DataONE Webinar Series

Incentives, Challenges, Barriers: Exploring social, institutional and economic reasons for sharing data

Jens Klump

Heidi Laine

Fiona Murphy

DataONE

DataONE

Enable new science and knowledge creation through universal access to data about life on earth and the environment that sustains it

DataONE network of Member Nodes:
www.dataone.org/current-member-nodes

DataONE Search tool:
https://search.dataone.org

DataONE

Lesson 3: Data Management Planning

View all Education Modules at https://www.dataone.org/education-modules

The Data Management Plan (DMP)

A DMP outlines what you will do with your data during and after you complete your research project. It is a formal, [illegible] document in which you lay out a pla[illegible] the present[illegible]

Component 2: Metadata content & format

- What metadata are needed?
- How will metadata be created and/or captured?

DataONE

Lesson 2: Data Sharing

View all Education Modules at https://www.dataone.org/education-modules

Address data sharing throughout the data lifecycle

Describe data content, character, and process.
Deposit in a location from which it can be accessed.
Preserve in [illegible] for long term.
Publish info[illegible] discover it [illegible]

Concerns about data sharing

Concern	Solution

DataONE

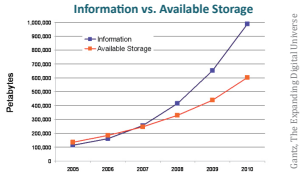
Lesson 1: Data Management

View all Education Modules at https://www.dataone.org/education-modules

The world of data around us

The data deluge has created a surge of information that needs to be well-managed, discoverable, and accessible.

The amount of available storage is not keeping pace with the amount of data being produced.



Gantz, The Expanding Digital Universe

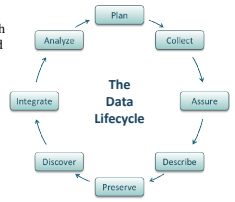
Causes of data loss

- Natural disasters
- Facilities infrastructure failures
- Storage failure
- Server hardware or software failure
- Application software failure
- Human errors
- Malicious attack
- Format obsolescence
- Loss of competencies
- Loss of funding
- Loss of insitutional commitment

Costs of not doing data management can be very high!

The Data Lifecycle

The stages through which well-managed data passes from project inception to conclusion.



Why manage data: the researcher perspective

- Keep yourself organized ⇨ find your own files!
- Track your processes for reproducibility
- Better version control of data
- More efficient data quality control
- More backups to avoid data loss
- Format your data for reuse by yourself & others
- Document your data for understability and reuse
- Prepare it to share it & gain credibility and recognition for your scientific efforts

Data management facilitates sharing and reuse.

Data Reuse Example

Researchers reused and aggregated data from several different sources to determine migration routes for specific bird species.

© Jerry & Sherry Liguron

The Case for Data Management

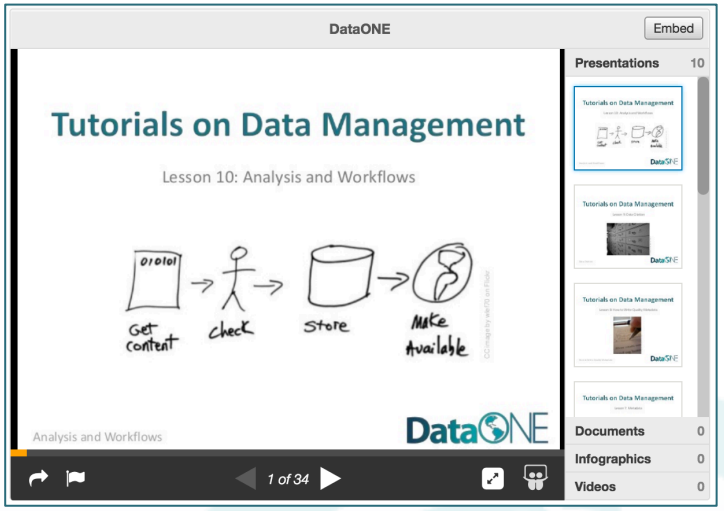
If data are:

- Well-organized
- Documented
- Preserved
- Accessible
- Verified as to accuracy & validity

The results are:

- High quality data
- Data that is easy to share and reuse
- Citation & credibility to researcher
- Cost savings to further science

Local contact information



#DWS2016
@DataONEorg
www.DataONE.org

DwS

DataONE Webinar Series

www.dataone.org/webinars

Upcoming Webinar Event

www.dataone.org/upcoming-webinar

Previous Webinar Events (Recording and Discussion)

www.dataone.org/previous-webinars

#DWS2016

@DataONEorg

Upcoming Webinar Event

www.dataone.org/upcoming-webinar

How not to collect data: organizing data for long-term use and re-use

October 11

Stephanie E. Hampton

Washington State University

If you attending
as part of a group,
please enter the number of
people listening within
the “questions” box
Thanks!

Incentives, Challenges, Barriers: Exploring social, institutional and economic reasons for sharing data

Jens Klump
CSIRO

Heidi Laine
University of Helsinki

Fiona Murphy
University of Reading

Data as Social Capital and the Gift Culture in the Scholarly Community

Jens Klump | OCE Science Leader Earth Science Informatics

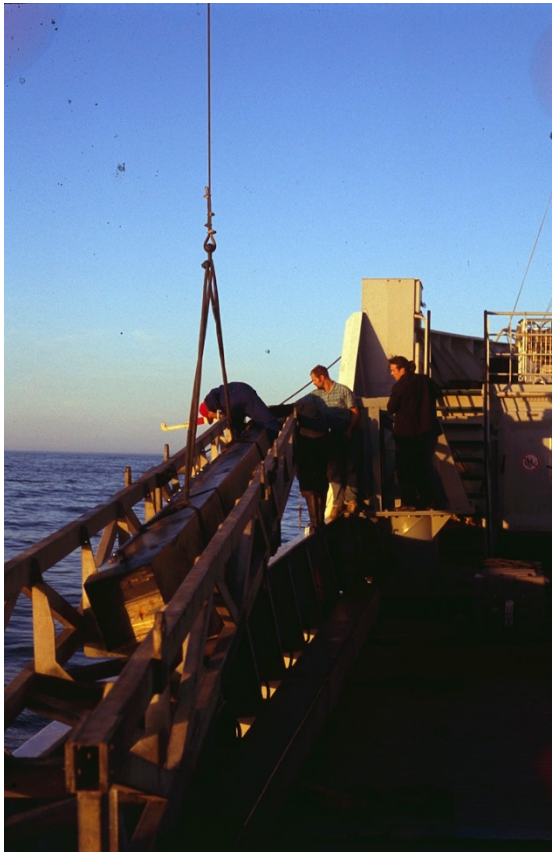
13 September 2016

Why am I interested in data sharing?

I am not a sociologist.

- I am a geochemist, my field of research is Earth Science Informatics.
- Research data infrastructures have been part of my work since 1999.
- When I switched from marine geology to limnology I was puzzled by the difference in attitudes towards data sharing.
- Over the years I made some observations in this respect.

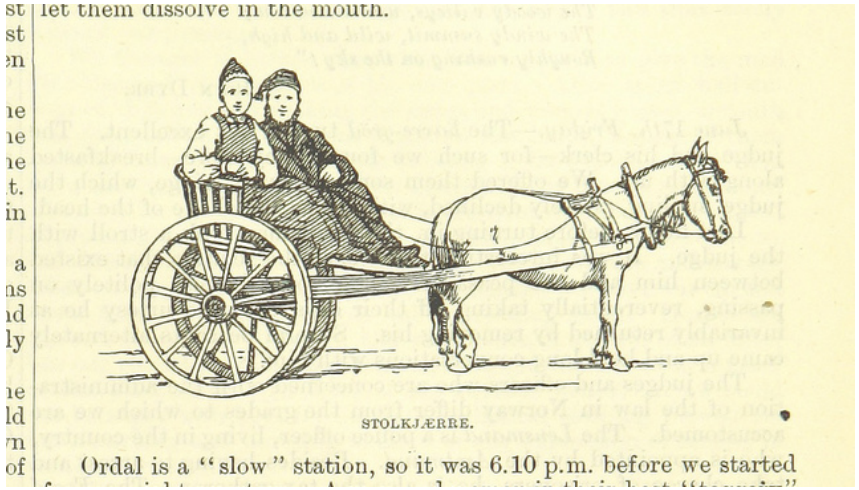
Why do communities behave so differently?



South Atlantic (Namibia)

Lake Baikal (Russia)

Are we getting the incentives right?


let them dissolve in the mouth.

STOLKJÆRRE.

Ordal is a "slow" station, so it was 6.10 p.m. before we started

Image: British Library (public domain)

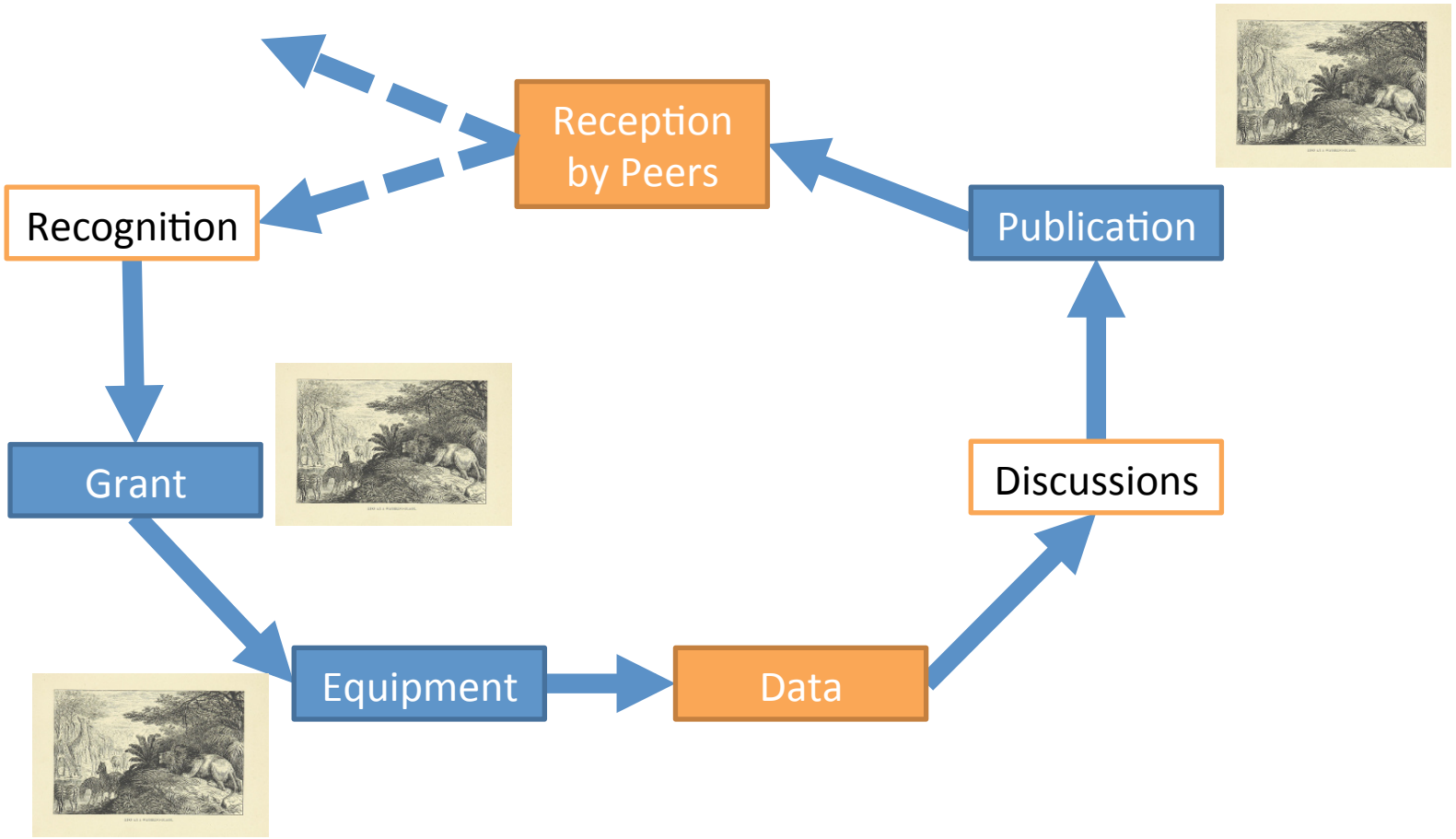
Image: ETH Zürich Library (public domain)

Waiting at the watering hole

- Sometimes waiting at the watering hole can be a successful strategy.
- The art is to identify suitable watering holes.
- Which resources do researchers need to access for their distinction gain?
- This is not only an opportunity to coerce compliant behaviour but also to develop better services for researchers.

Image: British Library (public domain)

Reputation Economy



Mineral Resources
Jens Klump
Science Leader Earth Science Informatics

t +61 8 6436 8828
e jens.klump@csiro.au
w www.csiro.au/en/Research/MRF

AFRAID OF SCOOPING?

Case Study on Perceived vs. Actualized Risks of Sharing Research Outputs

HEIDI LAINE / heidi.laine@helsinki.fi /
@heidiklaine

DataONE

This work is licensed under a

Creative Commons Attribution 4.0 International License.

SCOOPING AS MISCONDUCT

Misappropriation: unlawful presentation of another person's result, idea, plan, observation or data as one's own research

- Finnish Advisory Board on Research Integrity: *Responsible conduct of research and procedures for handling allegations of misconduct in Finland - RCR guidelines*, available at http://www.tenk.fi/en/resposible-conduct-research-guidelines

RESEARCH QUESTIONS

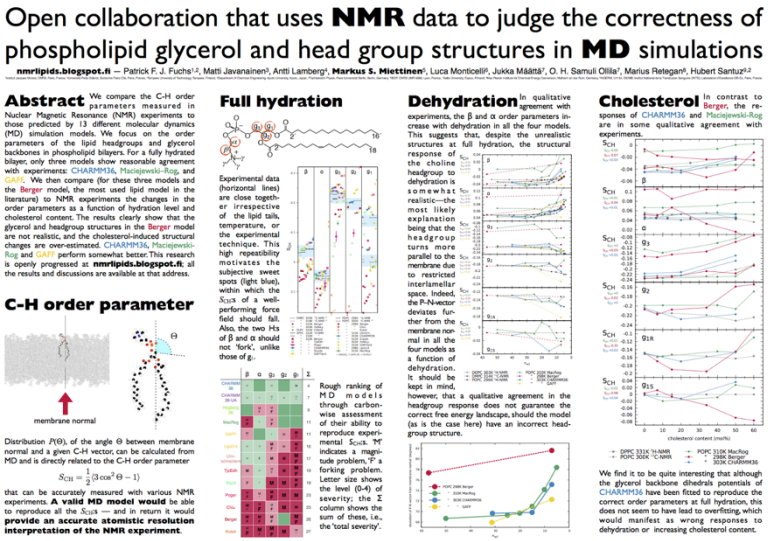
Descriptive level: How two pioneering open collaboration research projects coped with the risk of research misconduct?

Interpretive level: What allowed the researchers of the two case projects to ignore the risk of research misconduct?

OPEN COLLABORATION CASES

Open Research Swarm / Social Media for Citizen Participation (SOMUS), 2007-2010

NMR Lipids Project (NMRLP), 2013-ongoing



RADICAL OPENNESS

ORS / SOMUS	**NMRLP**
Open collaborative writing of research proposal to a funding call	Participants required to give credit for ideas also from “hallway” discussions
Funding allocated to an open online community	Data shared via Github without embargo
Open collaborative writing of conference papers, meeting notes, wiki	Co-authorship determined by input through blog comments

LESSONS TO BE LEARNED?

- Research ethics and integrity training for researchers is vital also for advancement of open science.
- Multidisciplinarity fosters innovation: get physicists reading philosophy, engineering scientists working with sociologists.
- Data citation practices and data publication workflows should be implemented to ensure that openness pays of.
- Lack of gender balance indicative of women researcher's lower tolerance of risk? Targeted efforts for encouraging women to share.
- Pioneers should be recognised and rewarded with funding and positions!
- Openness as an empowering experience should be promoted.

FIND OUT MORE

Longer version of this presentation: bit.ly/fearofscooping

Paper will be submitted for the Data Science Journal SciDataCon special collection and will be available as a pre-print on zenodo.org as of 1 October 2016.

Follow or contact me:
@heidiklaine / www.thehonestbrokerblog.org / heidi.laine@helsinki.fi

Giving Researchers Credit for their Data (or 'Data2Paper')

Fiona Murphy, Neil Jefferies, Anusha Ranganathan, Thomas Ingraham, Hollydawn Murray

JISC

DataOne Webinar, Denver, 13 September 2016

DataONE

The Business Case

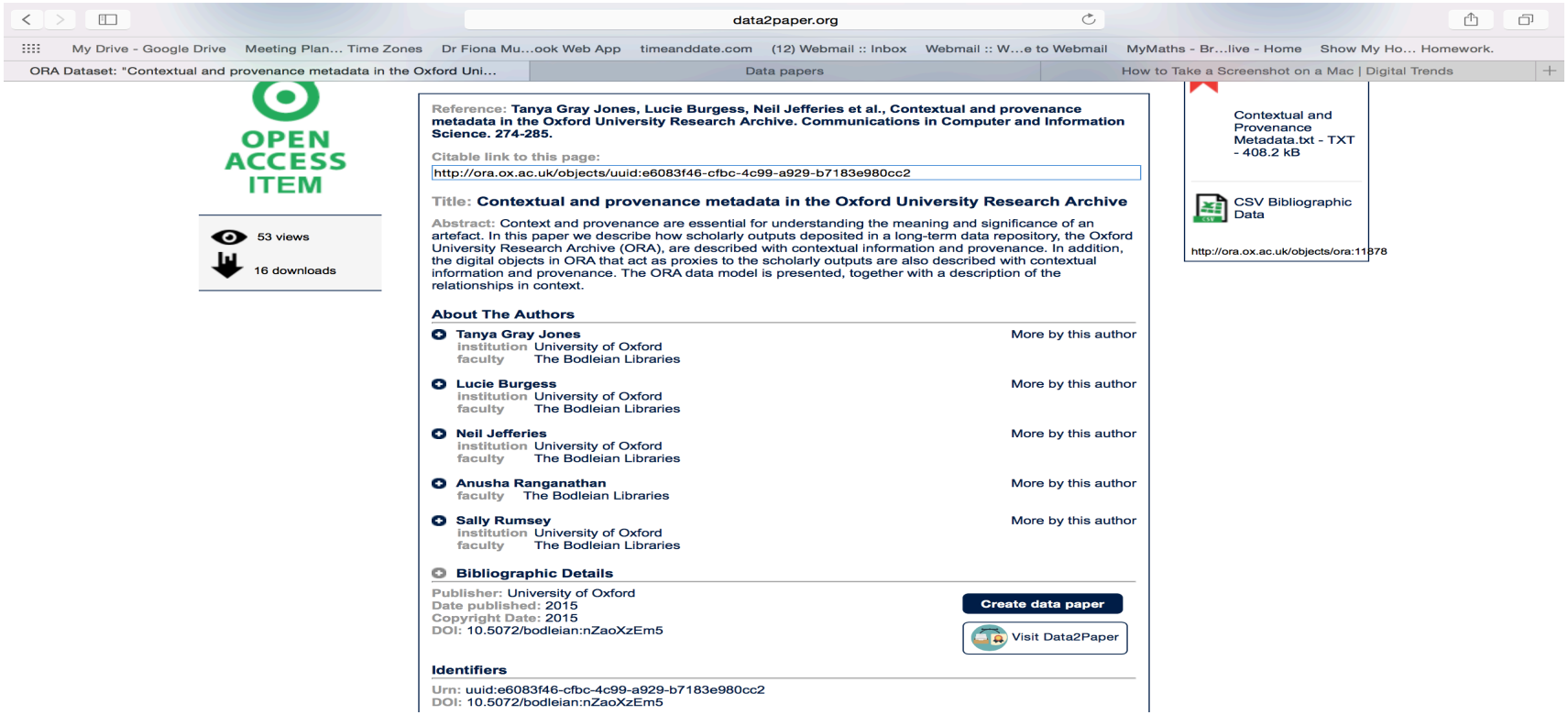
JISC

The App

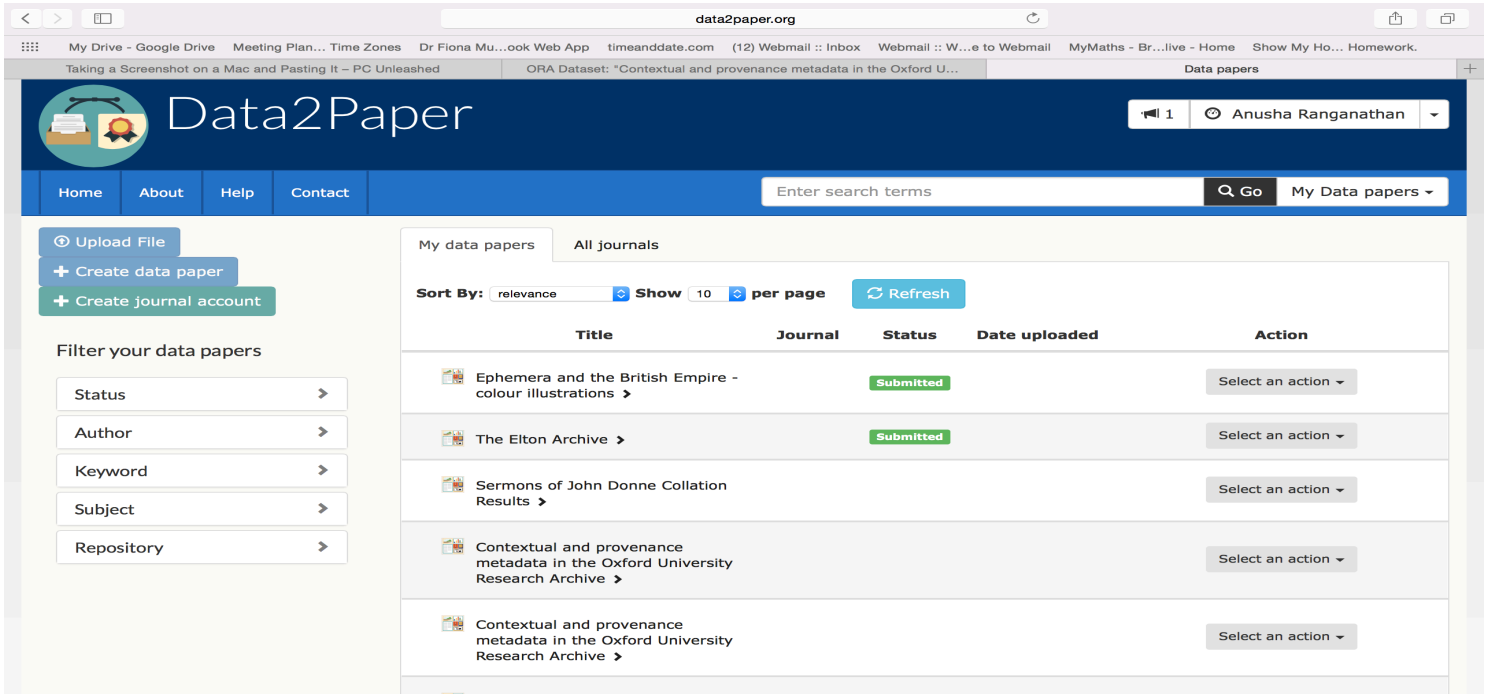
- Researcher logs in to app via institutional repository
- Automated data paper compilation within the research space (*no journal submission systems*)
- Article packages submitted to publisher system via SWORD
- If published, ORCID updates the record

JISC

Institutional Repository View



Data2Paper


data2paper.org
My Drive - Google Drive
Meeting Plan... Time Zones
Dr Fiona Mu...ook Web App
timeanddate.com
(12) Webmail :: Inbox
Webmail :: W...e to Webmail
MyMaths - Br...live - Home
Show My Ho... Homework.
Taking a Screenshot on a Mac and Pasting It – PC Unleashed
ORA Dataset: "Contextual and provenance metadata in the Oxford U...
Data papers
Data2Paper
1
Anusha Ranganathan
Home
About
Help
Contact
Enter search terms
Go
My Data papers
Upload File
Create data paper
Create journal account
Filter your data papers
Status
Author
Keyword
Subject
Repository
My data papers
All journals
Sort By: relevance
Show 10 per page
Refresh
Title
Journal
Status
Date uploaded
Action
Ephemera and the British Empire - colour illustrations
Submitted
Select an action
The Elton Archive
Submitted
Select an action
Sermons of John Donne Collation Results
Select an action
Contextual and provenance metadata in the Oxford University Research Archive
Select an action
Contextual and provenance metadata in the Oxford University Research Archive
Select an action

Next Steps

- Feedback, UI/UX work
- Researchers – domain and social factors
- Outreach and trials
- Sustainable long-term strategy
- All suggestions welcome!

JISC

Thank you!

fionalm27@gmail.com
@DrFionaLM

Jisc: Research Data Spring Grant

JISC

DataONE Webinar Series

www.dataone.org/webinars

Upcoming Webinar Event

www.dataone.org/upcoming-webinar

Previous Webinar Events (Recording and Discussion)

www.dataone.org/previous-webinars

#DWS2016

@DataONEorg