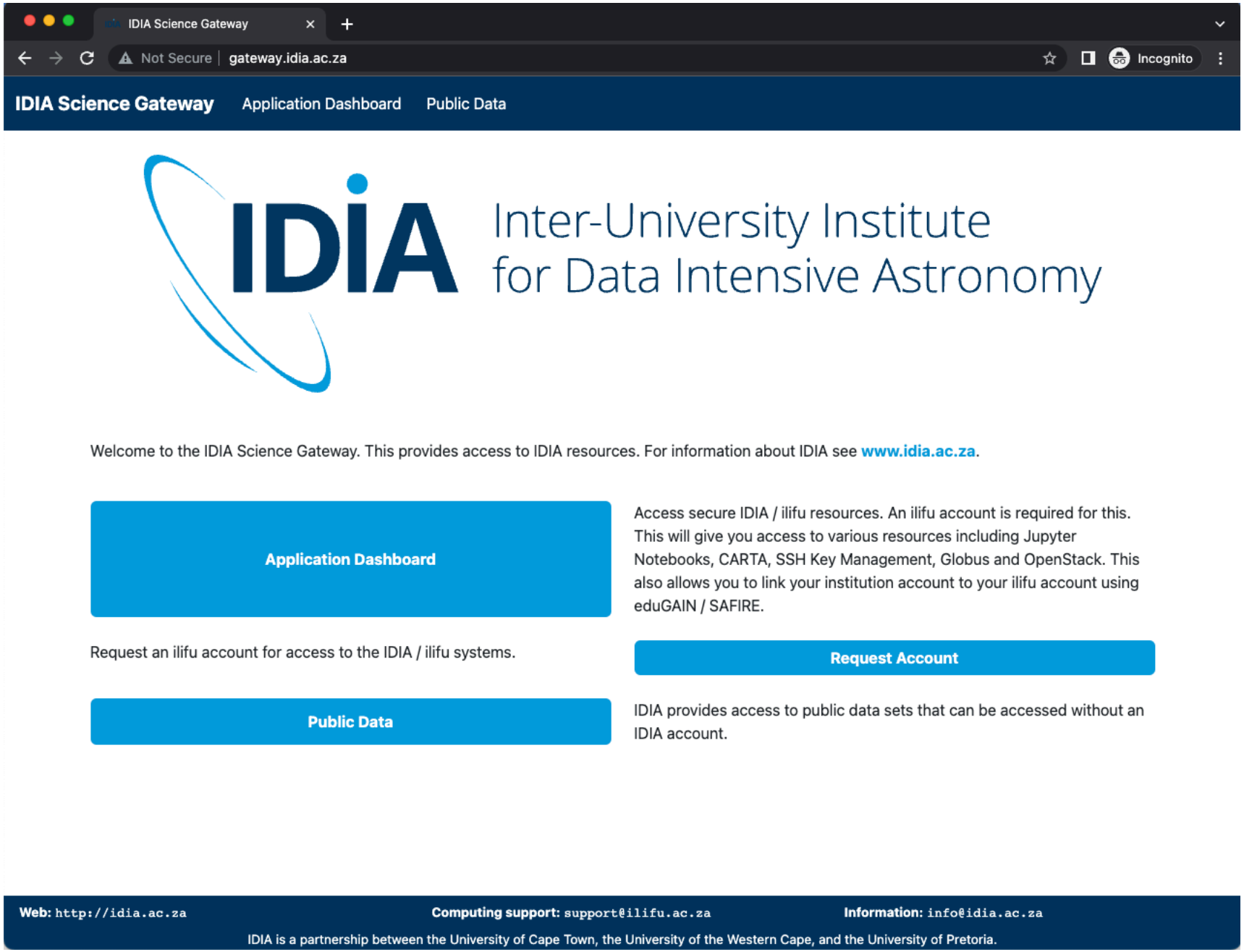


Text is not SVG - cannot display

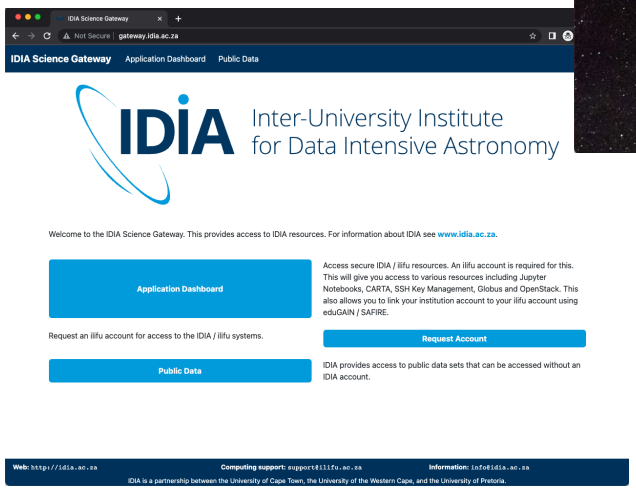
Technologies adopted

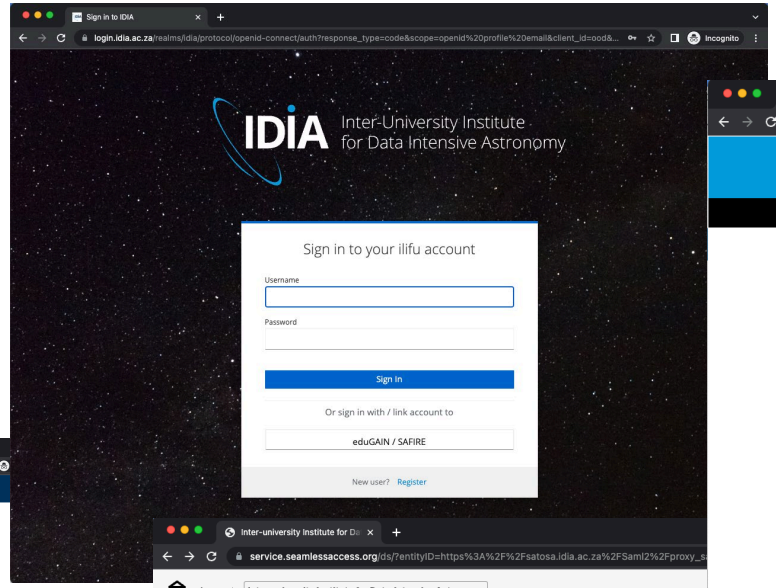
- Main Portal
  - Django
    - FrontPage
    - Public Data (using Wagtail CMS)
    - SSH Public Key management
    - Usage information
    - Initial container listing (management still done at CLI level for now)
  - Open OnDemand
    - Launch site for authenticated tools
- Other tools
  - Keycloak – AAI : Provides access to SAFIRE / EduGain
  - Jupyter – Notebooks
  - CARTA – Remote Visualization
  - Globus – WAN file transfer
  - OpenStack Dashboard – IaaS compute and Storage control

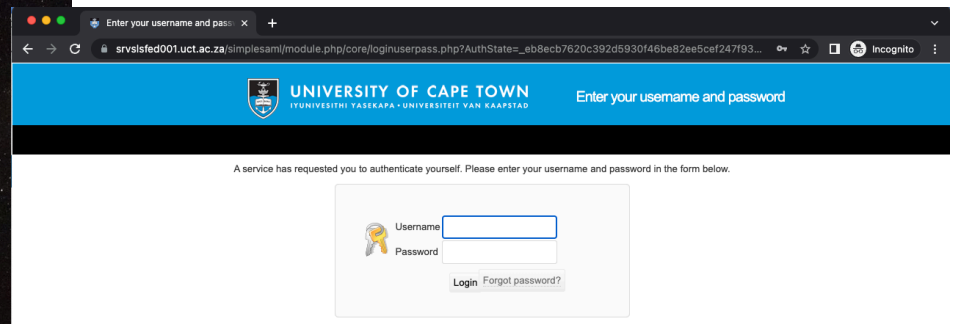
Front page

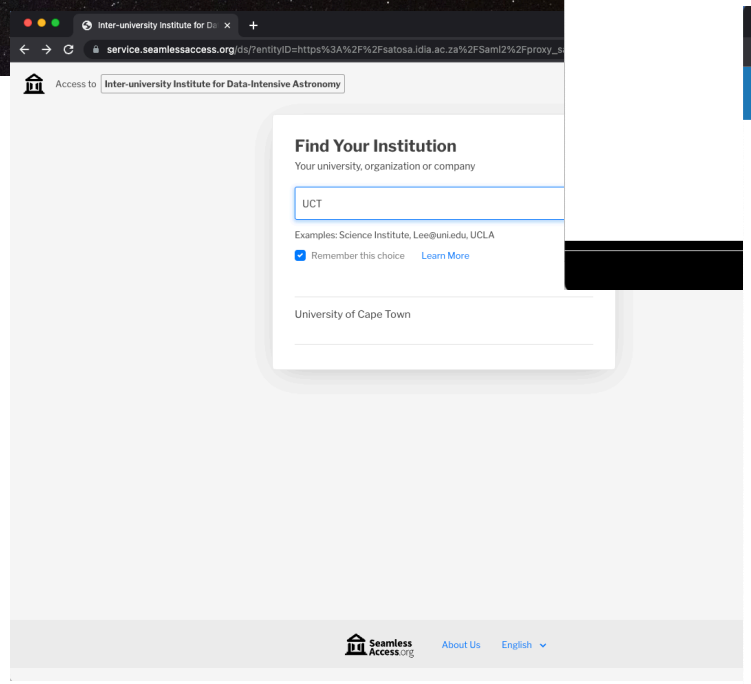

IDIA Science Gateway
Not Secure | gateway.idia.ac.za
Incognito
IDIA Science Gateway
Application Dashboard
Public Data
IDIA
Inter-University Institute
for Data Intensive Astronomy
Welcome to the IDIA Science Gateway. This provides access to IDIA resources. For information about IDIA see www.idia.ac.za.
Application Dashboard
Access secure IDIA / ilifu resources. An ilifu account is required for this. This will give you access to various resources including Jupyter Notebooks, CARTA, SSH Key Management, Globus and OpenStack. This also allows you to link your institution account to your ilifu account using eduGAIN / SAFIRE.
Request an ilifu account for access to the IDIA / ilifu systems.
Request Account
Public Data
IDIA provides access to public data sets that can be accessed without an IDIA account.
Web: http://idia.ac.za
Computing support: support@ilifu.ac.za
Information: info@idia.ac.za
IDIA is a partnership between the University of Cape Town, the University of the Western Cape, and the University of Pretoria.

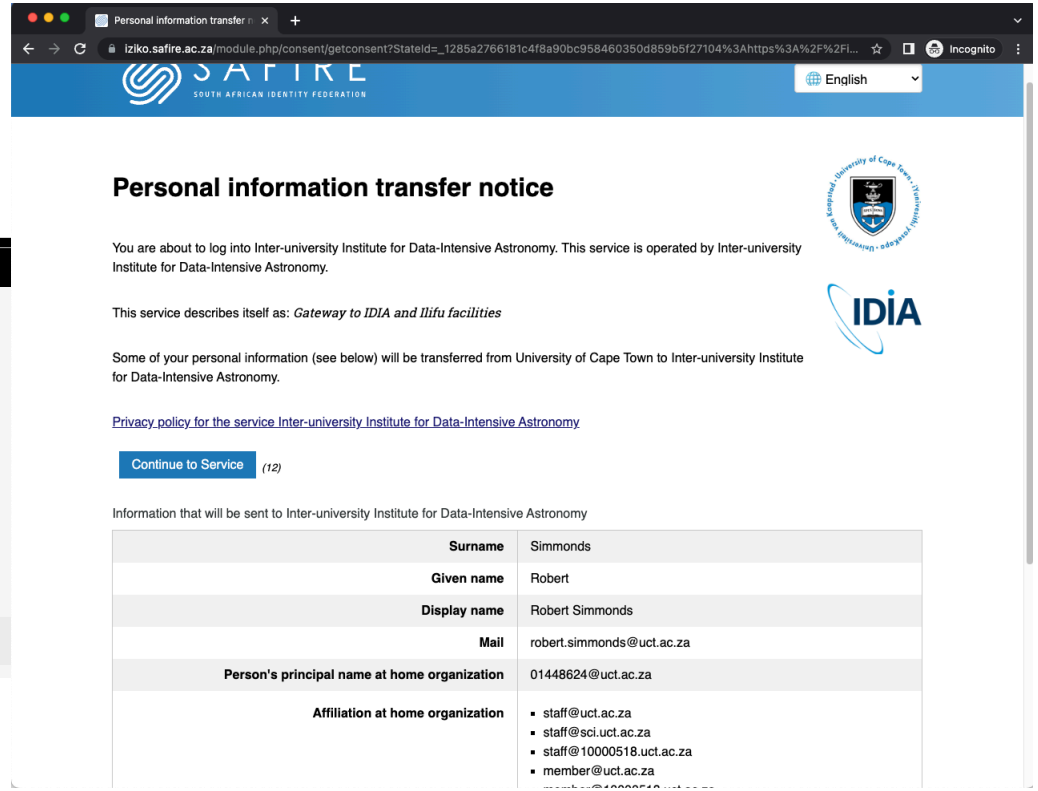
Login


IDIA Science Gateway
Application Dashboard
Public Data
IDIA
Inter-University Institute
for Data Intensive Astronomy
Welcome to the IDIA Science Gateway. This provides access to IDIA resources. For information about IDIA see www.idia.ac.za.
Application Dashboard
Access secure IDIA / ilifu resources. An ilifu account is required for this. This will give you access to various resources including Jupyter Notebooks, CARTA, SSH Key Management, Globus and OpenStack. This also allows you to link your institution account to your ilifu account using eduGAIN / SAFIRE.
Request an ilifu account for access to the IDIA / ilifu systems.
Request Account
Public Data
IDIA provides access to public data sets that can be accessed without an IDIA account.
Web: http://idia.ac.za
Computing support: support@ilifu.ac.za
Information: info@idia.ac.za
IDIA is a partnership between the University of Cape Town, the University of the Western Cape, and the University of Pretoria.

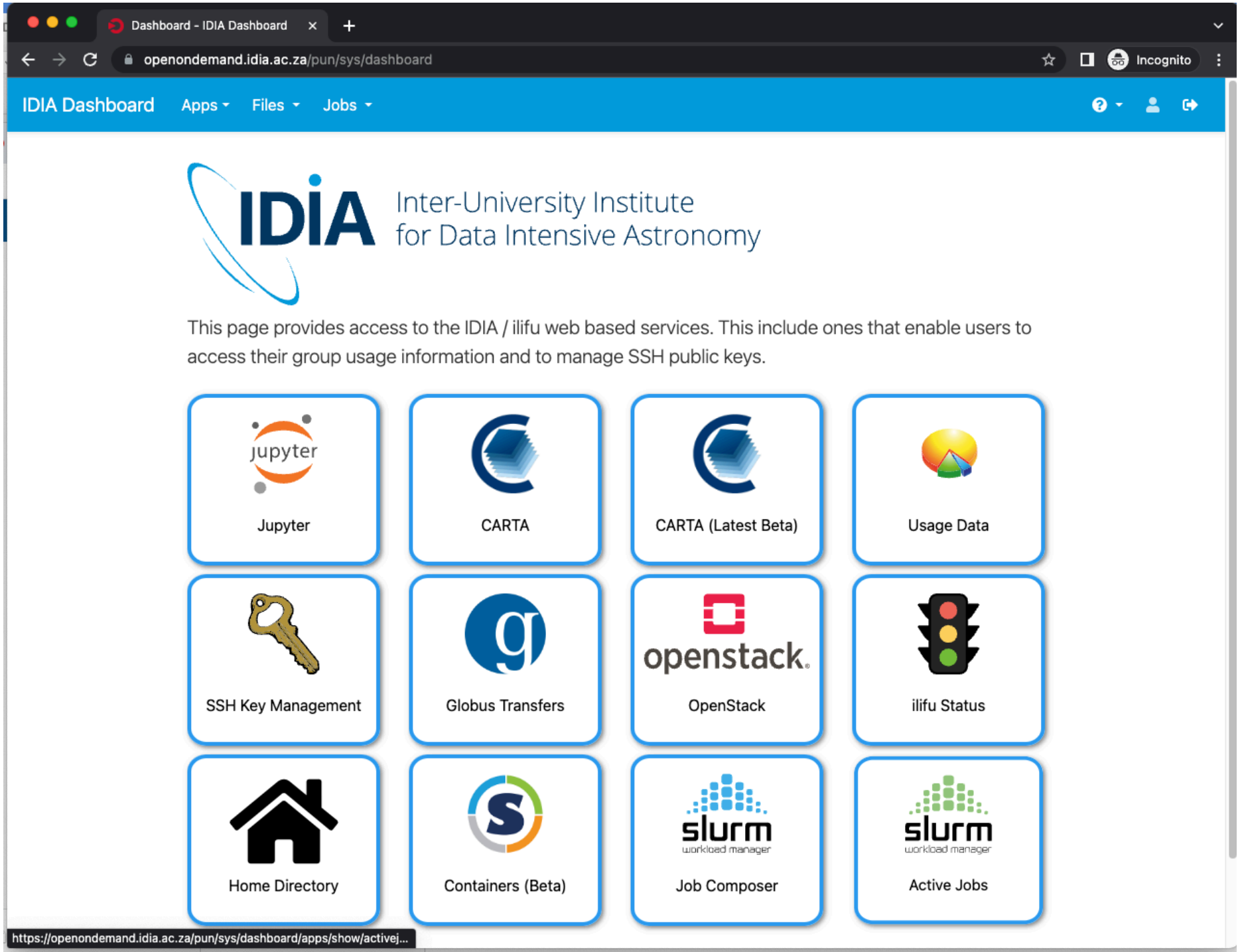

Sign in to IDIA
login.idia.ac.za/realms/idia/protocol/openid-connect/auth?response_type=code&scope=openid%20profile%20email&client_id=ood&...
Incognito
IDIA
Inter-University Institute
for Data Intensive Astronomy
Sign in to your ilifu account
Username
Password
Sign In
Or sign in with / link account to
eduGAIN / SAFIRE
New user? Register


Enter your username and pass
srvslsfed001.uct.ac.za/simplesaml/module.php/core/loginuserpass.php?AuthState=_eb8ecb7620c392d5930f46be82ee5cef247f93...
Incognito
UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD
Enter your username and password
A service has requested you to authenticate yourself. Please enter your username and password in the form below.
Username
Password
Login
Forgot password?

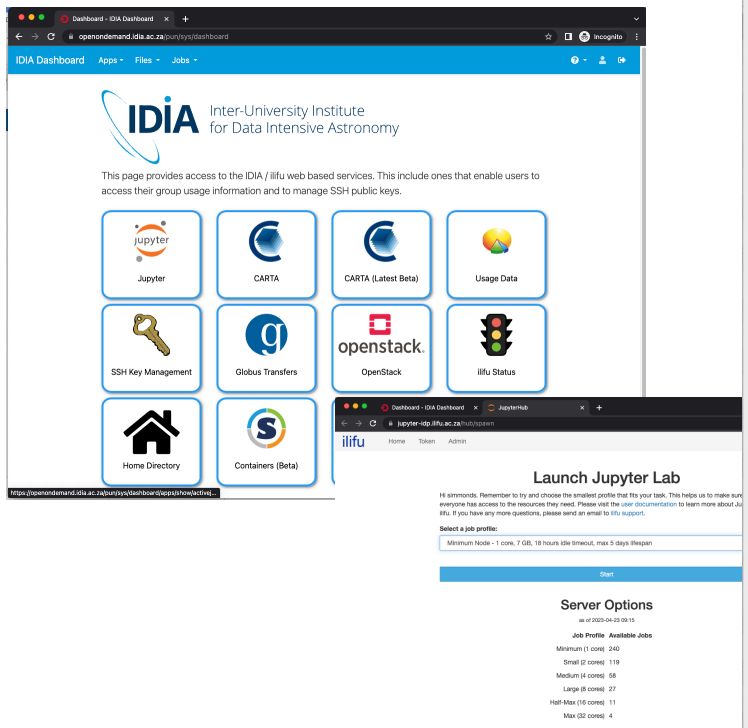

Inter-university Institute for Da
service.seamlessaccess.org/ds/?entityID=https%3A%2F%2Fsatosa.idia.ac.za%2FSaml2%2Fproxy_s
Access to Inter-university Institute for Data-Intensive Astronomy
Find Your Institution
Your university, organization or company
UCT
Examples: Science Institute, Lee@uni.edu, UCLA
Remember this choice
Learn More
University of Cape Town
Seamless Access.org
About Us
English

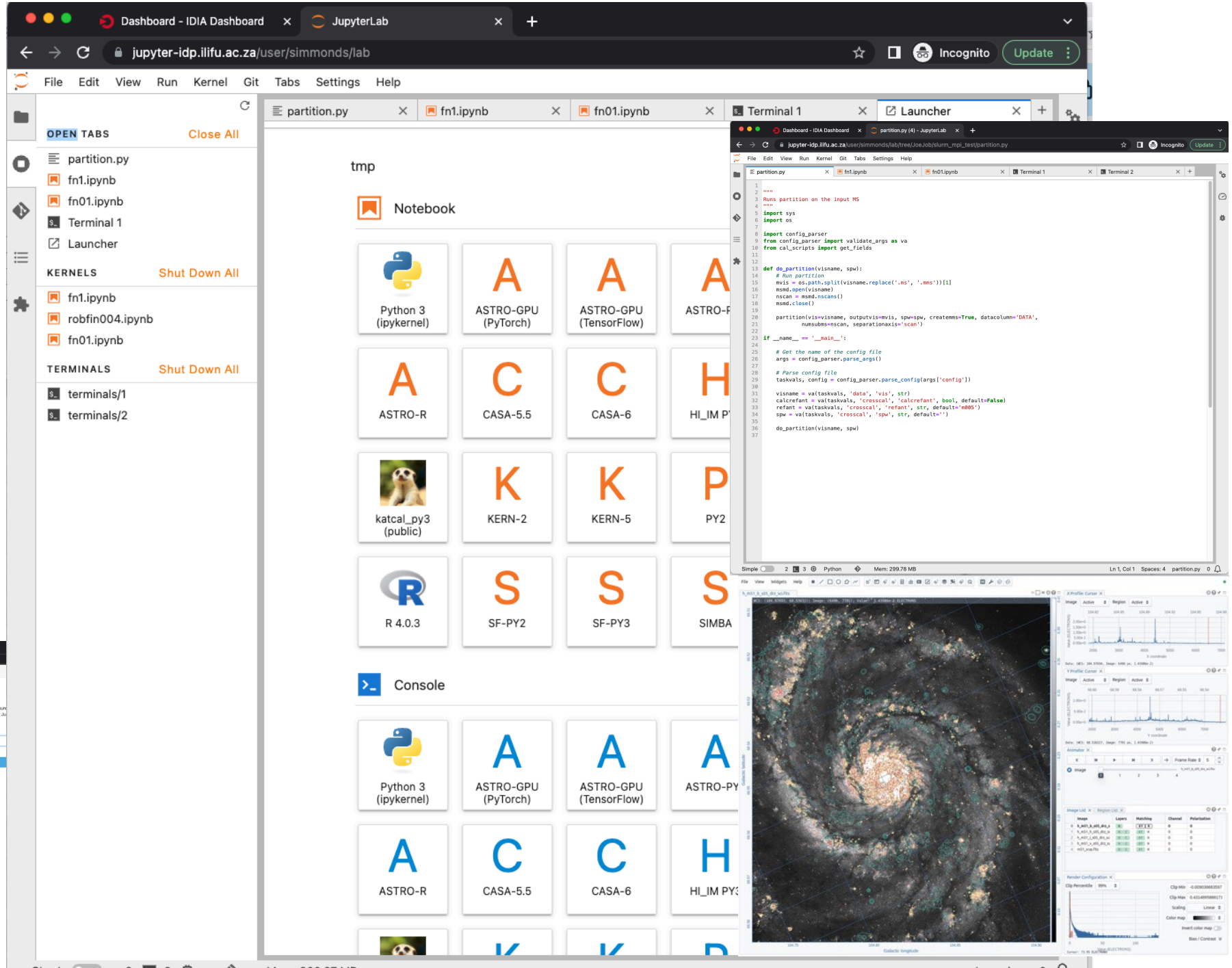

Personal information transfer n
iziko.safire.za/module.php/consent/getconsent?StateId=_1285a2766181c4f8a90bc958460350d859b5f27104%3Ahttps%3A%2F%2Fi...
Incognito
SAFIRE
SOUTH AFRICAN IDENTITY FEDERATION
English
Personal information transfer notice
You are about to log into Inter-university Institute for Data-Intensive Astronomy. This service is operated by Inter-university Institute for Data-Intensive Astronomy.
This service describes itself as: Gateway to IDIA and Ilifu facilities
Some of your personal information (see below) will be transferred from University of Cape Town to Inter-university Institute for Data-Intensive Astronomy.
Privacy policy for the service Inter-university Institute for Data-Intensive Astronomy
Continue to Service (12)
Information that will be sent to Inter-university Institute for Data-Intensive Astronomy
Surname Simmonds
Given name Robert
Display name Robert Simmonds
Mail robert.simmonds@uct.ac.za
Person's principal name at home organization 01448624@uct.ac.za
Affiliation at home organization
staff@uct.ac.za
staff@sci.uct.ac.za
staff@10000518.uct.ac.za
member@uct.ac.za

Launch Page

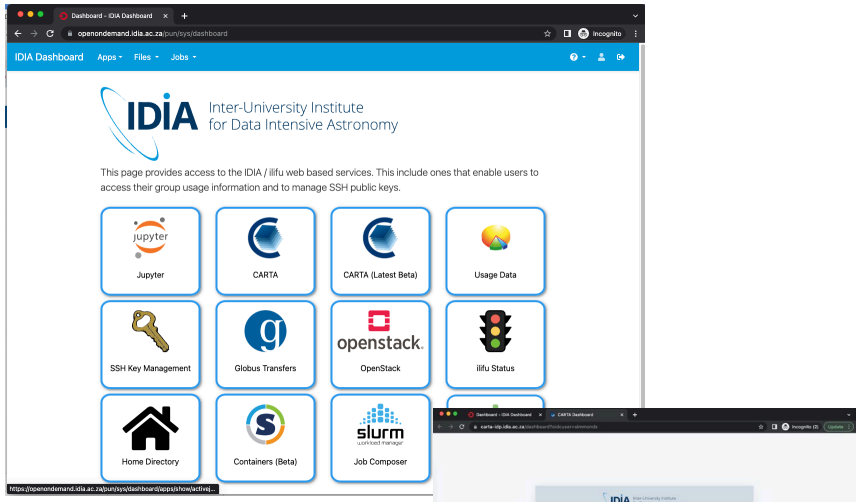

Dashboard - IDIA Dashboard
openondemand.idia.ac.za/pun/sys/dashboard
Incognito
IDIA Dashboard
Apps
Files
Jobs
IDIA
Inter-University Institute
for Data Intensive Astronomy
This page provides access to the IDIA / ilifu web based services. This include ones that enable users to access their group usage information and to manage SSH public keys.
Jupyter
CARTA
CARTA (Latest Beta)
Usage Data
SSH Key Management
Globus Transfers
OpenStack
ilifu Status
Home Directory
Containers (Beta)
Job Composer
Active Jobs
https://openondemand.idia.ac.za/pun/sys/dashboard/apps/show/activej...

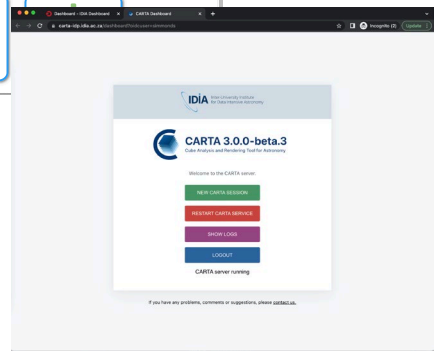
Jupyter

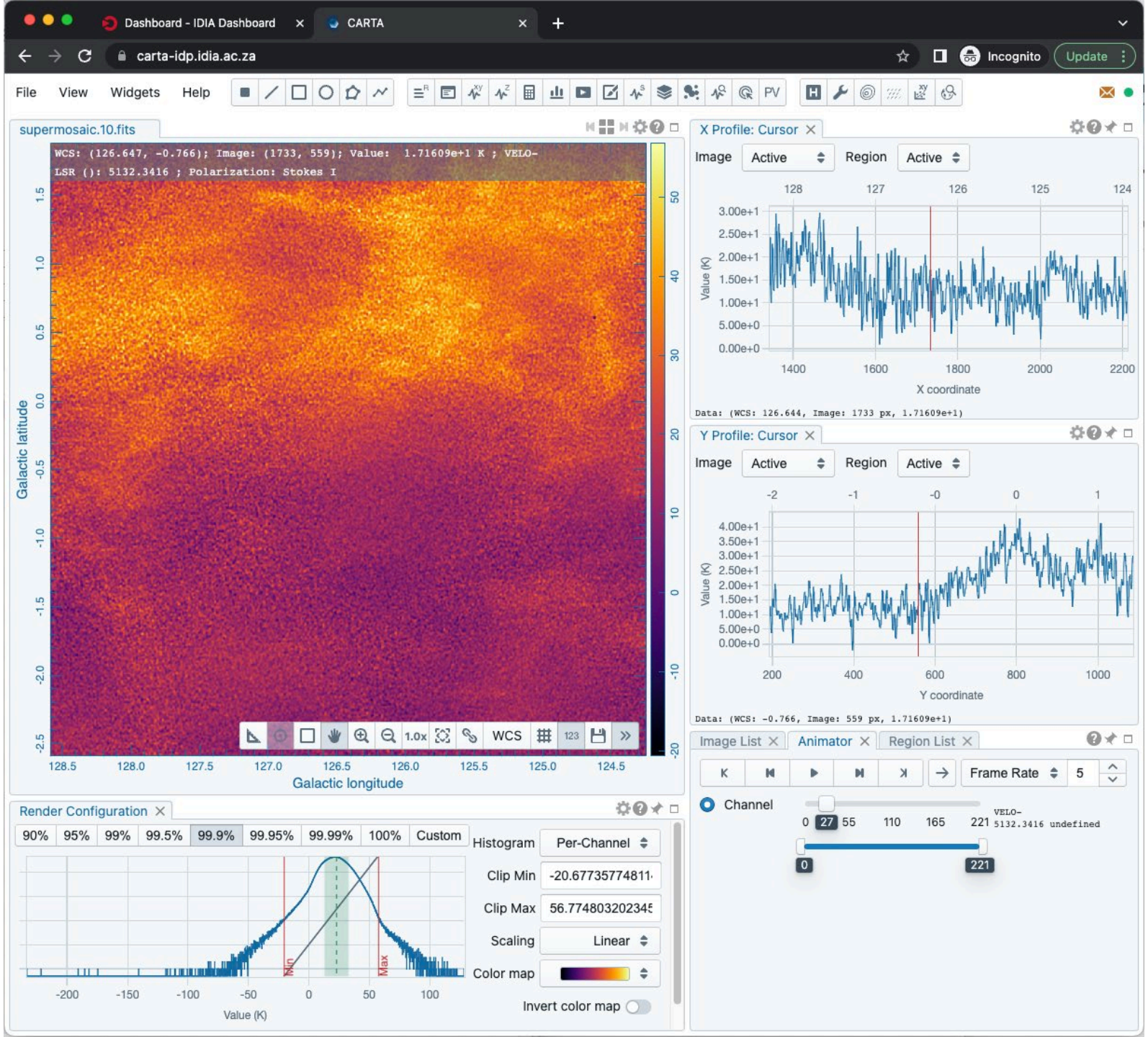

IDIA
Inter-University Institute
for Data Intensive Astronomy
This page provides access to the IDIA / ilifu web based services. This include ones that enable users to access their group usage information and to manage SSH public keys.
Jupyter
CARTA
CARTA (Latest Beta)
Usage Data
SSH Key Management
Globus Transfers
OpenStack
ilifu Status
Home Directory
Containers (Beta)
Launch Jupyter Lab
Server Options


Dashboard - IDIA Dashboard
JupyterLab
jupyter-idp.ilifu.ac.za/user/simmonds/lab
File Edit View Run Kernel Git Tabs Settings Help
OPEN TABS
Close All
partition.py
fn1.ipynb
fn01.ipynb
Terminal 1
Launcher
KERNELS
Shut Down All
robfin004.ipynb
TERMINALS
terminals/1
terminals/2
tmp
Notebook
Python 3 (ipykernel)
ASTRO-GPU (PyTorch)
ASTRO-GPU (TensorFlow)
ASTRO-R
CASA-5.5
CASA-6
katcal_py3 (public)
KERN-2
KERN-5
PY2
R 4.0.3
SF-PY2
SF-PY3
SIMBA
Console

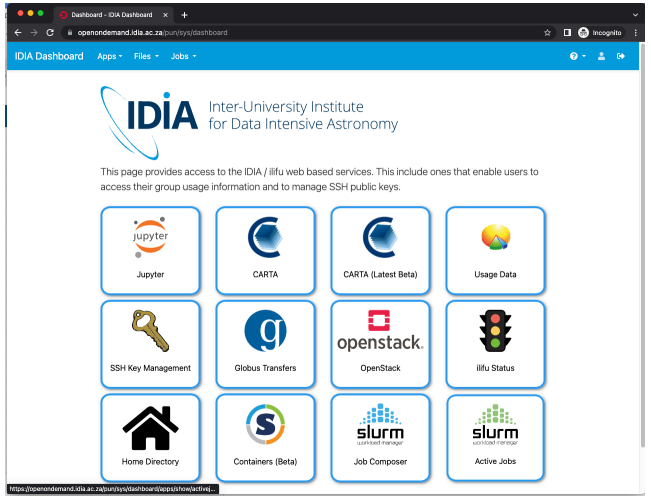
CARTA Visualization

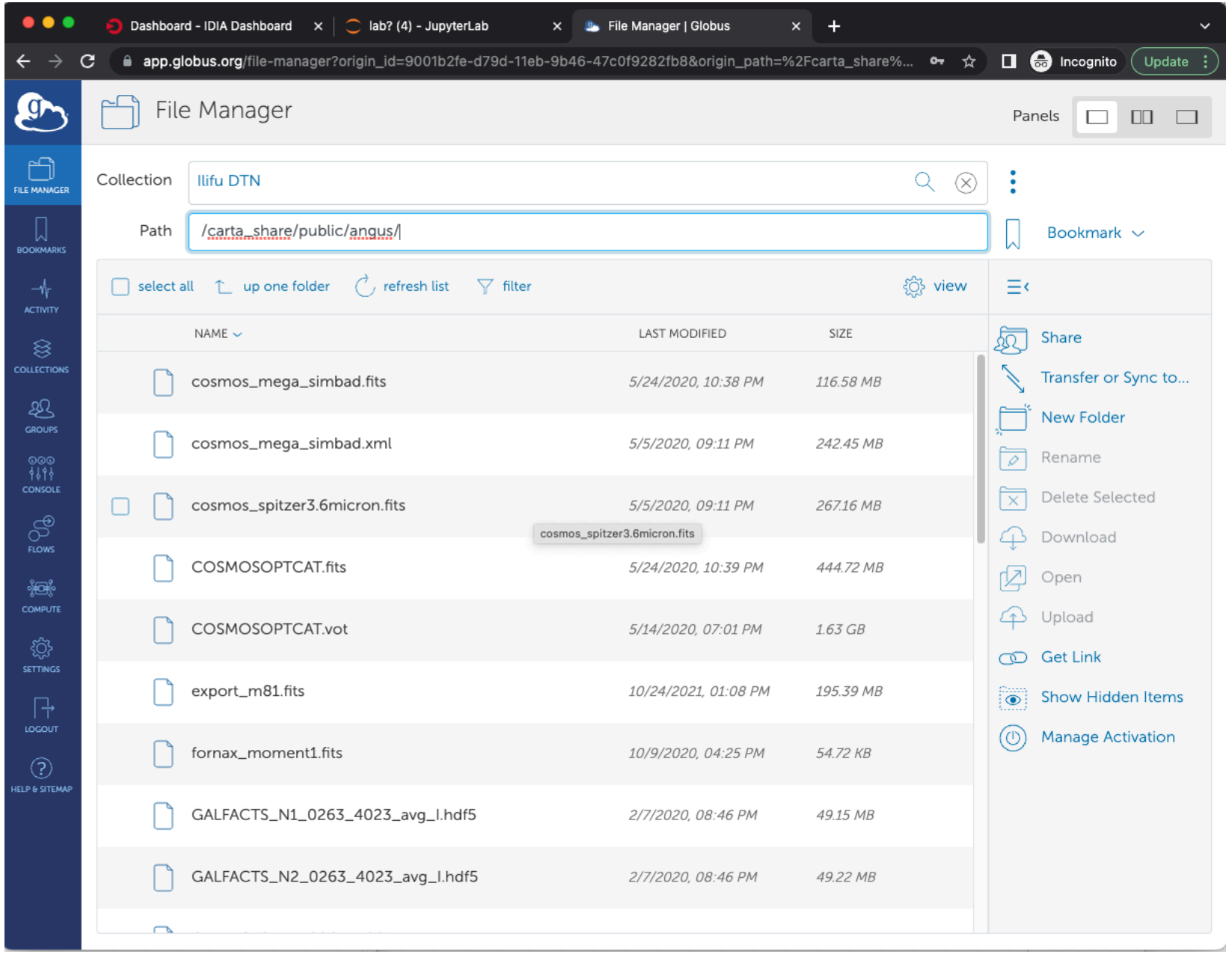




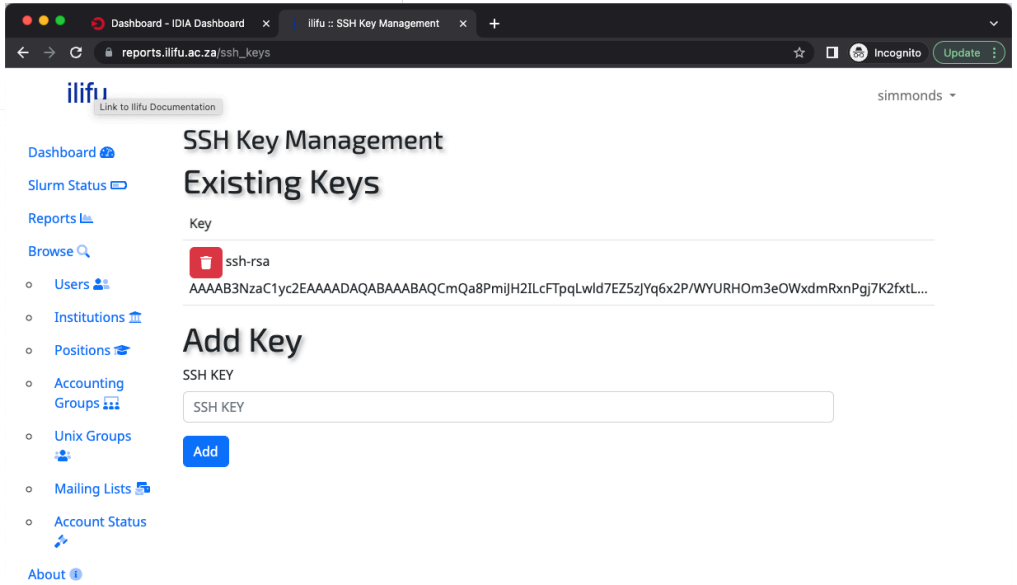


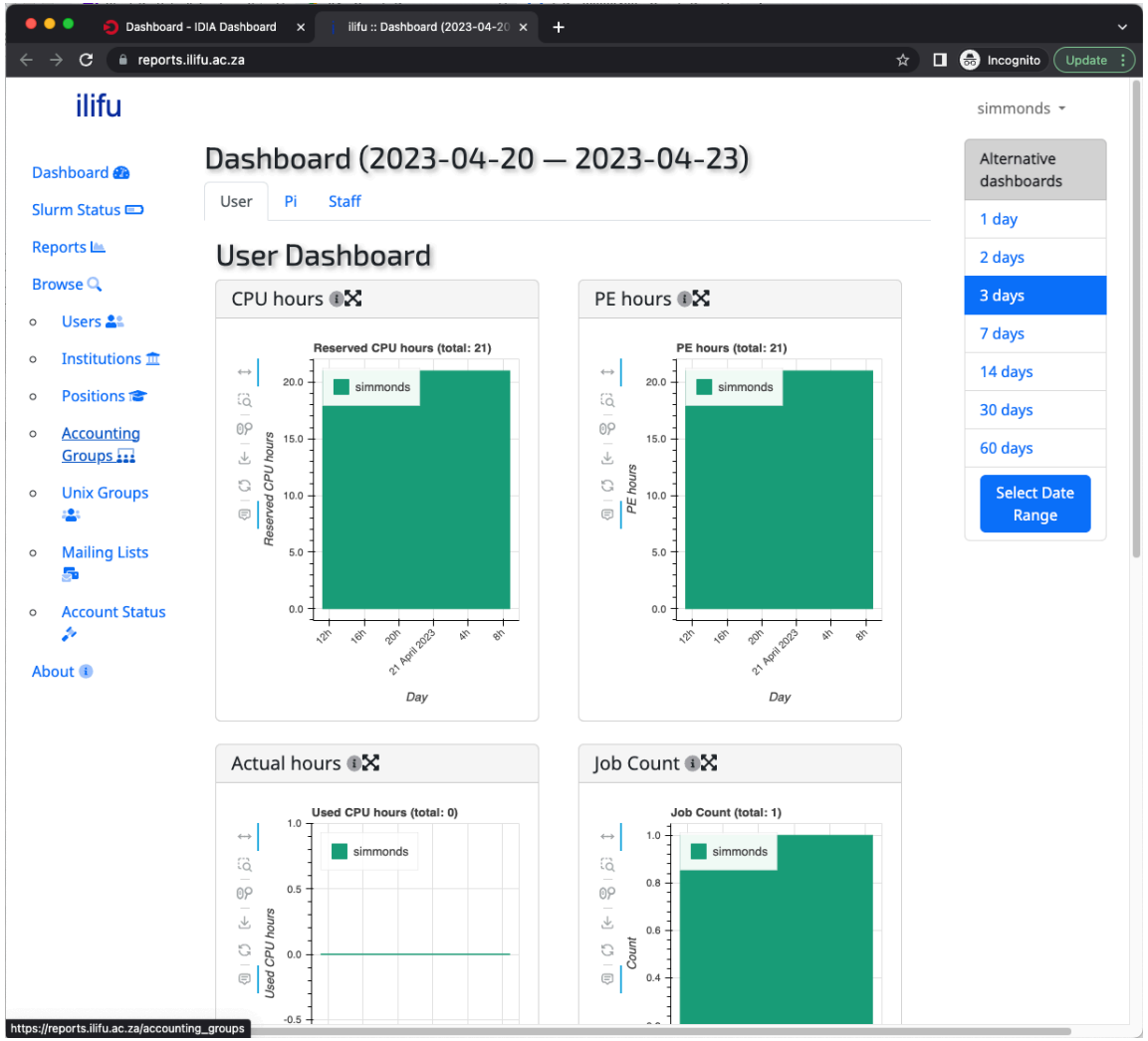
Globus File Transfer

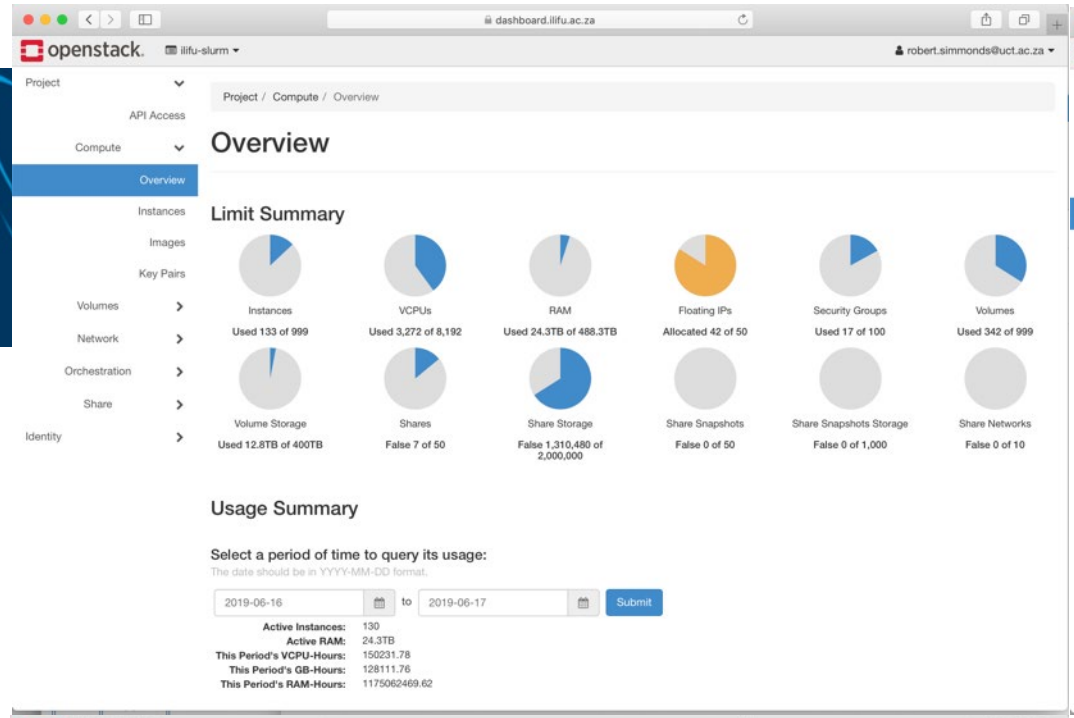


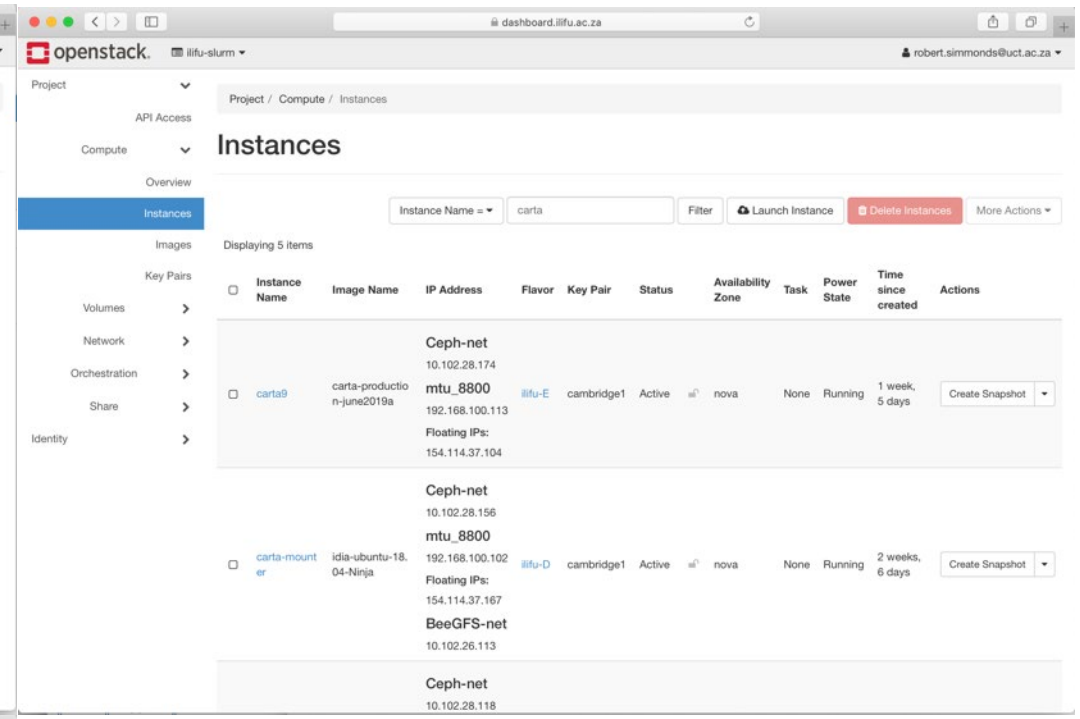


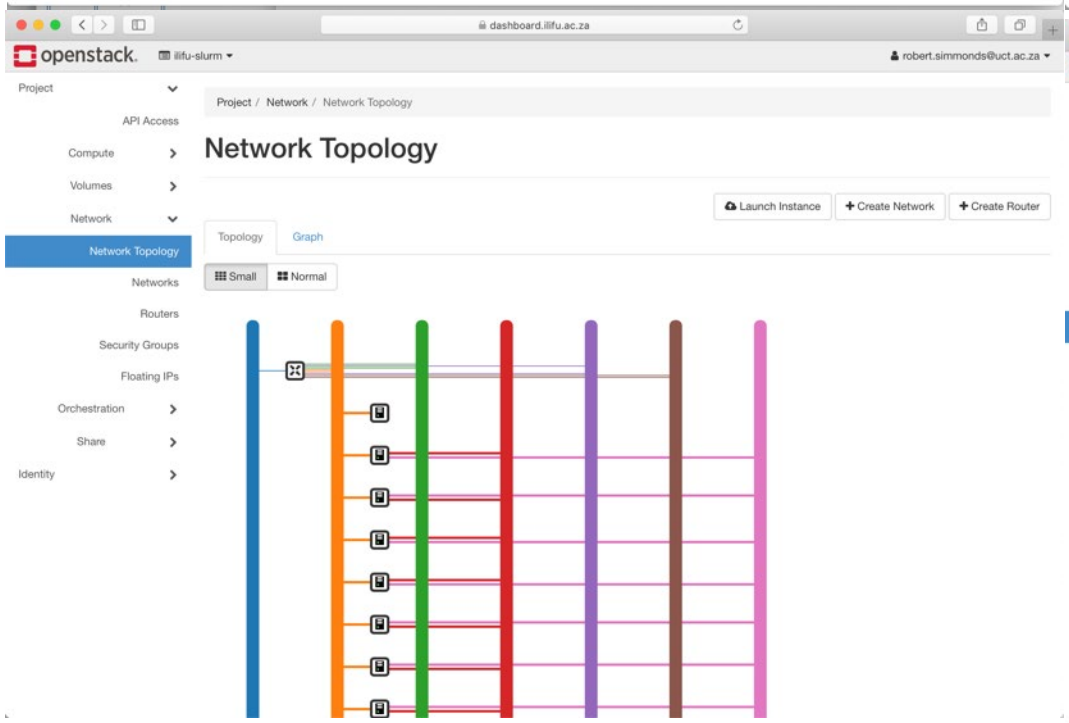
Key Management and Usage Data

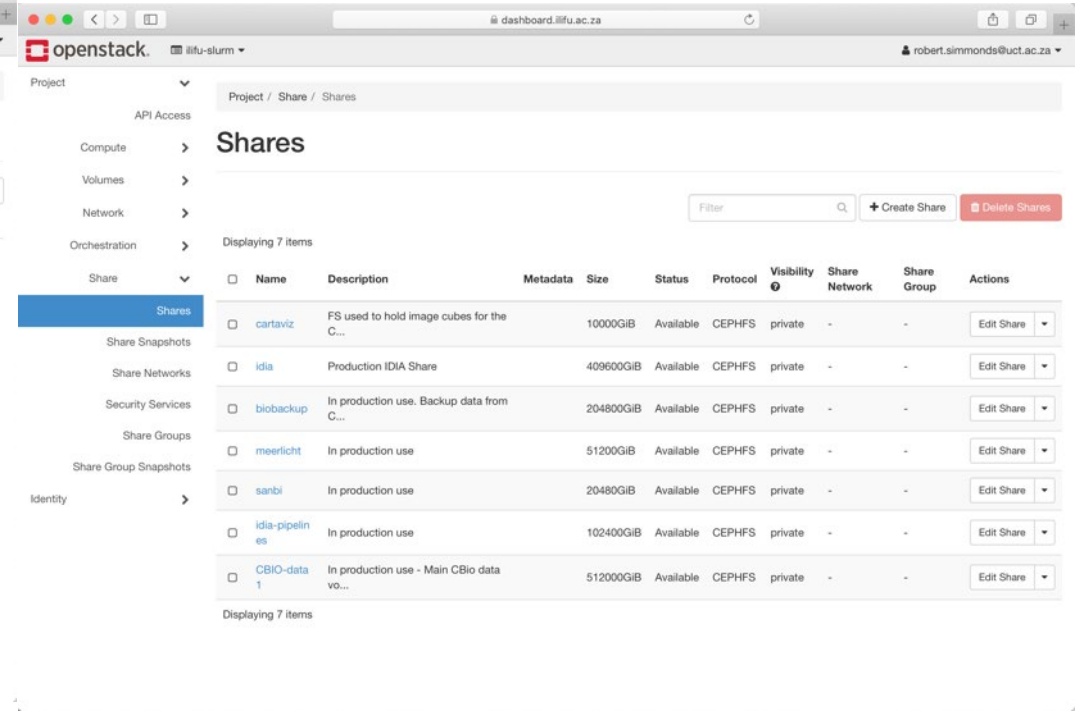





dashboard.ilifu.ac.za
openstack.
ilifu-slurm
robert.simmonds@uct.ac.za
Project
API Access
Compute
Overview
Instances
Images
Key Pairs
Volumes
Network
Orchestration
Share
Identity
Project / Compute / Overview
Overview
Limit Summary
Instances
Used 133 of 999
VCPUs
Used 3,272 of 8,192
RAM
Used 24.3TB of 488.3TB
Floating IPs
Allocated 42 of 50
Security Groups
Used 17 of 100
Volumes
Used 342 of 999
Volume Storage
Used 12.8TB of 400TB
Shares
False 7 of 50
Share Storage
False 1,310,480 of 2,000,000
Share Snapshots
False 0 of 50
Share Snapshots Storage
False 0 of 1,000
Share Networks
False 0 of 10
Usage Summary
Select a period of time to query its usage:
The date should be in YYYY-MM-DD format.
2019-06-16
to
2019-06-17
Submit
Active Instances: 130
Active RAM: 24.3TB
This Period's VCPU-Hours: 150231.78
This Period's GB-Hours: 128111.76
This Period's RAM-Hours: 1175062469.62


dashboard.ilifu.ac.za
openstack.
ilifu-slurm
robert.simmonds@uct.ac.za
Project
API Access
Compute
Overview
Instances
Images
Key Pairs
Volumes
Network
Orchestration
Share
Identity
Project / Compute / Instances
Instances
Instance Name =
carta
Filter
Launch Instance
Delete Instances
More Actions
Displaying 5 items
Instance Name
Image Name
IP Address
Flavor
Key Pair
Status
Availability Zone
Task
Power State
Time since created
Actions
carta9
carta-production-june2019a
Ceph-net
10.102.28.174
mtu_8800
192.168.100.113
Floating IPs:
154.114.37.104
ilifu-E
cambridge1
Active
nova
None
Running
1 week, 5 days
Create Snapshot
carta-mounter
idia-ubuntu-18.04-Ninja
Ceph-net
10.102.28.156
mtu_8800
192.168.100.102
Floating IPs:
154.114.37.167
BeeGFS-net
10.102.26.113
ilifu-D
cambridge1
Active
nova
None
Running
2 weeks, 6 days
Create Snapshot
Ceph-net
10.102.28.118

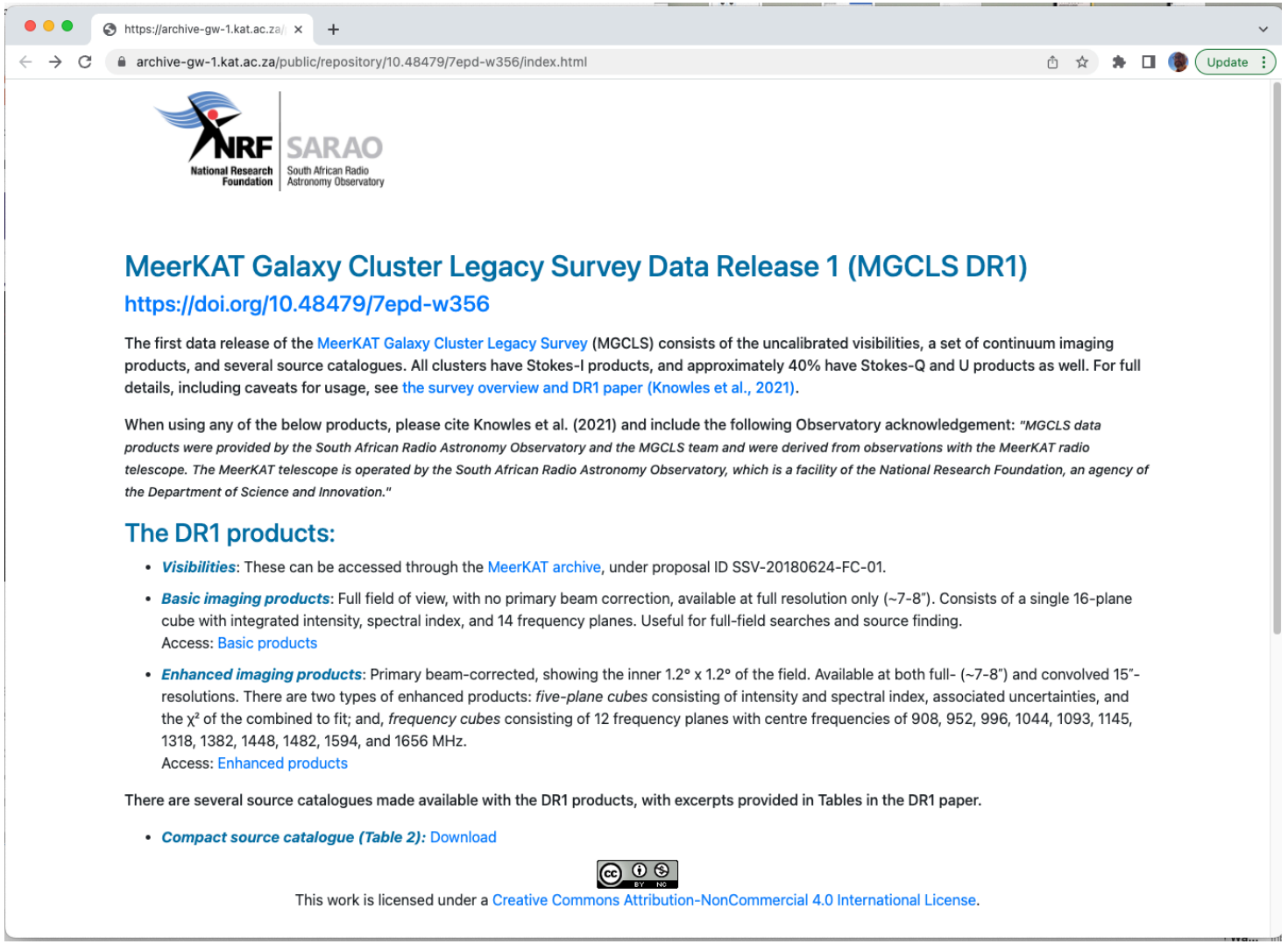

dashboard.ilifu.ac.za
openstack.
ilifu-slurm
robert.simmonds@uct.ac.za
Project
API Access
Compute
Volumes
Network
Network Topology
Networks
Routers
Security Groups
Floating IPs
Orchestration
Share
Identity
Project / Network / Network Topology
Network Topology
Launch Instance
Create Network
Create Router
Topology
Graph
Small
Normal


dashboard.ilifu.ac.za
openstack.
ilifu-slurm
robert.simmonds@uct.ac.za
Project
API Access
Compute
Volumes
Network
Orchestration
Share
Shares
Share Snapshots
Share Networks
Security Services
Share Groups
Share Group Snapshots
Identity
Project / Share / Shares
Shares
Filter
Create Share
Delete Shares
Displaying 7 items
Name
Description
Metadata
Size
Status
Protocol
Visibility
Share Network
Share Group
Actions
cartaviz
FS used to hold image cubes for the C...
10000GiB
Available
CEPHFS
private
Edit Share
idia
Production IDIA Share
409600GiB
Available
CEPHFS
private
Edit Share
biobackup
In production use. Backup data from C...
204800GiB
Available
CEPHFS
private
Edit Share
meerlicht
In production use
51200GiB
Available
CEPHFS
private
Edit Share
sanbi
In production use
20480GiB
Available
CEPHFS
private
Edit Share
idia-pipelines
In production use
102400GiB
Available
CEPHFS
private
Edit Share
CBIO-data1
In production use - Main CBio data vo...
512000GiB
Available
CEPHFS
private
Edit Share
Displaying 7 items

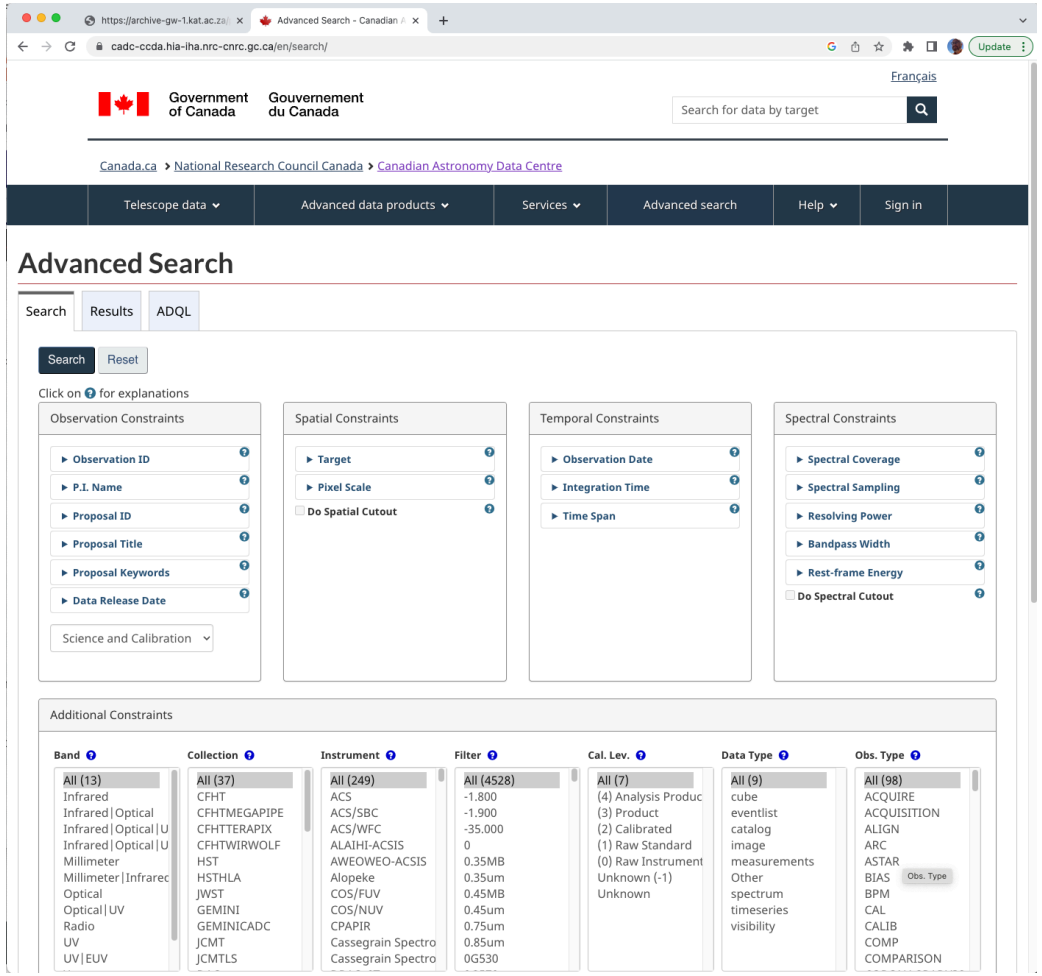
Data Presentation

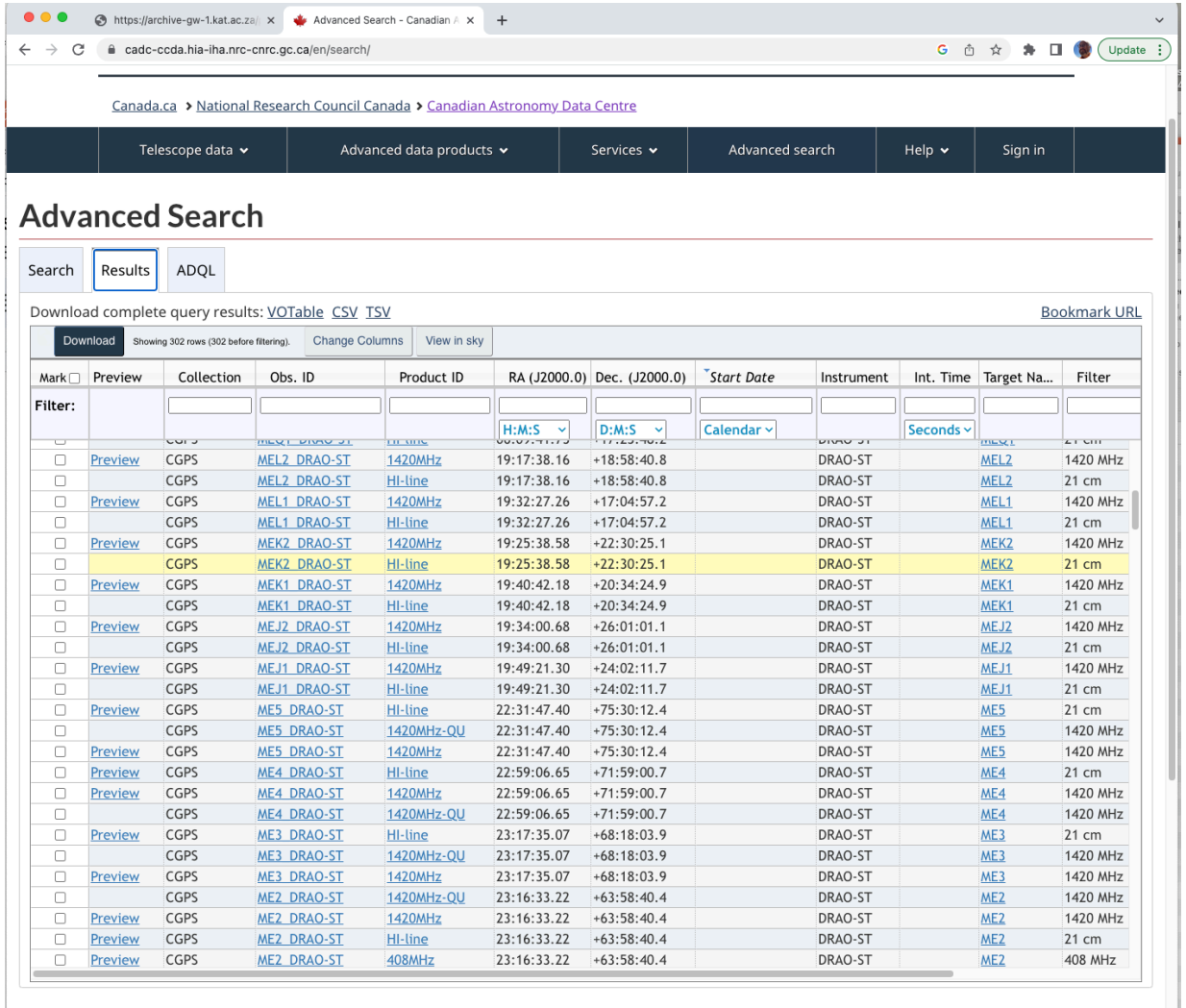
- IDIA has never intended to be a data archive
  - IDIA systems are not architected with this use case in mind
  - Also need to be careful of trying to take on activities that require additional support worker skills and effort
- Recently there have been requests to host some data releases
  - IDIA can do this, but would be good to work with partners such as NICIS to provide increased data resilience
- Will be creating focus group withing IDIA on what forms of Data Presentation and Access are most helpful

Data Views : Data Release (SARAO)



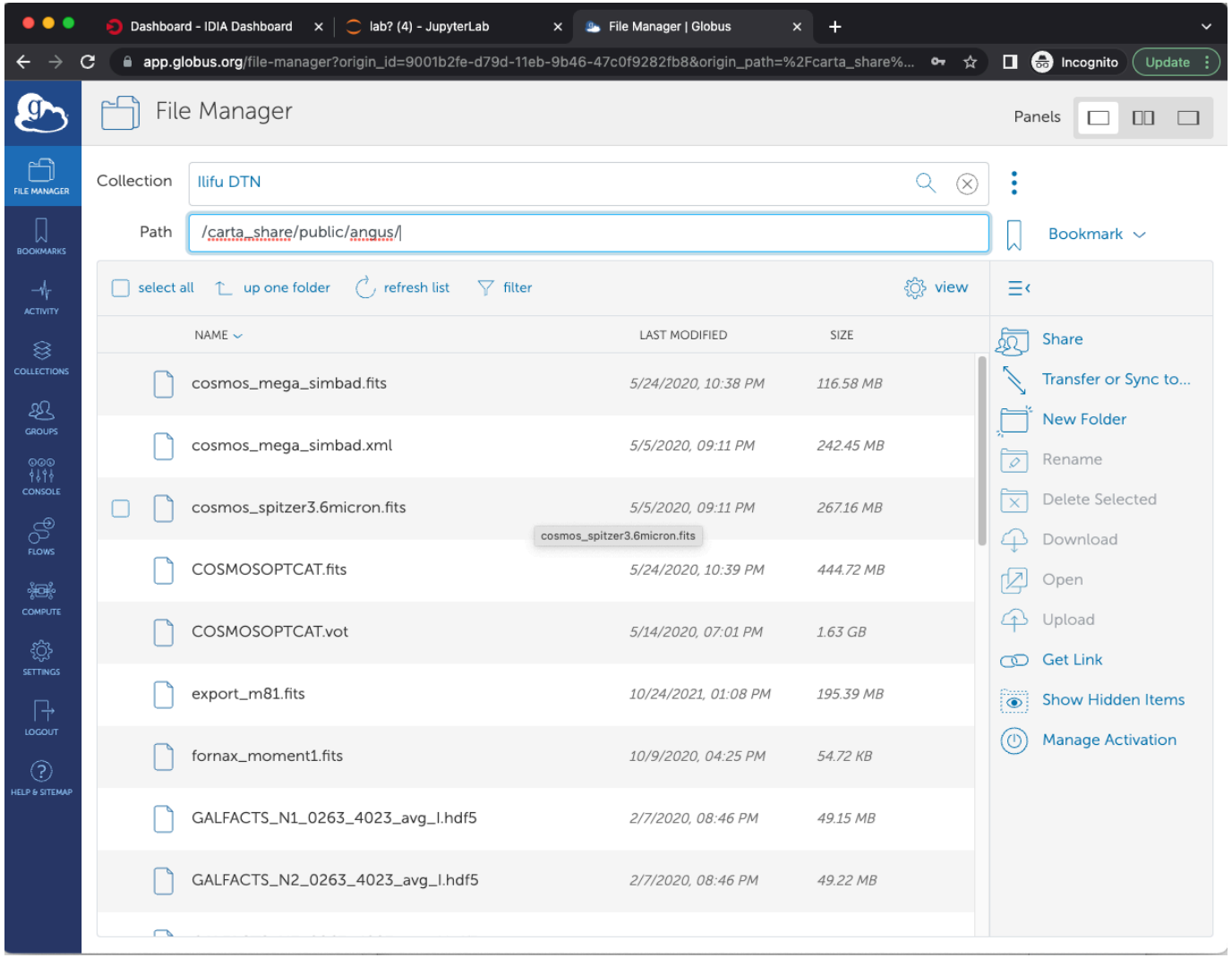
Data Views : Searchable index (CADC)



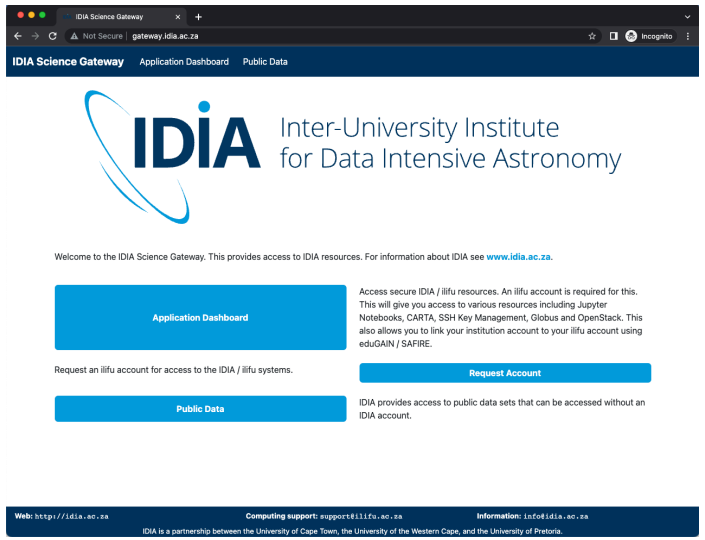


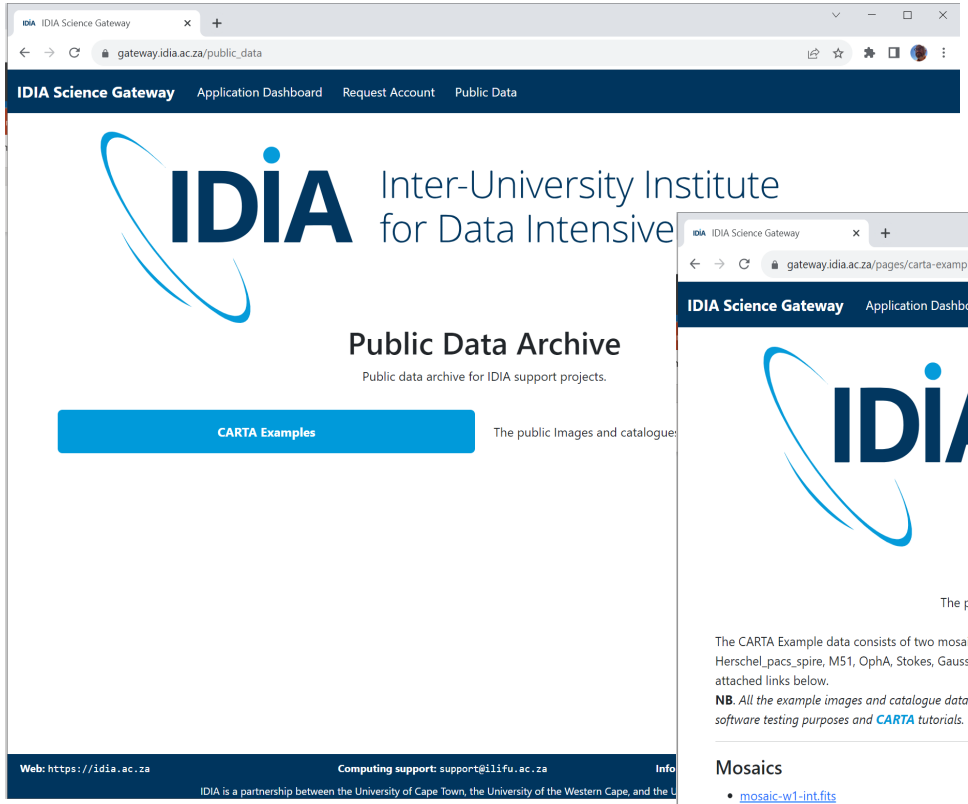
Simple File Views

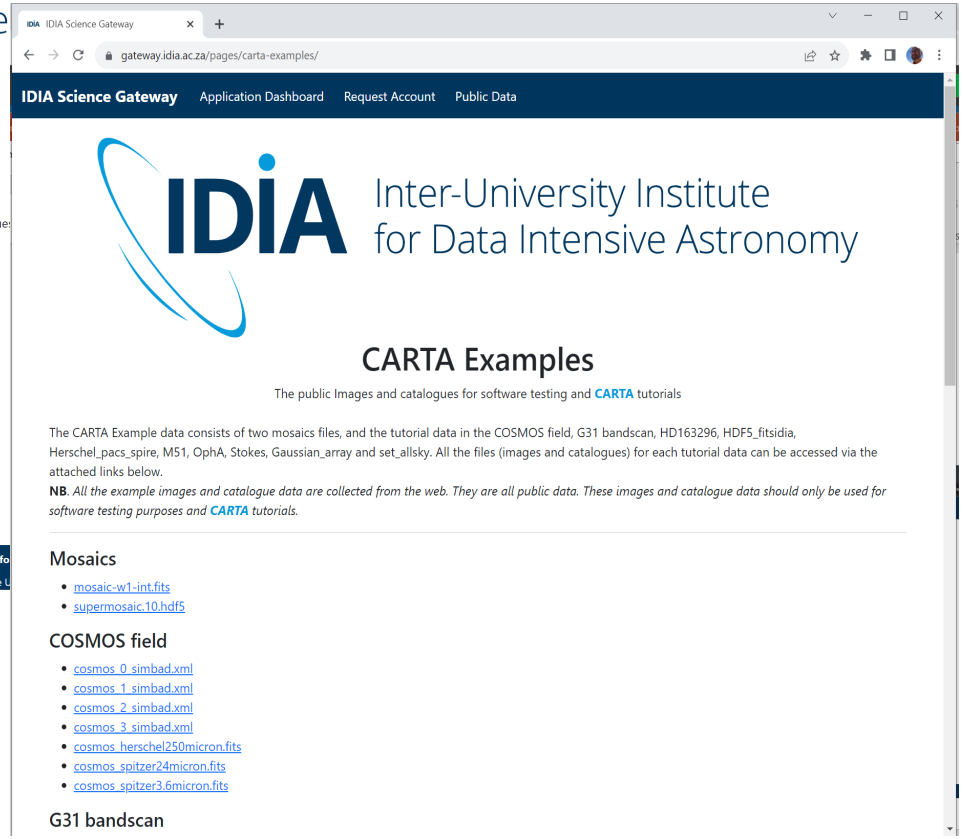
- Provided by several tools already
  - OnDemand
  - Juypter
  - Globus
  - Etc.
- Good enough if you know what is in the file from its name
- Not good for searching



Initial Public Data Publication

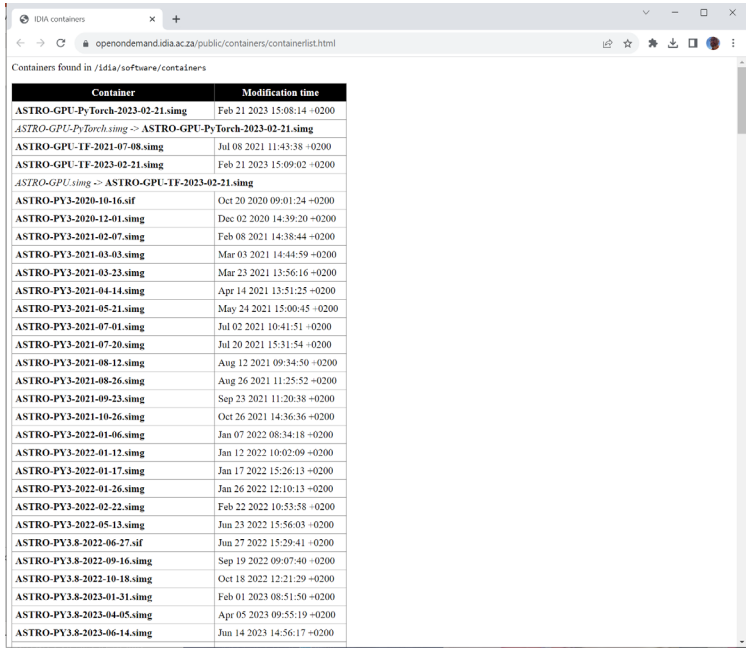




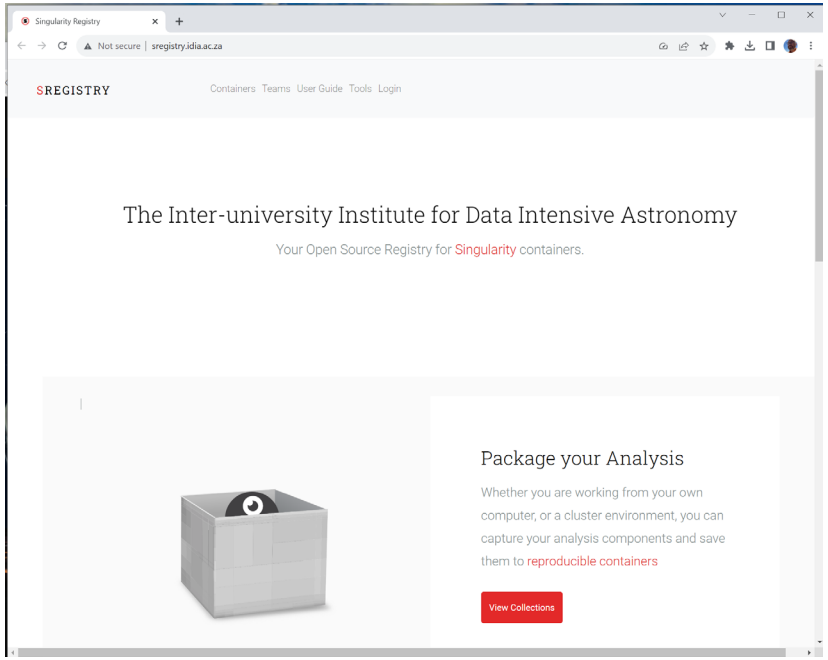


Application catalogue

- Initial gateway has simple listing of available containers
- Working on providing featureful container management system using sregistry



Container	Modification time
ASTRO-GPU-PyTorch-2023-02-21.simg	Feb 21 2023 15:08:14 +0200
ASTRO-GPU-PyTorch.simg -> **ASTRO-GPU-PyTorch-2023-02-21.simg**	
ASTRO-GPU-TF-2021-07-08.simg	Jul 08 2021 11:43:38 +0200
ASTRO-GPU-TF-2023-02-21.simg	Feb 21 2023 15:09:02 +0200
ASTRO-GPU.simg -> **ASTRO-GPU-TF-2023-02-21.simg**	
ASTRO-PY3-2020-10-16.sif	Oct 20 2020 09:01:24 +0200
ASTRO-PY3-2020-12-01.simg	Dec 02 2020 14:39:20 +0200
ASTRO-PY3-2021-02-07.simg	Feb 08 2021 14:38:44 +0200
ASTRO-PY3-2021-03-03.simg	Mar 03 2021 14:44:59 +0200
ASTRO-PY3-2021-03-23.simg	Mar 23 2021 13:56:16 +0200
ASTRO-PY3-2021-04-14.simg	Apr 14 2021 13:51:25 +0200
ASTRO-PY3-2021-05-21.simg	May 24 2021 15:00:45 +0200
ASTRO-PY3-2021-07-01.simg	Jul 02 2021 10:41:51 +0200
ASTRO-PY3-2021-07-20.simg	Jul 20 2021 15:31:54 +0200
ASTRO-PY3-2021-08-12.simg	Aug 12 2021 09:34:50 +0200
ASTRO-PY3-2021-08-26.simg	Aug 26 2021 11:25:52 +0200
ASTRO-PY3-2021-09-23.simg	Sep 23 2021 11:20:38 +0200
ASTRO-PY3-2021-10-26.simg	Oct 26 2021 14:36:36 +0200
ASTRO-PY3-2022-01-06.simg	Jan 07 2022 08:34:18 +0200
ASTRO-PY3-2022-01-12.simg	Jan 12 2022 10:02:09 +0200
ASTRO-PY3-2022-01-17.simg	Jan 17 2022 15:26:13 +0200
ASTRO-PY3-2022-01-26.simg	Jan 26 2022 12:10:13 +0200
ASTRO-PY3-2022-02-22.simg	Feb 22 2022 10:53:58 +0200
ASTRO-PY3-2022-05-13.simg	Jun 23 2022 15:56:03 +0200
ASTRO-PY3.8-2022-06-27.sif	Jun 27 2022 15:29:41 +0200
ASTRO-PY3.8-2022-09-16.simg	Sep 19 2022 09:07:40 +0200
ASTRO-PY3.8-2022-10-18.simg	Oct 18 2022 12:21:29 +0200
ASTRO-PY3.8-2023-01-31.simg	Feb 01 2023 08:51:50 +0200
ASTRO-PY3.8-2023-04-05.simg	Apr 05 2023 09:55:19 +0200
ASTRO-PY3.8-2023-06-14.simg	Jun 14 2023 14:56:17 +0200



African Shared Computing Projects

- African Open Science Platform
  - Has been in planning for 5+ years
  - 3 initial sites have been funded
- South African Open Science Cloud
  - Planned in collaboration with EOSC partners
  - Documents published by NICIS mid 2023
  - IDIA gateway provides much of what is planned for SAOSC

Summary

- High level requirement for IDIA Science Gateway developed over 6 months
- One aim was to provide federated access to all IDIA / ilifu services
- Working with DIRISA to ensure common federation approaches with their sites
- Aim at utilizing existing best in class systems as far as possible
  - Plan to specialize to user requirements while keeping customizations that need to be maintained to a minimum
- Provides input into development of SKA Regional Centres

IDIA
Inter-University Institute
for Data Intensive Astronomy