2014

Managing Research Data for Success

Darren Gibson
Edith Cowan University

Agnes Noronha
Edith Cowan University

Gordon McIntyre
Edith Cowan University

Luke Edwards
Edith Cowan University

Presented at Edith Cowan University, Perth, Western Australia on 21st August, 2014 (Joondalup) and 25th August, 2014 (Mount Lawley)

Managing Research Data for Success

Thursday 21st August, 2014, JO 21.429
Monday 25th August, 2014, ML 13.118

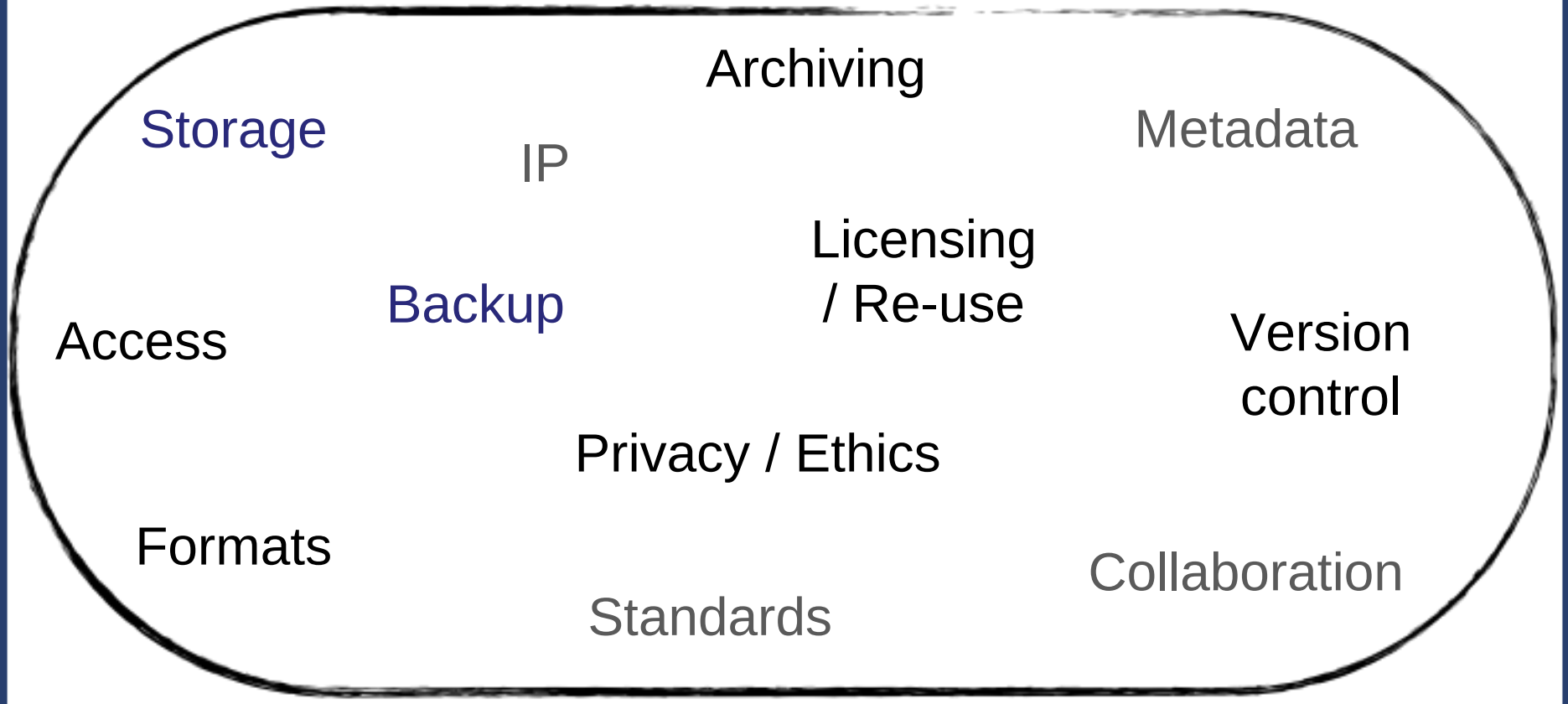
Darren Gibson, Manager, Research Collaboration and eResearch
Agnes Noronha, Senior Librarian: Research Services
Gordon McIntyre, Librarian: Research Services
Luke Edwards, Data Adviser, iVEC

Contents

- Overview of Research Data Management
- Current status at ECU
- ECU services
- National services
- Q&A

What is Research Data Management?

*Research Data Management is everything surrounding the actual "**research analysis**"*



ECU: Research Data Management Policy

- Nov/Dec 2013 The Academic Board approved:
  - Updated version of the Research Data Management Plan
  - Data Management Plan
  - Implementation Plan

http://intranet.ecu.edu.au/research/for-research-staff/policies-forms-and-guides

RDM Implementation

- 2014 – Interviews with Cat 1 grant holders and any other interested parties (feedback and DMP completion) *ORI/eResearch and Library*
- 2014/2015 – Feedback: challenges, updates required etc.
- 2015/2016 – Integrate DMP with other ECU systems
- 2016 onwards – ECU wide implementation

Why worry about RDM?

- Why is this important to me as a researcher?

- Why does ECU care?

National and funding body policies

Australian Code for the Responsible Conduct of Research (NHMRC, 2007) Section 2: Management of Research Data & Primary Materials

"Each institution must have a **policy** on the retention of materials and research data. It is important that institutions acknowledge their continuing role in the management of research material and data ."

Australian Research Council

ARC Discovery Grant requirements, **February 2014**

Researchers are now required as part of the application process for National Competitive Grants Program funding to outline how they ***plan to manage*** research data arising from ARC-funded research.

RDM and eResearch

- Assist with the completion of the DMP
- Monitor DMPs
- Refine RDM Policy and DMP
- Assist with finding appropriate storage
- Assist with sharing data with collaborators
- Work with the Library, ITSC and National providers

RDM and the library

The library can assist researchers with:

- Deciding where datasets will be deposited
- Developing the metadata records to describe datasets
- Determining the type of access to be granted to datasets
- Caretaking of datasets described and uploaded to Research Online

Datasets and Research Online

- The Library Research Services team manage ECU's institutional repository **Research Online**
- The Research Services team can support researchers with **4.7 Data Reuse** of the ECU RDM policy by
  - Creating metadata records to datasets in Research Online
  - Storing and managing datasets (within technical limitations)
  - Mapping metadata records to **Research Data Australia**, the national discovery tool for Australian research datasets

Publishers and open data

- Publishers have begun to introduce open data policies such as Public Library of Science (PLOS) and Nature
- Some details of these policies include:
  - Submitting a Data Availability Statement with the article manuscript (PLOS)
  - Recognising that “others should be able to **replicate** and **build** upon the authors' published claims” (Nature)
  - Providing a number of options to enable compliance based on the nature of the data

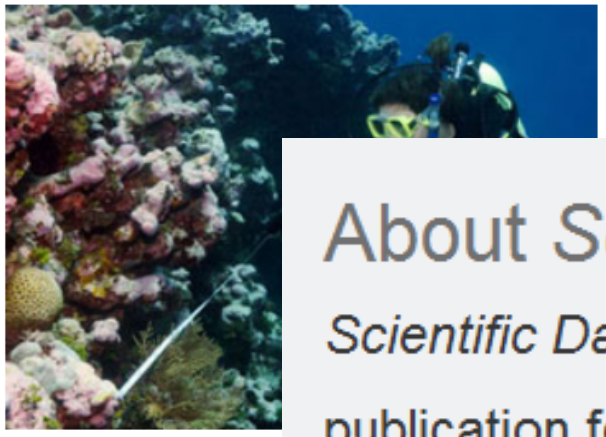
SCIENTIFIC DATA

Search Go

Advanced search

Home | Archive | About | For Authors | For Referees | Advisory & Editorial Board | Data Policies

Featured Data Descriptor



Systematic global assessment of reef fish communities by the Reef Life Survey program

Graham J. Edgar and Rick D. Stuart-Smith

volunteer divers to asse
describe the data collec
the wider scientific com

About *Scientific Data*

Scientific Data is an open-access, peer-reviewed publication for descriptions of scientifically valuable datasets. Our primary article-type, the **Data Descriptor**, is designed to make your data more discoverable, interpretable and reusable.

RSS

Twitter

ipt

als

About *Scientific Data*

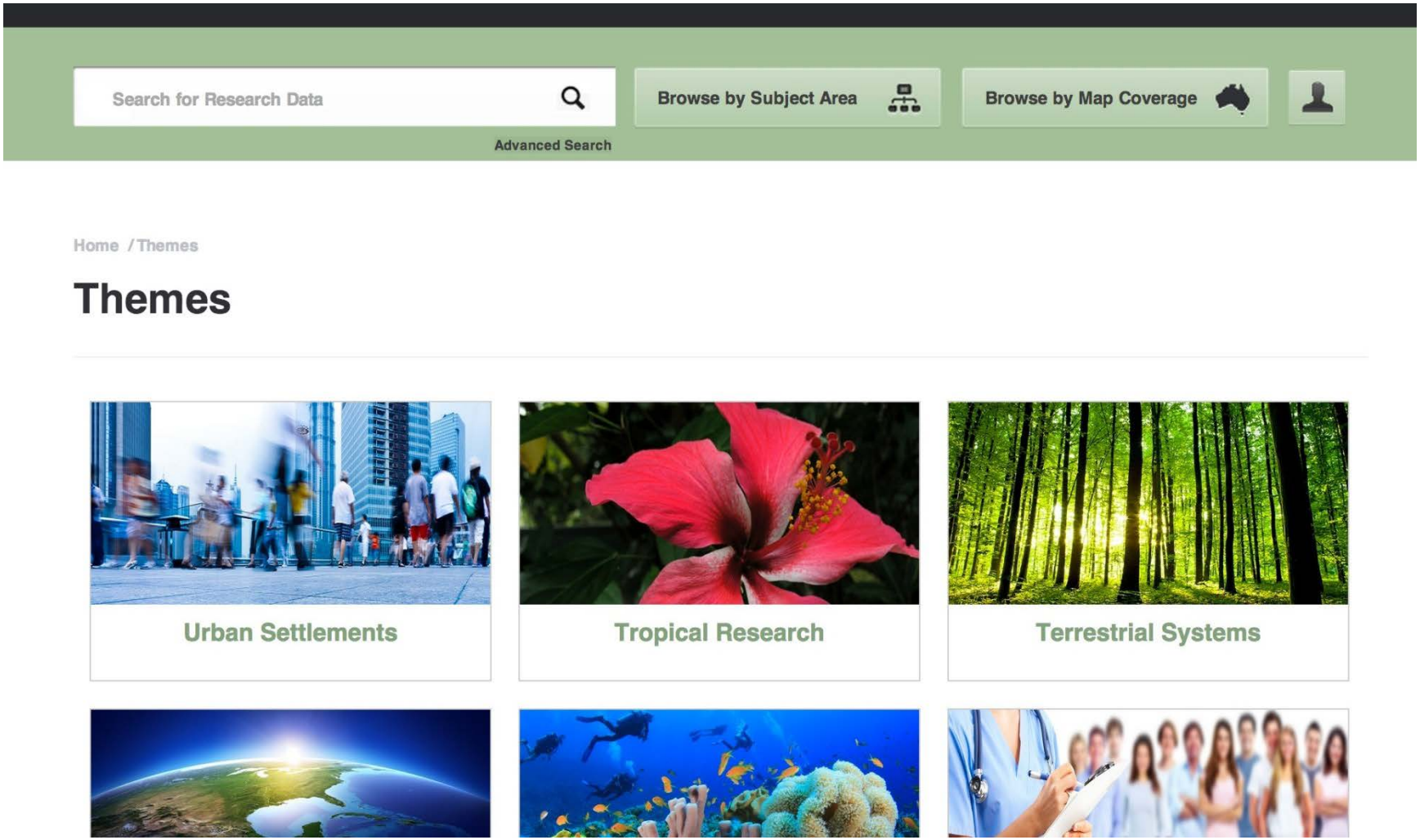
Scientific Data is an open-access, peer-reviewed publication for descriptions of scientifically valuable datasets. Our primary article-type, the **Data Descriptor**, is designed to make your data more discoverable, interpretable and reusable.

http://www.**nature**.com/sdata/
Launched May 2014

Benefits of sharing research data

- Piwowar, et. al., “Sharing Detailed Research Data Is Associated with Increased Citation Rate”
  - http://www.plosone.org/article/info:doi/10.1371/journal.pone.0000308
- Looked at the citation history of cancer microarray clinical trial publications
- Found that publicly available data was associated with a 69% increase in citations, independently of journal impact factor, date of publication, and author country of origin

Research Data Australia



http://researchdata.ands.org.au

International initiatives for research data

OECD principles and guidelines for access to research data from public funding 2007

Research Councils UK Common Principles on Data Policy

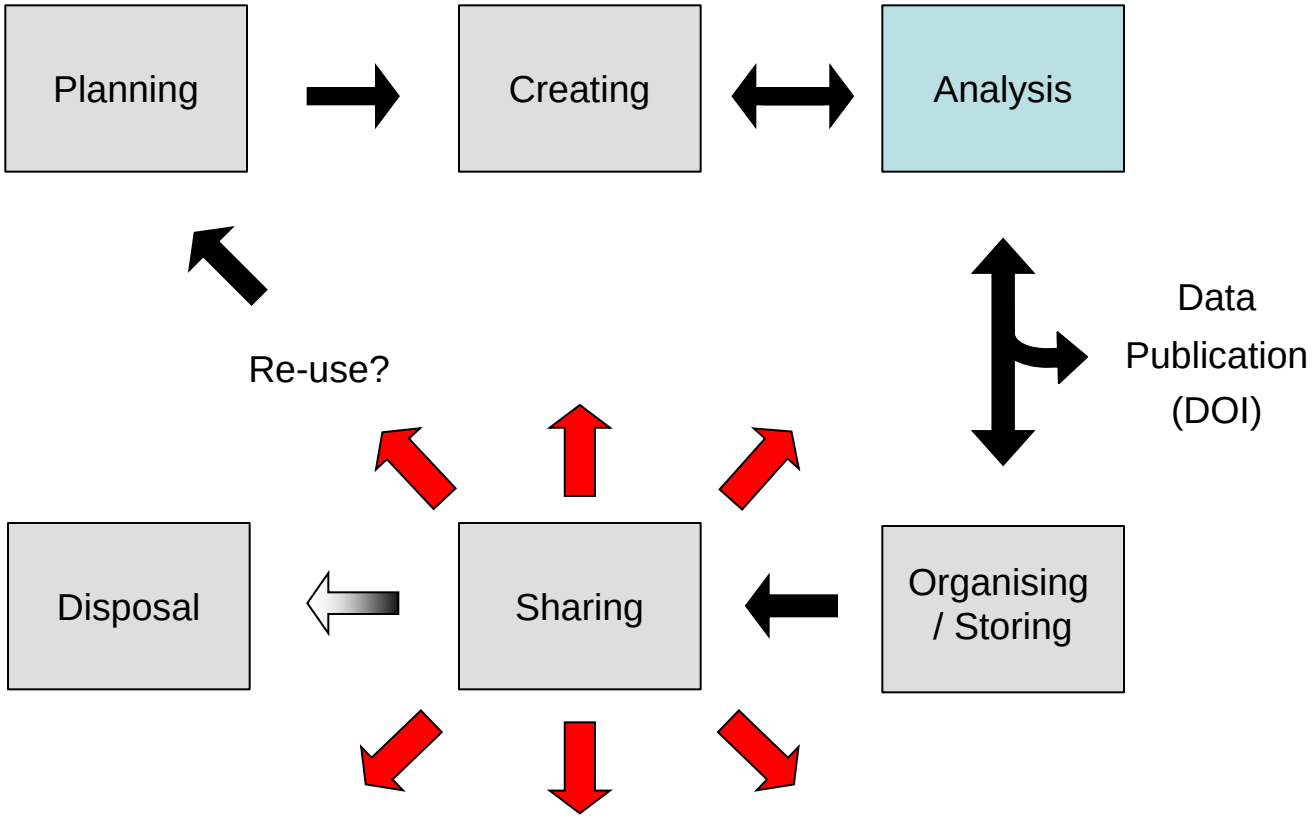
National Institutes of Health USA Data Sharing policy

G8 Open Data Charter

Luke Edwards

Data Adviser (luke.edwards@ivec.org)

Research data life-cycle



Motivation

- Efficiencies / reduce duplication
- Transparency
- Exposure / maximise impact

Australian Code for the Responsible Conduct of Research and funding bodies (e.g. ARC) have requirements

http://www.data-archive.ac.uk/create-manage/planning-for-sharing/why-share-data

Climate sceptic wins landmark data victory 'for price of a stamp'

Belfast ecologist forced to hand over tree-ring data describes order from information commission as a 'staggering injustice'

Fred Pearce
The Guardian, Tuesday 20 April 2010
Jump to comments (760)

The Queen's University of Belfast, Northern Ireland, must hand over 40 years' worth of data on 7,000 years of Irish tree rings. Photograph: Ron Sachs / Rex Features/Rex Features

An arch-critic of climate scientists has won a major victory in his campaign to win access to British university data that could reveal details of Europe's past climate.

In a landmark ruling, the UK Information Commissioner's Office has ruled that Queen's University Belfast must hand over data obtained during 40 years of research into 7,000 years of Irish tree rings to a City banker and part-time climate analyst, Doug Keenan.

Services available to researchers

iVEC

- Data storage and access
- Supercomputing
- Visualisation
- Training / Data clinics

NeCTAR

- Research Cloud (https://www.nectar.org.au/research-cloud)

ANDS

- Research Data Australia (metadata)
- Persistent Identifier Services e.g. Cite My Data (DOIs)
- Training

Services available to researchers - Compute

Magnus

Courtesy of Dr George Beckett - iVEC Annual Symposium, 20th February 2014

Services available to researchers - Visualisation

People

- Andrew Squelch - Jonathan Knispel - Paul Bourke - Yathunanthan Sivarajah

Resources invested across iVEC partners

- Stereoscopic 3D displays
- Immersive hemispherical displays
- High resolution displays
- High performance graphics workstations
- Spherical display
- Specialist still capture cameras
- Specialist video cameras
- 3D capture technology
- User interface and tracking systems
- Commercial licenses for a high end visualisation software packages.

Services available to researchers - Data

Storage

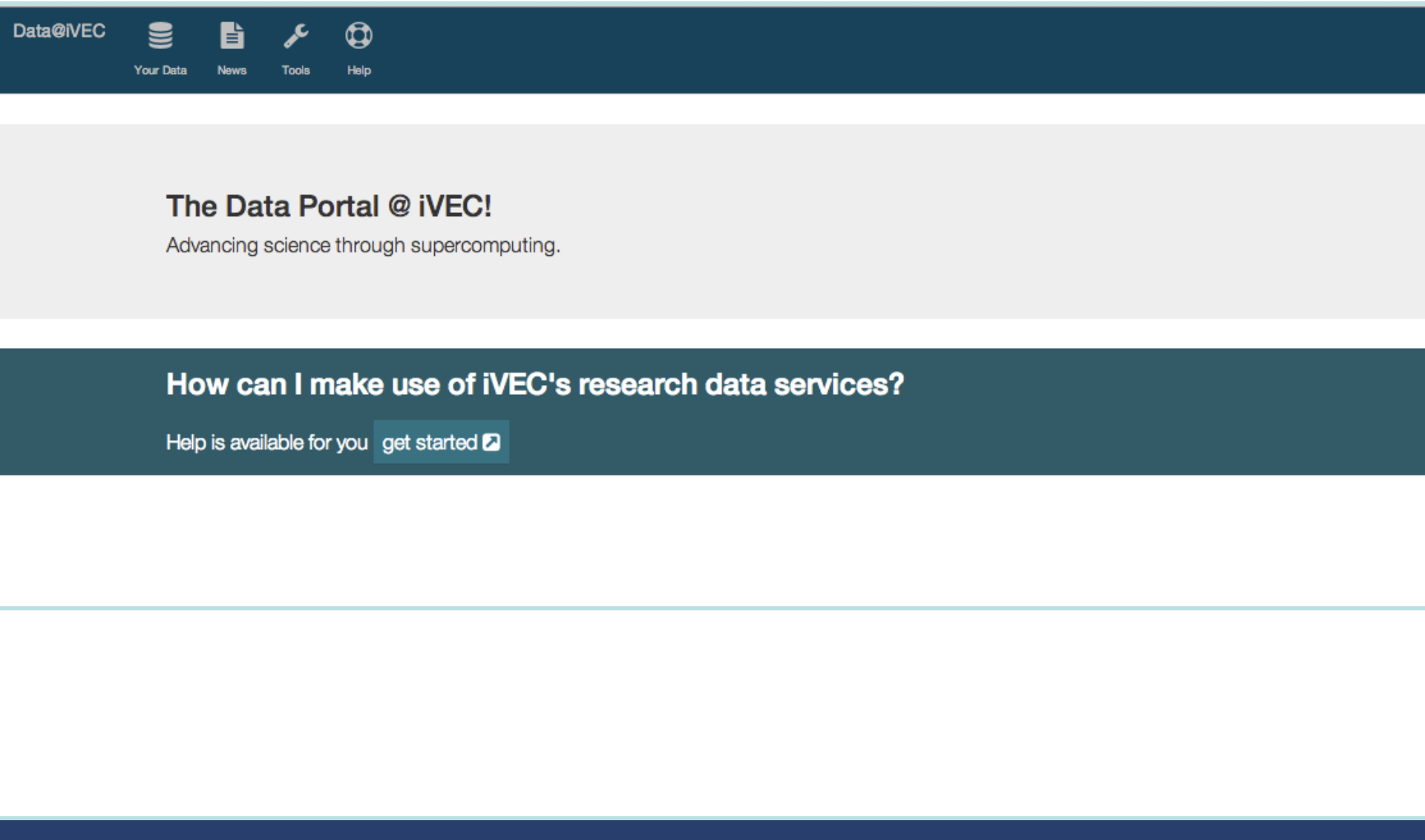
Existing service (soon to be superseded)

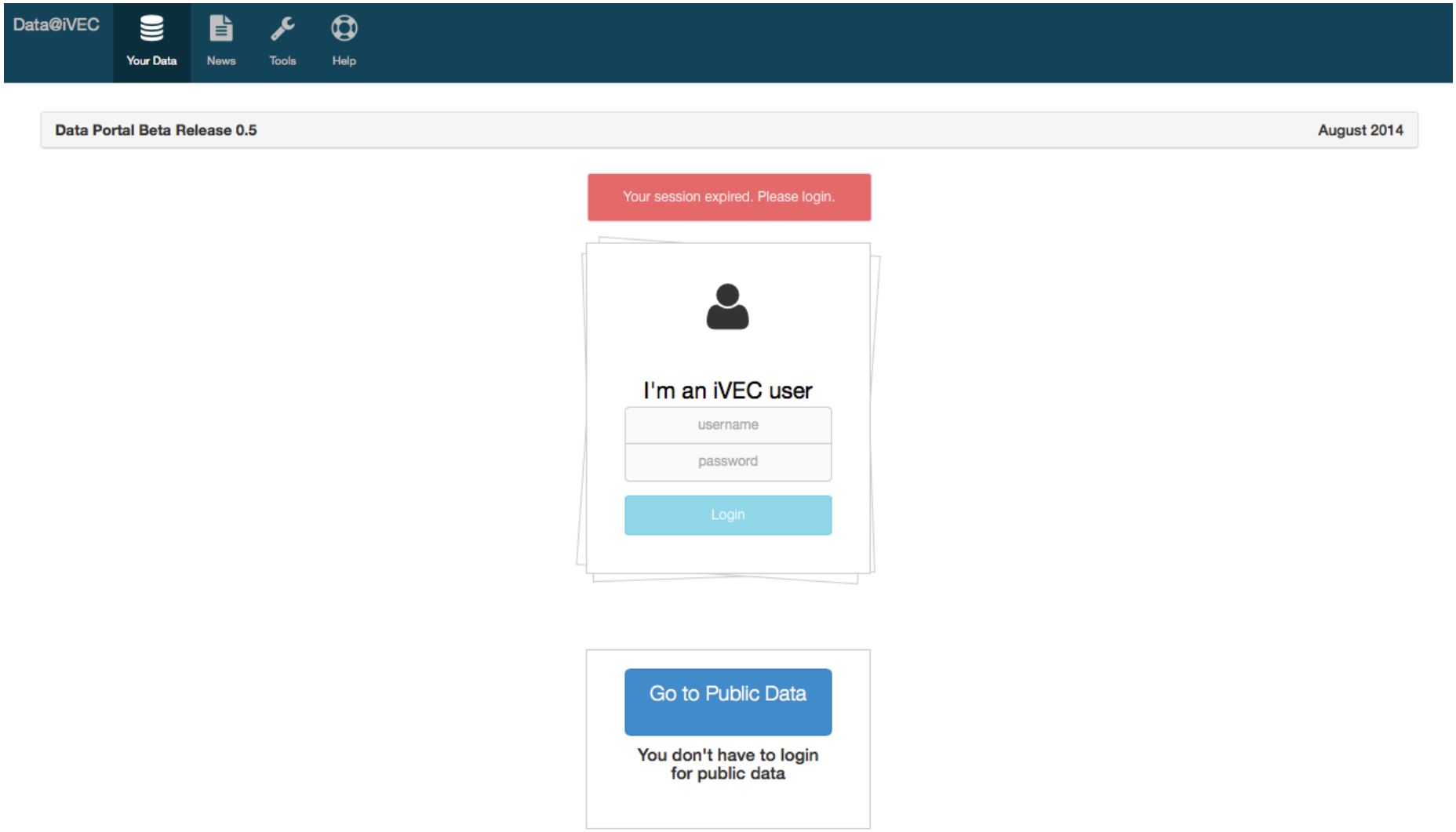
- cortex.ivec.org
  - Needed iVEC account to download
  - "Command" line required
  - "Staging" of data required

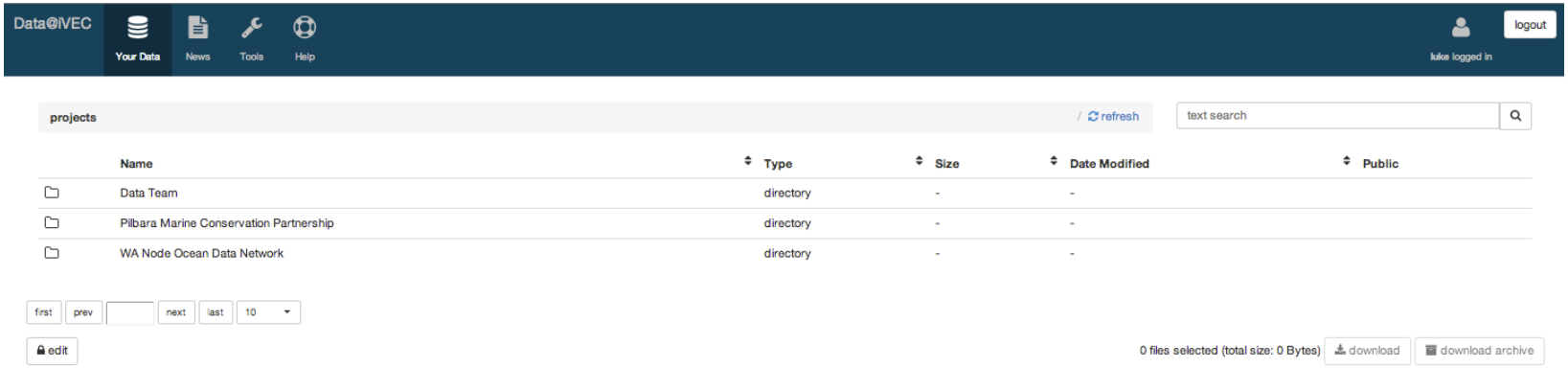
New service

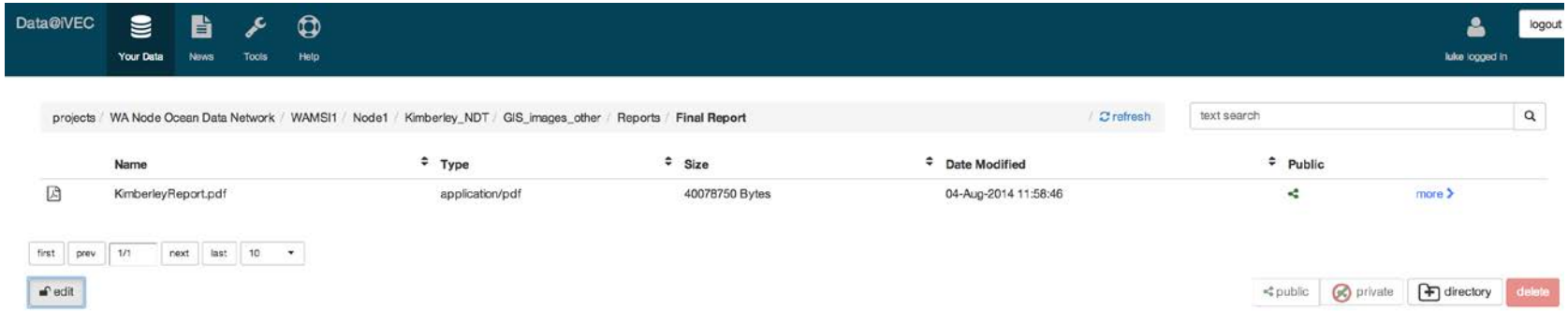
- iVEC Data Storage (https://data.ivec.org) - "LiveARC"
  - Web interface (navigate, search and download)
  - WebDAV (upload data - size limits)
  - Application process (>5TB) http://www.ivec.org/services/data-storage/
  - Focus on 'active' research

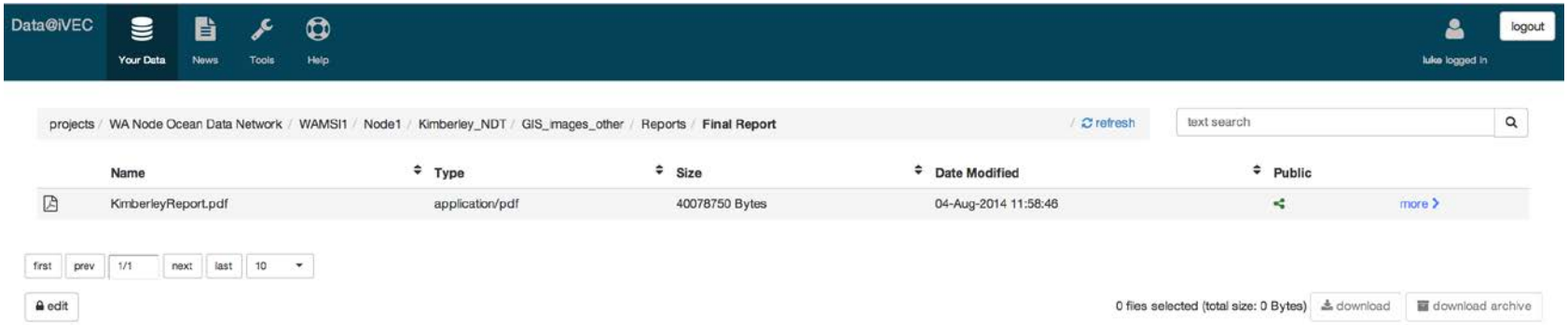
https://data.ivec.org/




Data@iVEC
Your Data
News
Tools
Help
Data Portal Beta Release 0.5
August 2014
Your session expired. Please login.
I'm an iVEC user
username
password
Login
Go to Public Data
You don't have to login
for public data


Data@iVEC
Your Data
News
Tools
Help
luke logged in
logout
projects
/ refresh
text search
Name
Type
Size
Date Modified
Public
Data Team
directory
-
-
Pilbara Marine Conservation Partnership
directory
-
-
WA Node Ocean Data Network
directory
-
-
first
prev
next
last
10
edit
0 files selected (total size: 0 Bytes)
download
download archive


Data@iVEC
Your Data
News
Tools
Help
luke logged in
logout
projects / WA Node Ocean Data Network / WAMSI1 / Node1 / Kimberley_NDT / GIS_images_other / Reports / Final Report
refresh
text search
Name
Type
Size
Date Modified
Public
KimberleyReport.pdf
application/pdf
40078750 Bytes
04-Aug-2014 11:58:46
more
first
prev
1/1
next
last
10
edit
public
private
directory
delete


Data@iVEC
Your Data
News
Tools
Help
luke logged in
logout
projects / WA Node Ocean Data Network / WAMSI1 / Node1 / Kimberley_NDT / GIS_images_other / Reports / Final Report
refresh
text search
Name
Type
Size
Date Modified
Public
KimberleyReport.pdf
application/pdf
40078750 Bytes
04-Aug-2014 11:58:46
more
first
prev
1/1
next
last
10
edit
0 files selected (total size: 0 Bytes)
download
download archive

Tools

You'll need these to manage your data resources.

1. First grab these files. You'll need to put them on the computer with your data (ie: fornax, magnus etc)

aterm.jar and ashell.py

2. Run the ashell.py program

```
chmod u+x ashell.py
./ashell.py
```

You should see

```
Welcome to ashell v0.4, type 'help' for a list of commands
ivec:offline>
```

3. Now you need to login to the data store:

```
ivec:offline>login
Username: markg
Password:
ivec:online>
```

4. You're now ready to upload files:

```
ivec:online>upload ./myfile.txt to '/projects/my project/folder/'
```

...when this is complete you should be able to see your file/s in the data tab here, or in WebDAV

type 'help' in the ashell for more information on all this and other options and commands

I'm thinking about using iVEC for my research data storage...

I've submitted my application...

I administer/own a project...

I'm a researcher in a project...

Services available from iVEC - Data

Training

- Research Data Management (part of)

http://www.ivec.org/services/training/

Access all iVEC services via:

http://www.ivec.org/services/access/

Conclusion

Research data

consider early …

… consider often …

… consider context …

… consider communication…

Borrowed from Monash University (http://monash.edu/library/researchdata/resources/training/ - Intellectual property and research data management for supervisors)

Future

https://twitter.com/HistoricalPics/status/490451073036210176/photo/1

Useful information

- http://datalib.edina.ac.uk/mantra/ (Online course)

- http://www.data-archive.ac.uk/create-manage

- http://www.ands.org.au/resource/data-management-planning.html

- http://libguides.anu.edu.au/datamanagement

- http://ecu.au.libguides.com/research_data_management

- https://www.youtube.com/watch?v=PgqtiY7oZ6k - Data citation

- http://youtu.be/sum3AkmET1c?t=2m30s

- Example of workflow using Pawsey centre

http://5stardata.info/ - Open data

Interactive discussion (Q&A)

- What are the types of data you collect?
- What projects are you working on?
- What avenues are you going to use to communicate your work? Who will be interested?
- What challenges are you facing? Are there any particular 'roadblocks'?
- http://www.gigasciencejournal.com/about - Data journals