



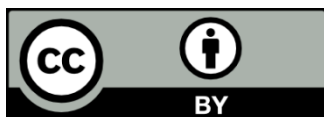
Research Data Management and the role of libraries

Sarah Jones

Digital Curation Centre

sarah.jones@glasgow.ac.uk

Twitter: @sjDCC



ADBU conference: Quelle(s) stratégie(s) de recherche face à la nouvelle massification des données?, Legal session, 2nd December 2014, Paris

Funded by:



What is the DCC?

A UK service to support the Higher Education sector with Research Data Management (RDM)

“Helping to build capacity, capability and skills in data management and curation across the UK’s higher education research community.”

DCC Phase 3 Business Plan



Training | Events | Tools | Advocacy | Tailored Support | IJDC | International Conference

www.dcc.ac.uk

Why manage research data?

“If an institution spent A\$10 million on data, what would be the return? The answer is: more publications; an increased citation count; more grants; greater profile; and more collaboration.”

Dr Ross Wilkinson, ANDS

www.ariadne.ac.uk/issue72/oar-2013-rpt



CC-BY by Simon Cunningham www.flickr.com/photos/lendingmemo/11697134804

The 'big data' challenge for universities



Often curated by large international projects, national facilities or disciplinary initiatives.

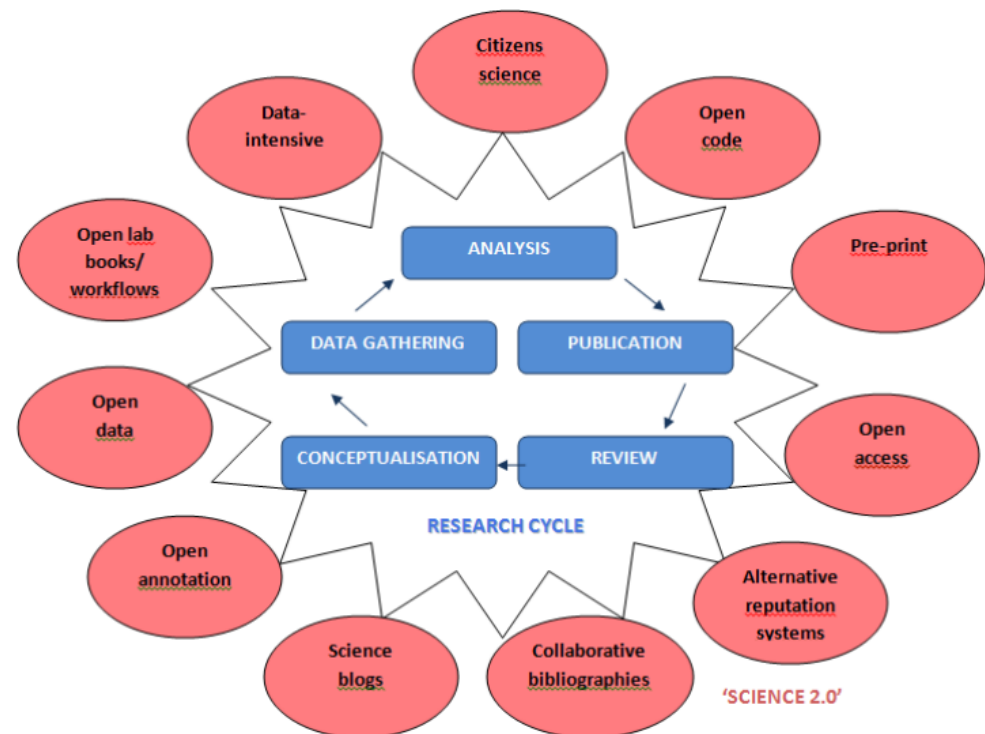


Some large data, but mostly a long-tail of small, distributed datasets with diverse needs.

Science is changing

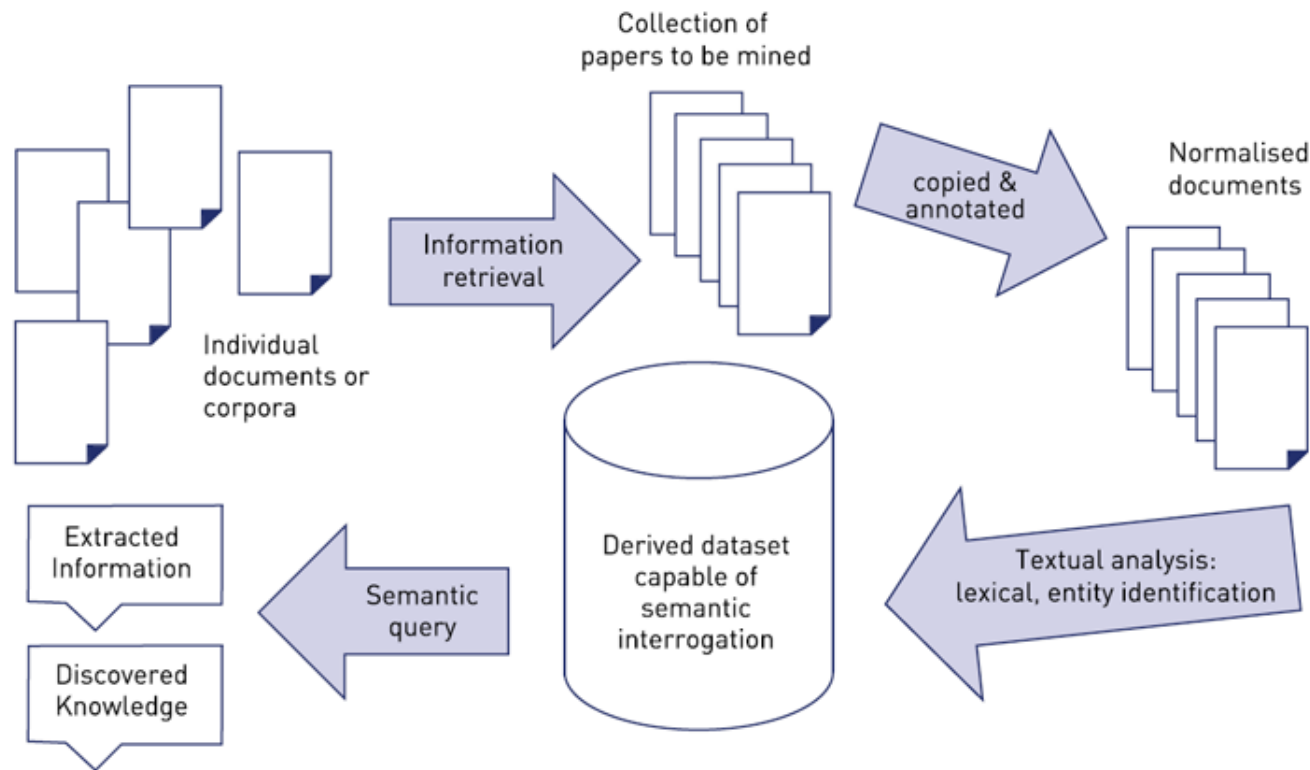
‘Science 2.0’ represents a paradigm shift in the modus operandi of research and science impacting the entire scientific process.

- More open
- Data intensive
- Collaborative
- Sharing throughout research process
- ...



Text and data mining

The process of copying data required for TDM may infringe copyright, even if the user has access rights to the material

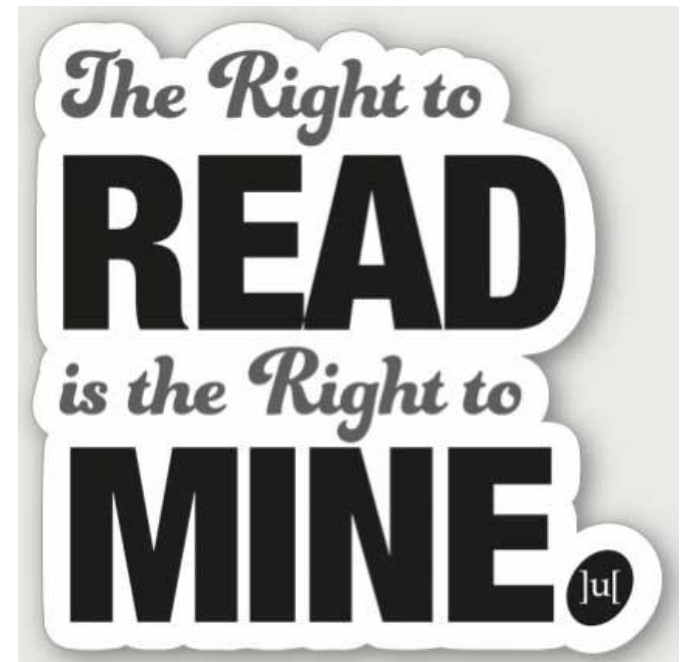


UK copyright exception

Allows researchers to make copies of copyright material for TDM without obtaining any additional permissions from the rights holder

- Copying has to be for text and data mining
- Have to have lawful access to content
- Has to be non-commercial research
- Can't re-sell the text or data mined

www.gov.uk/government/uploads/system/uploads/attachment_data/file/375954/Research.pdf



So, what do we (Europe) want?

- Legal clarity
 - A specific exception in EU law to allow TDM
 - A reinterpretation of EU law
- Legal interoperability
 - A solution at WIPO
- Open licences
 - CC-BY and CC0



NO

MAYBE

YES



Ligue des Bibliothèques Européennes de Recherche
Association of European Research Libraries

Slide courtesy of Susan Reilly of LIBER from presentation given at IFLA 2014 in Lyon



What do we not want?

- Licences for subscriptions which explicitly forbid machine crawling
- A licence with every single publisher for every single research project
- Publishers placing conditions on how TDM results are disseminated
- Click-through licences
- “Open access” licences that are NOT interoperable (STM model licences)



Storage: use of the cloud

University responses to the use of Dropbox

- Provide guidance on when to use third-party cloud services and what for

e.g.

www.lse.ac.uk/intranet/LSEServices/IMT/guides/softwareGuides/other/usingDropboxCloudStorageServices.aspx

- Provide greater quantities of research data storage so researchers have enough space



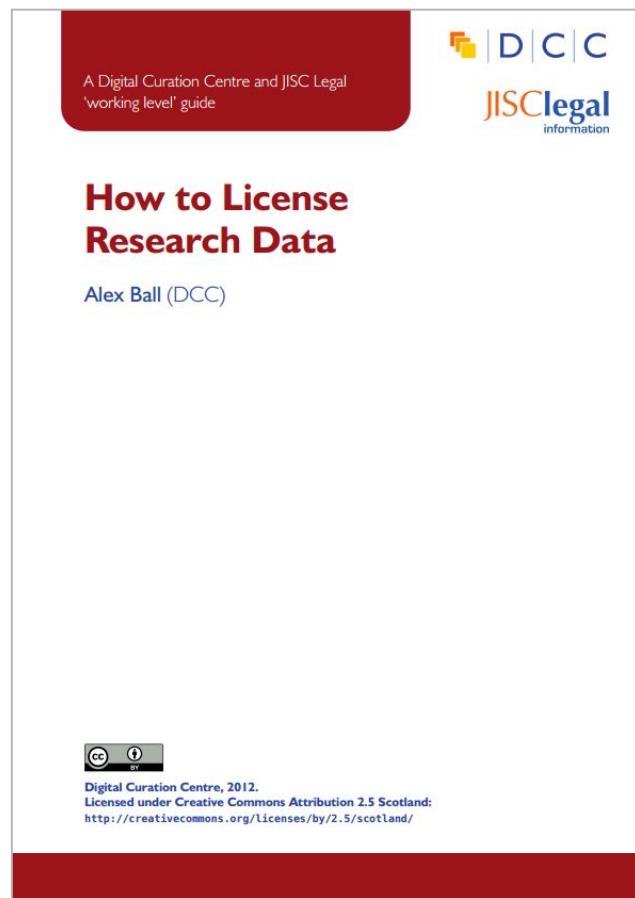
- Offer open source alternatives

e.g. OwnCloud pilot at Lincoln Uni <http://tiny.cc/owncloud-pilot>

Selecting which data to keep

- What is required by research funders?
- Legal obligations to retain data as evidence?
- Have consent agreements promised data destruction?
- Interests of commercial partners or consortium?
- Possibility for sharing and data reuse?
- Does researcher have rights to grant open licenses?

Licensing data for reuse



Outlines pros and cons of each approach and gives practical advice on how to implement data licences

CREATIVE COMMONS LIMITATIONS



NC

Non-Commercial

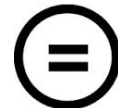
What counts as commercial?



SA

Share Alike

Reduces interoperability

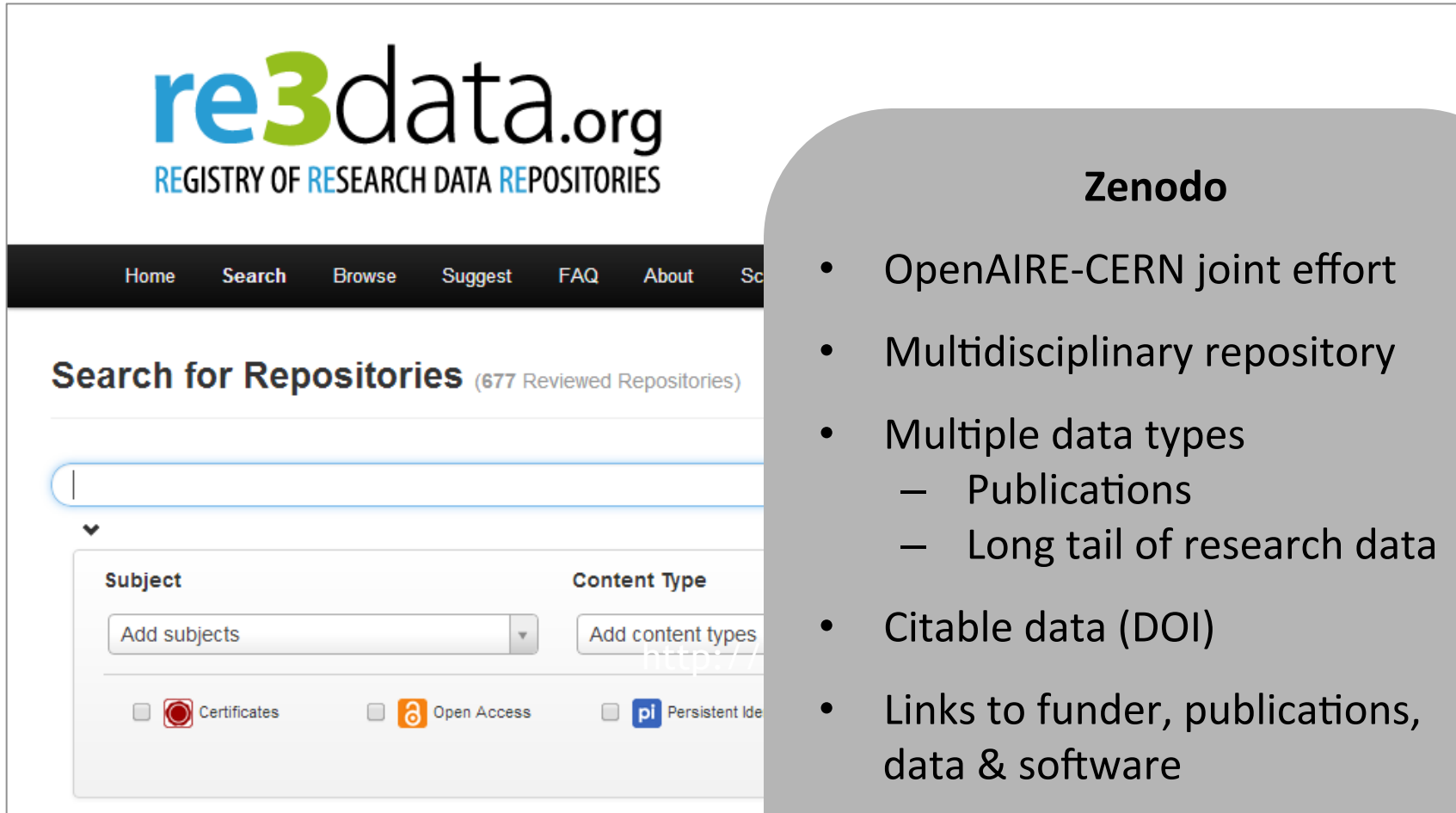


ND

No Derivatives

Severely restricts use

Suggesting data repositories



The screenshot shows the re3data.org website, which is a registry of research data repositories. The header includes the logo "re3data.org" and the text "REGISTRY OF RESEARCH DATA REPOSITORIES". A navigation bar contains links for Home, Search, Browse, Suggest, FAQ, About, and Sc. Below the navigation bar, there is a search section titled "Search for Repositories (677 Reviewed Repositories)". This section features a search input field, a dropdown menu, and two filter sections: "Subject" with an "Add subjects" dropdown and "Content Type" with an "Add content types" dropdown. At the bottom of the filters, there are three checkboxes: "Certificates" (checked), "Open Access", and "Persistent Id".

Zenodo

- OpenAIRE-CERN joint effort
- Multidisciplinary repository
- Multiple data types
 - Publications
 - Long tail of research data
- Citable data (DOI)
- Links to funder, publications, data & software

www.zenodo.org

<http://service.re3data.org/>

Cataloguing data: the Research Data Registry and Discovery (RDRDS) pilot

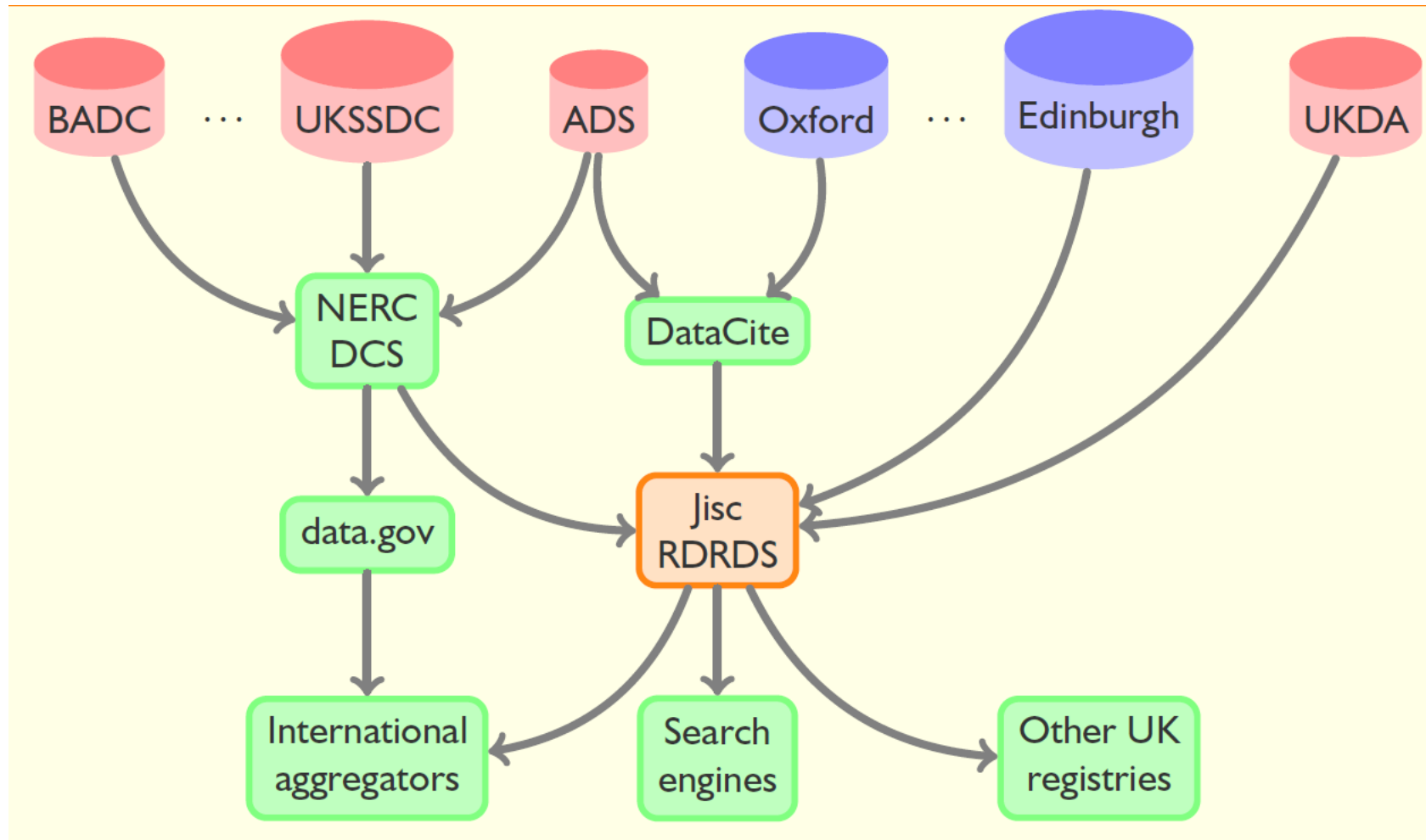
The screenshot shows the RDRDS pilot website interface. At the top, there is a dark navigation bar with the text "Research Data Registry & Discovery" in orange and white, and a menu with links for "About", "Collections", "Parties", "Activities", "Services", and "Themes". To the right of the menu is the "D|C|C" logo. Below the navigation bar is a light green search bar with the placeholder text "Search for Research Data" and a magnifying glass icon. To the right of the search bar are two buttons: "Browse by Subject Area" with a tree icon and "Browse by Map Coverage" with a map of Australia icon. Below the search bar is a link for "Advanced Search".

The main content area is divided into two columns. The left column is titled "What's in the Research Data Registry and Discovery Service" and contains four categories, each with an icon and a description:

- Collections (49)**: Research datasets or collections of research materials. (Icon: folder)
- Parties (36)**: Researchers or research organisations that create or maintain research datasets or collections. (Icon: two people)
- Activities (0)**: Projects or programs that create research datasets or collections. (Icon: flask)
- Services (0)**: Services that support the creation or use of research datasets or collections. (Icon: laptop and gears)

The right column is titled "Spotlight on research data" and features a large orange-bordered box containing the URL <http://rdrds.cloudapp.net/>. Below this is a section titled "Who contributes to the Research Data Registry and Discovery Service?" which states that 5 research organisations from around the UK contribute information to Research Data Australia. A "See All" link is provided. At the bottom right of this section is a "Share" button with a green icon.

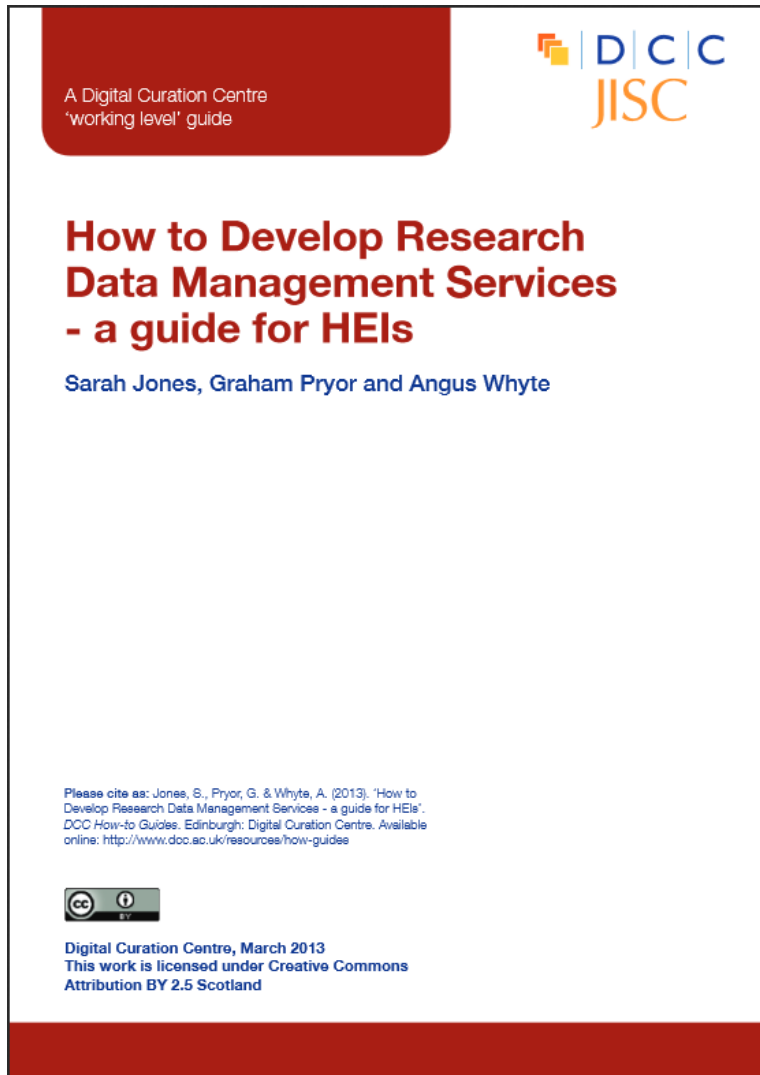
Where does RDRDS fit in?



How can libraries support RDM?

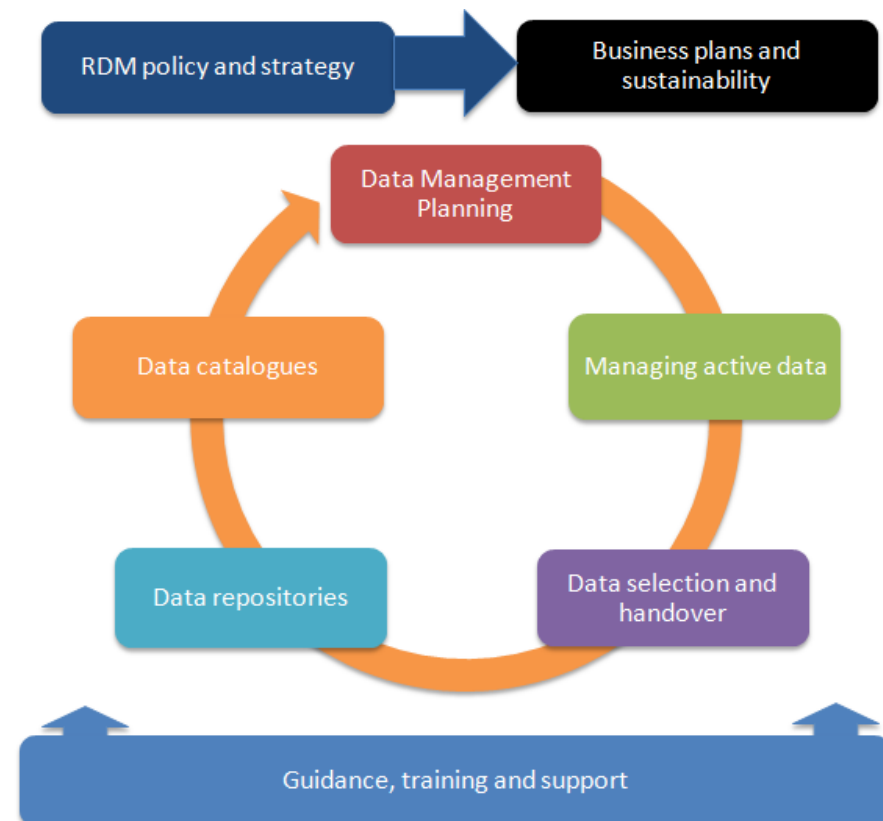
- Lobby for copyright reform to enable TDM
 - the right to read is the right to mine
- Help researchers to clarify data ownership & rights
- Catalogue data holdings to improve discoverability
- Ensure data are licensed and have DOIs so they can be reused and cited

Developing RDM services



Guide and case studies:

www.dcc.ac.uk/resources/developing-rdm-services



“Researchers need help to manage their data. This is a really exciting opportunity for libraries....”

Liz Lyon, VALA 2012



CC-BY by domeheid
[www.flickr.com/
photos/domeheid/](http://www.flickr.com/photos/domeheid/)

Merci!
Thank you!

DCC guidance, tools and case studies:

www.dcc.ac.uk/resources

Follow us on twitter:

@digitalcuration and #ukdcc



D|C|C

because good research needs good data