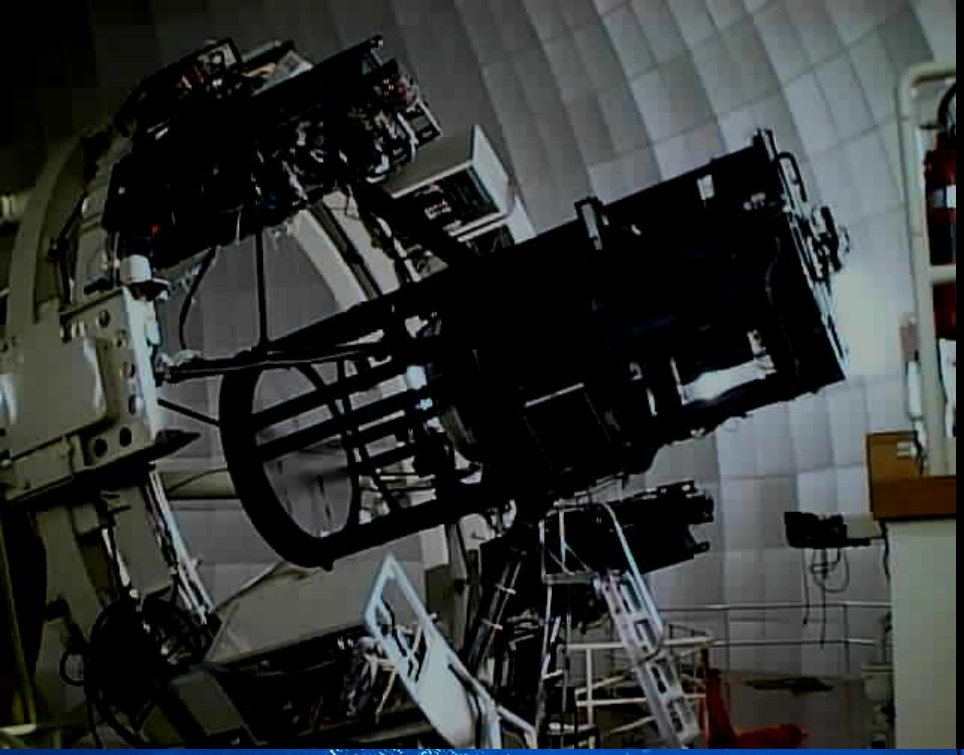
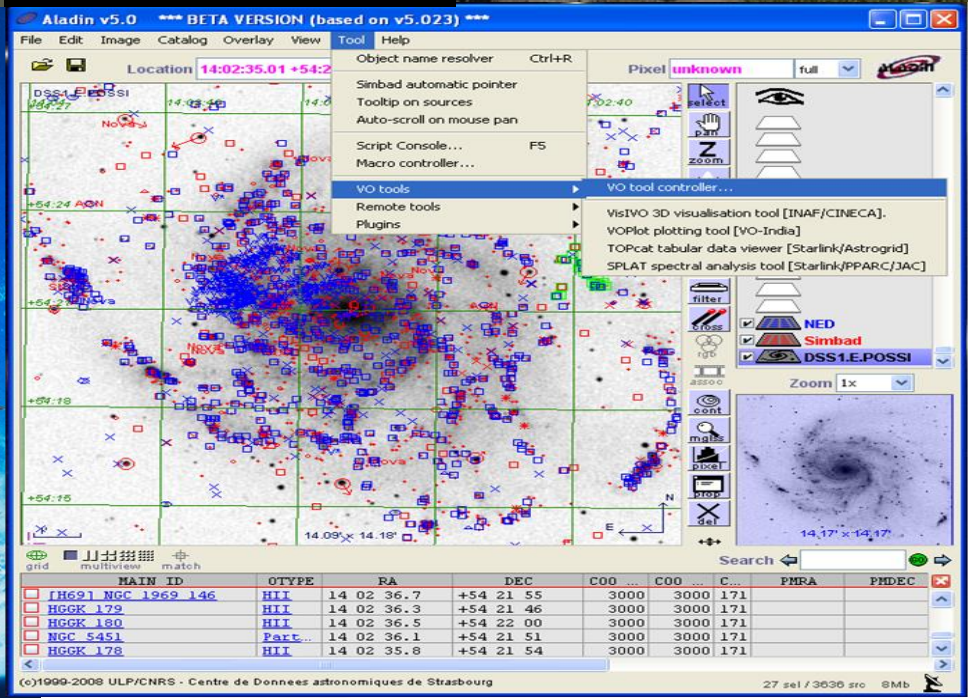
PUBLISHING RESEARCH DATA LANDSCAPE

Dr Jonathan Tedds jat26@le.ac.uk @jtedds

Senior Research Fellow,
Director: Health And Research Data Informatics
Department of Health Sciences, University of Leicester

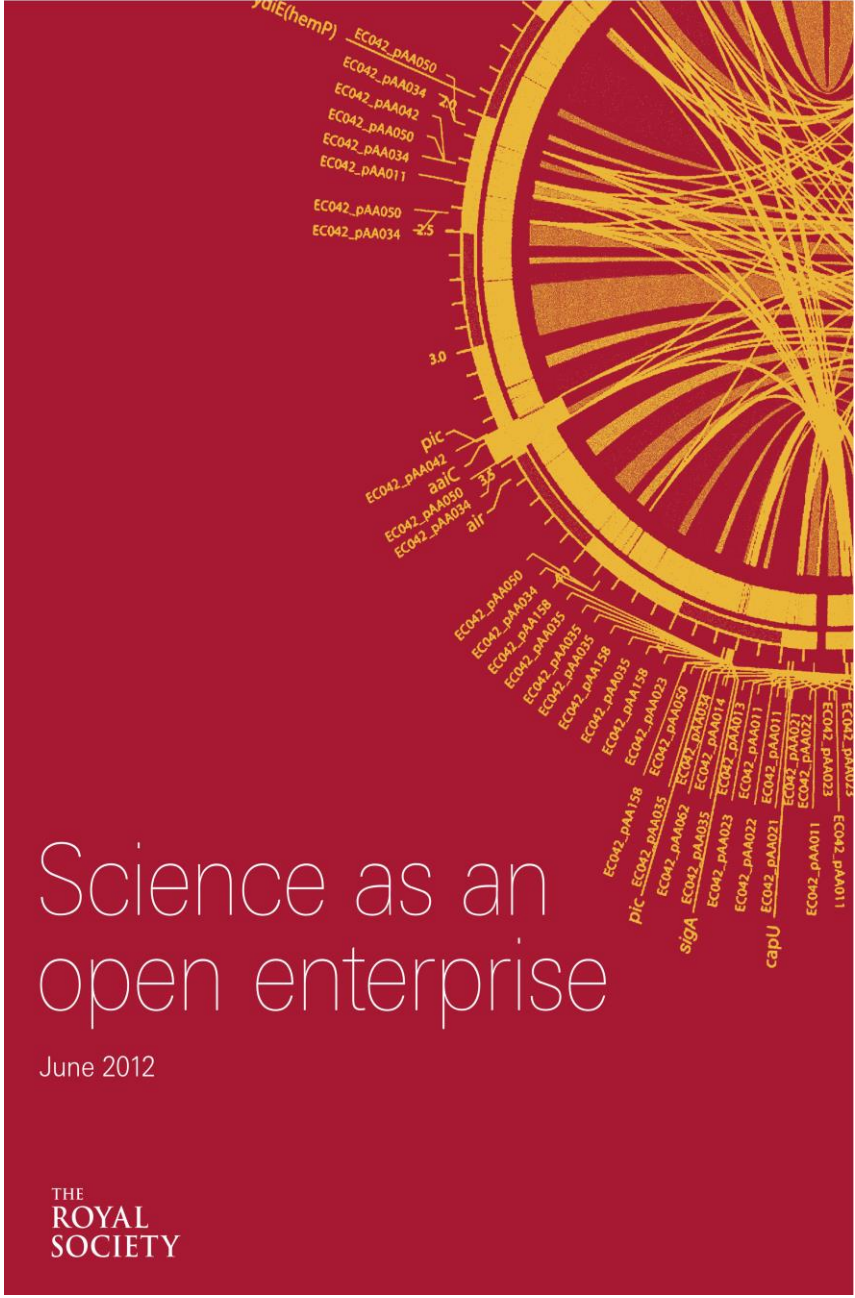
PI #PREPARDE http://www.le.ac.uk/projects/preparde
Editor-in-Chief, Open Health Data Journal
Co-Chair Research Data Alliance – WDS Publishing Data IG
and proposed Workflows WG



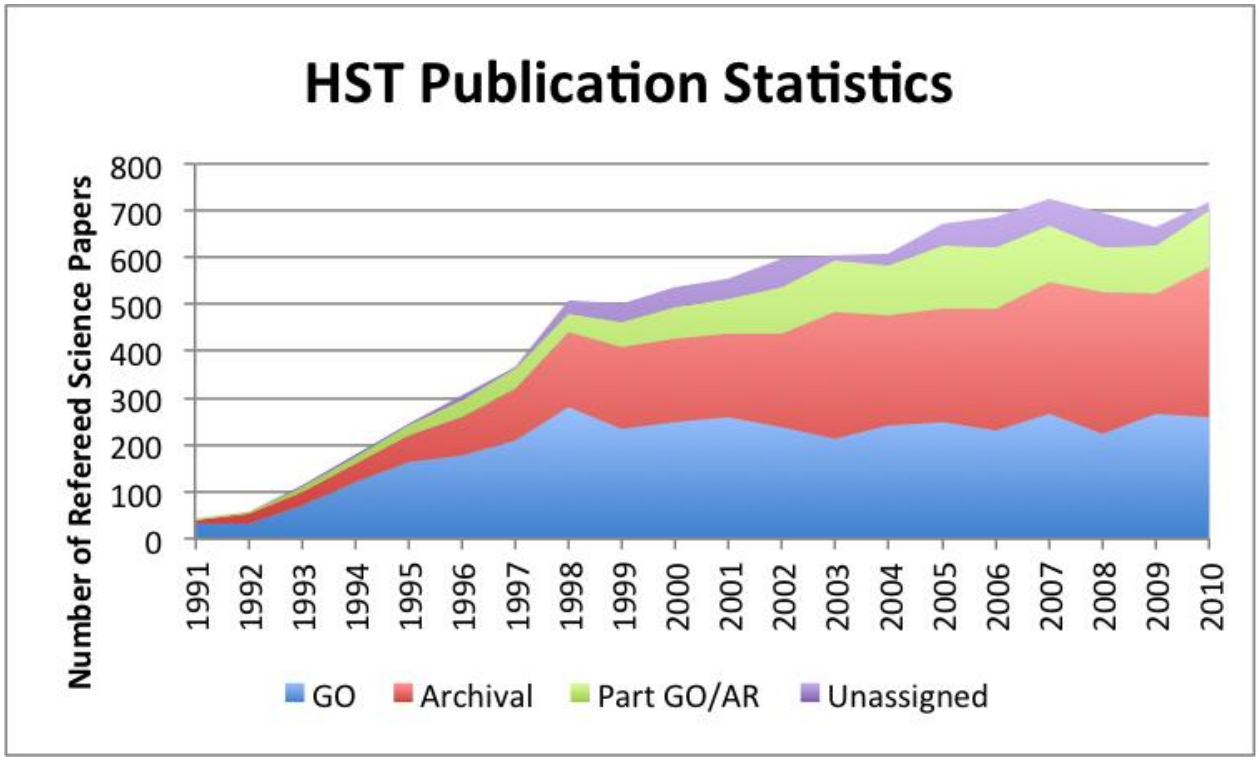

Aladin v5.0 *** BETA VERSION (based on v5.023) ***
File Edit Image Catalog Overlay View Tool Help
Object name resolver Ctrl+R
Simbad automatic pointer
Tooltip on sources
Auto-scroll on mouse pan
Script Console... F5
Macro controller...
VO tools
Remote tools
Plugins
VO tool controller...
VisIVO 3D visualisation tool [INAF/CINECA].
VOPlot plotting tool [VO-India]
TOPcat tabular data viewer [Starlink/Astrogrid]
SPLAT spectral analysis tool [Starlink/PPARC/JAC]
Location 14:02:35.01 +54:2
Pixel unknown full
NED
Simbad
DSS1.E.POSSI
Zoom 1x
Search
MAIN ID OTYPE RA DEC C00 ... C00 ... C... PMRA PMDEC
[H69] NGC 1969 146 HII 14 02 36.7 +54 21 55 3000 3000 171
HGGK 179 HII 14 02 36.3 +54 21 46 3000 3000 171
HGGK 180 HII 14 02 36.5 +54 22 00 3000 3000 171
NGC 5451 Part... 14 02 36.1 +54 21 51 3000 3000 171
HGGK 178 HII 14 02 35.8 +54 21 54 3000 3000 171
(c)1999-2008 ULP/CNRS - Centre de Donnees astronomiques de Strasbourg
27 sel / 3636 src 8Mb

Why open?

- **As a first step towards this intelligent openness, data that underpin a journal article should be made concurrently available in an accessible database**
- We are now on the brink of an achievable aim: for all science literature to be online, for all of the data to be online and for the two to be interoperable. [p.7]
- Royal Society June 2012, *Science as an Open Enterprise*, http://royalsociety.org/policy/projects/science-public-enterprise/report/
- Issues linking data to the scientific record:
  - Data persistence
  - Data and metadata quality
  - **Attribution and credit for data producers**
- Geoffrey Boulton (Edinburgh), Lead author:
  - "Science has been sleepwalking into crisis of replicability...and of the credibility of science"
  - "Publishing articles without making the data available is *scientific malpractice*"



Data Reuse: asking new questions



Hubble Space Telescope

- Papers based upon reuse of archived observations now exceed those based on the use described in the original proposal
  - http://archive.stsci.edu/hst/bibliography/pubstat.html
- See also work by Piwowar & Vision re life sciences: “Data reuse and the open data citation advantage”
  - http://peerj.com/preprints/1/

Oh, and…. G8 UK says so :P

G8 UK
UNITED KINGDOM 2013

We are committed to openness in scientific research data to speed up the progress of scientific discovery, create innovation, ensure that the results of scientific research are as widely available as practical, enable transparency in science and engage the public in the scientific process.

- To the greatest extent and with the fewest constraints possible publicly funded scientific research data should be open, while at the same time respecting concerns in relation to privacy, safety, security and commercial interests, whilst acknowledging the legitimate concerns of private partners.

- Open scientific research data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards.

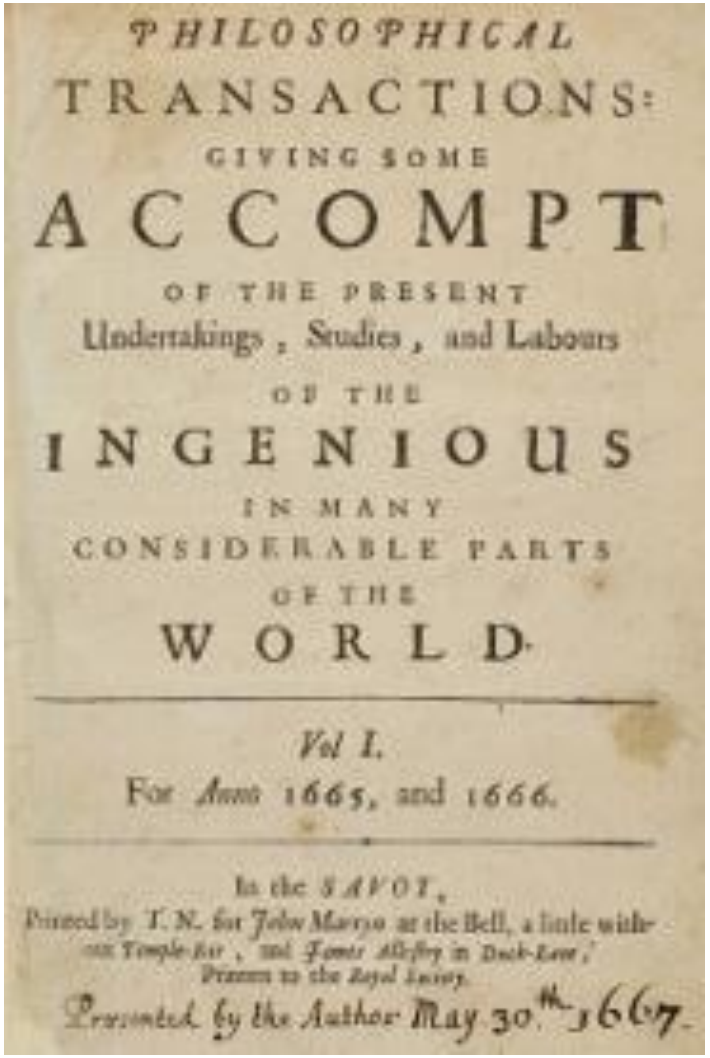
- To ensure successful adoption by scientific communities, open scientific research data principles will need to be underpinned by an appropriate policy environment, including recognition of researchers fulfilling these principles, and appropriate digital infrastructure.

Scale of the problem: who, what, when where….?

http://blogs.scientificamerican.com/absolutely-maybe/2013/09/10/opening-a-can-of-data-sharing-worms/

- Timothy Vines and colleagues studied reproducibility of data sets in zoology and changes through time
  - gathered 516 papers published between 1991 and 2011
  - then they tried to track the data down…
- Even tracking down the authors was a challenge
  - Over time a dwindling minority of papers were accompanied by author email addresses that still functioned
- only 37% of the data - even from papers in 2011 - were still findable and retrievable
  - proportion dropped each earlier year
- For papers published in 1991
  - only 7% of the data could be determined to truly still be in existence and retrievable
  - few authors could be found, and most of them were reporting that their data were lost or inaccessible

This isn't new…


PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little within
the Temple-Bar, and James Allestry in Duck-Lane,
Printers to the Royal Society.
Presented by the Author May 30th 1667.

Henry Oldenburg

- inveterate correspondent
- now think of as scientist

- Had idea to publish Philosophical Transactions (1665):
  - Should be written in vernacular not Latin
  - Underlying evidence must be concurrently published
  - Helped propel Europe at the time
  - Concept of scientific self correction
    - able to write it's errors
- Wrote: "thought fit to employ the [printing] press……..Universal Good of Mankind"
  - How do we achieve these ends in the post-Gutenburg era?

Data as a "public good" (2011)

RESEARCH COUNCILS UK

Excellence with Impact

Accessibility

Research and funding

Research Careers

Public Engagement with Research

Knowledge Exchange and Impact

International

Press and Media

Publications

About

Home

Research and Funding

Research Funding

Areas of Research

Cross-Council Research Themes

Research Infrastructure

Research Priorities

Peer review

Eligibility for Research Council funding

How to apply for research funding

Applications which may cross Research Council remits

Terms and Conditions of Research Council fEC Grants

Terms and Conditions of Research Council Training Grants

Open Access

RCUK Common Principles on Data Policy

Efficiency

Home > Research and Funding > RCUK Common Principles on Data Policy

RCUK Common Principles on Data Policy

Making research data available to users is a core part of the Research Councils' remit and is undertaken in a variety of ways. We are committed to transparency and to a coherent approach across the research base. These RCUK common principles on data policy provide an overarching framework for individual Research Council policies on data policy.

Principles

- Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property.
- Institutional and project specific data management policies and plans should be in accordance with relevant standards and community best practice. Data with acknowledged long-term value should be preserved and remain accessible and usable for future research.
- To enable research data to be discoverable and effectively re-used by others, sufficient metadata should be recorded and made openly available to enable other researchers to understand the research and re-use potential of the data. Published results should always include information on how to access the supporting data.
- RCUK recognises that there are legal, ethical and commercial constraints on release of research data. To ensure that the research process is not damaged by inappropriate release of data, research organisation policies and practices should ensure that these are considered at all stages in the research process.
- To ensure that research teams get appropriate recognition for the effort involved in collecting and analysing data, those who undertake Research Council funded work may be entitled to a limited period of privileged use of the data they have collected to enable them to publish the results of their research. The length of this period varies by research discipline and, where appropriate, is discussed further in the published policies of individual Research Councils.
- In order to recognise the intellectual contributions of researchers who generate, preserve and share key research datasets, all users of research data should acknowledge the sources of their data and abide by the terms and conditions under which they are accessed.
- It is appropriate to use public funds to support the management and sharing of publicly-funded research data. To maximise the research benefit which can be gained from limited budgets, the mechanisms for these activities should be both efficient and cost-effective in the use of public funds.

Search:

Go

This website

All Research Councils

- Public good
- Preservation
- Discovery
- Confidentiality
- First use
- Recognition
- Public funding

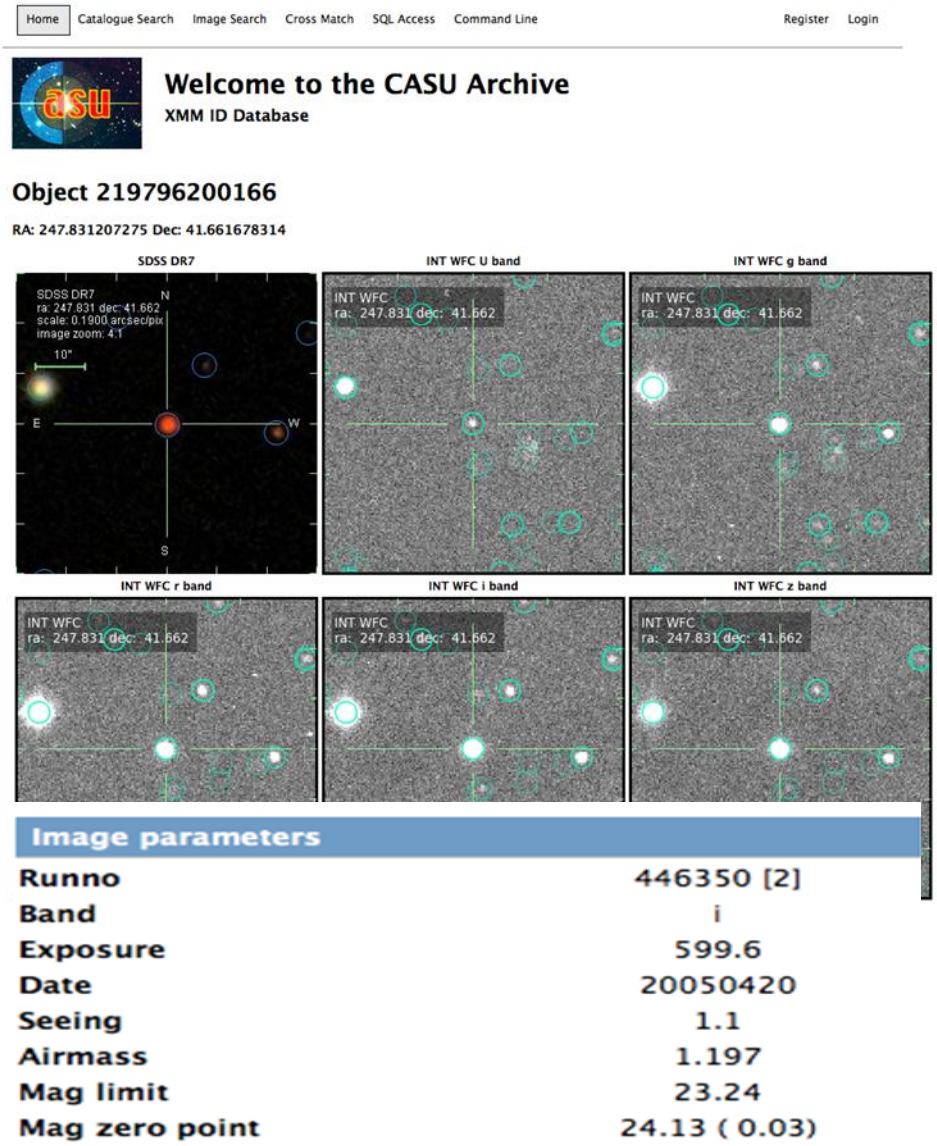
So what do we mean by publishing data?

- The familiar:
  - Supplementary tables via journal or
  - Archived raw or calibrated facility data
  - Discipline specific and institutional / national archives
- Data under the graph?
  - In order to reproduce and adapt article analysis
- “Research ready” open data
  - In order to reuse and repurpose
  - for interdisciplinary researchers, community, business
  - Ideally peer reviewed?

Research data example - level 1:

- A typical example from physical sciences (**astronomy**) distinguishes between broad categories within the research data spectrum:

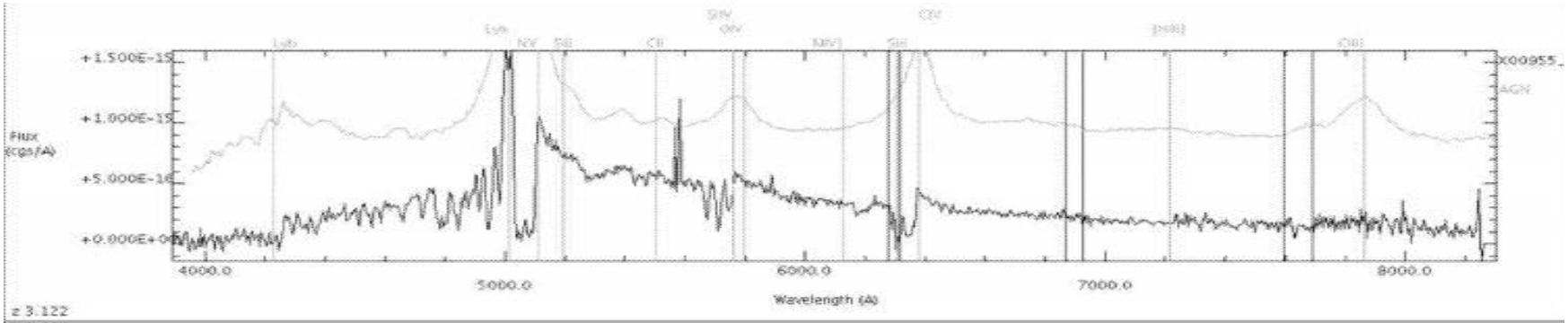
- **raw/initially auto-processed data** produced at a research facility such as an observatory
  - typically made publically available in this format after an embargo period of e.g. 1 year
  - in some cases available immediately - e.g. Swift Gamma Ray Burst satellite



Research data example – level 2

"research ready" processed data which has been fully calibrated, combined and cleaned/annotated

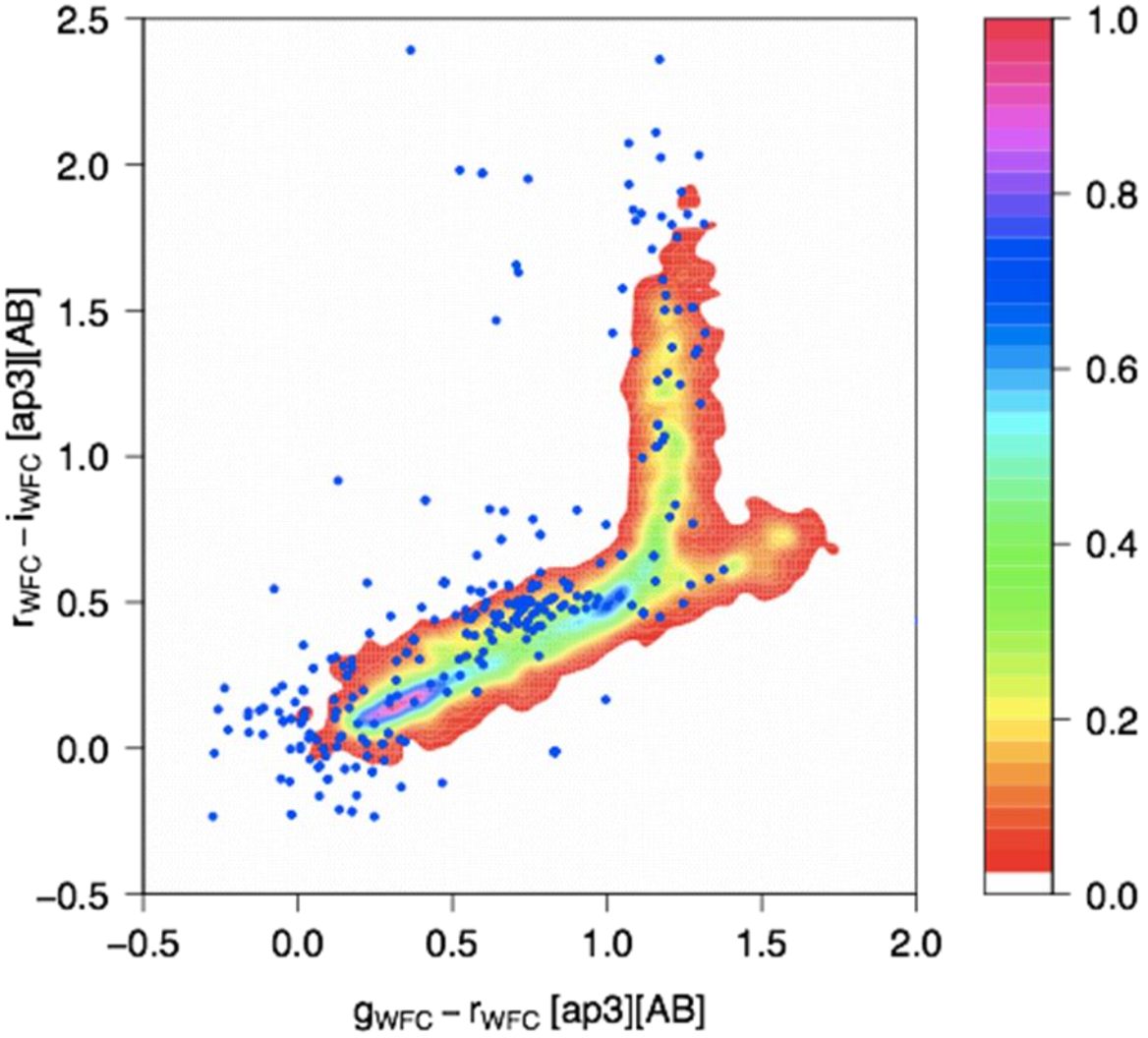
- often produced by individuals or collaborations
- rarely available to anyone outside the collaboration except upon request/collaboration
- needed for re-analysis or reuse for science unless you have detailed sub domain specific knowledge and detailed contextual information to reproduce from raw
- considered to enable a competitive advantage for producers
- may be produced by dedicated data scientists on behalf of the community for major survey/missions e.g. ESA XMM-Survey Science Centre (Leicester), NASA, NCAR…



Research data example – level 3

output dataset – following detailed analysis of research ready datasets

- **forms the *data under the graph* in a journal publication** following analysis of research ready datasets
- Might be available as a Table via journal, CDS etc
- *May not be available* outside the collaboration except upon request/collaboration
- may well generate future additional samples and papers for the owning collaboration on top of the original
- other researchers may request the data for their own research but may not get it!



….and STOP!

- Next project
  - Proposal long since written
  - Probably already underway…
- Feel free to email ME if you would like to work on an idea using this dataset or code
  - As long as I'm a co-author on the paper!
  - You have to go through me to find out what you really need to know to reuse the data/code

Research data example – level 4a

- **published catalogue** type representation of published output dataset

  - *NOT a "data paper"….but could be*
  - optional in many cases, mandatory for most major surveys
  - usually made available via project specific online resource
  - may be provided as table of parameters based on research ready dataset, usually linked from and associated with a journal
  - specifically produced in order for the wider community to reuse (cite!) and repurpose if wanted
  - The well-known Sloan Digital Sky Survey is a classic example or more recently the 2XMMi X-ray catalogue I have a close involvement with (largest X-ray survey of the sky).

e.g. **The *XMM-Newton* Wide Angle Survey (XWAS)**

P. Esquej[1,2,3,4], M. Page[5], F. J. Carrera[3], S. Mateos[3], J. Tedds[2], M. G. Watson[2], A. Corral[6], J. Ebrero[7], M. Krumpe[8,9,10], S. R. Rosen[2], M. T. Ceballos[3], A. Schwope[10], C. G. Page[2], A. Alonso-Herrero[3⋆], A. Caccianiga[6], R. Della Ceca[6], O. González-Martín[11], G. Lamer[10], P. Severgnini[6]

[1] Centro de Astrobiología (INTA-CSIC), ESAC Campus, PO Box 78, 28691 Villanueva de la Cañada, Spain
[2] Dept. of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, U.K.
[3] Instituto de Física de Cantabria (CSIC-UC), Avda. de los Castros, 39005 Santander, Spain
[4] Departamento de Física Moderna, Universidad de Cantabria, Avda. de Los Castros, 39005 Santander, Spain
[5] Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK
[6] INAF – Osservatorio Astronomico di Brera, via Brera 28, 20121 Milan, Italy
[7] SRON - Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA, Utrecht, The Netherlands
[8] European Southern Observatory, Karl-Schwarzschild-Straβe 2, 85748 Garching bei München, Germany
[9] University of California, San Diego, Center for Astrophysics & Space Sciences, 9500 Gilman Drive, CA 92093-0424, USA
[10] Leibniz-Institute for Astrophysics Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
[11] Instituto Astrofísico de Canarias, (IAC), C/Vía Láctea, s/n, E-38205, La Laguna, Tenerife, Spain

Received ??, 2012; accepted ??, 2013

ABSTRACT

Aims. This programme is aimed at obtaining one of the largest X-ray selected samples of identified active galactic nuclei to date in order to characterise such a population at intermediate fluxes, where most of the Universe's accretion power originates. We present the *XMM-Newton* Wide Angle Survey (XWAS), a new catalogue of almost a thousand X-ray sources spectroscopically identified through optical observations.

Methods. A sample of X-ray sources detected in 68 *XMM-Newton* pointed observations was selected for optical multi-fibre spectroscopy. Optical counterparts and corresponding photometry of the X-ray sources were obtained from the SuperCOSMOS Sky Survey. Candidates for spectroscopy were initially selected with magnitudes down to $R \sim 21$, with preference for X-ray sources having a flux $F_{0.5-4.5\,\mathrm{keV}} \geq 10^{-14}$ erg s^{-1} cm^{-2}. Optical spectroscopic observations were made using the Two Degree Field of the Anglo Australian Telescope, and the resulting spectra were classified based on optical emission lines.

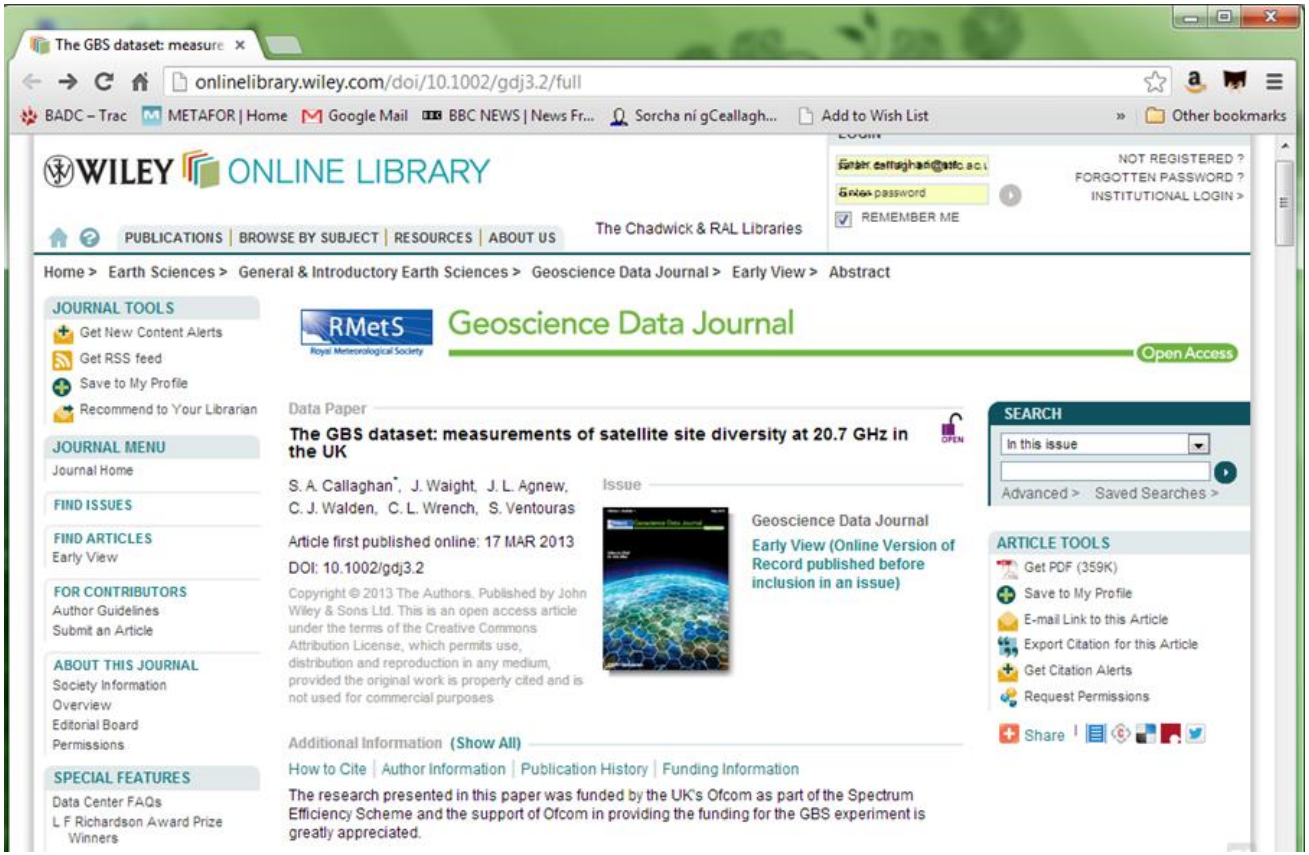
Results. We have identified through optical spectroscopy 940 X-ray sources over $\Omega \sim 11.8$ deg^2 of the sky. Source populations in our sample can be summarised as 65% broad line active galactic nuclei (BLAGN), 16% narrow emission line galaxies (NELGs), 6% absorption line galaxies (ALGs) and 13% stars. An active nucleus is also likely to be present in the large majority of the X-ray sources spectroscopically classified as NELGs or ALGs. Sources lie in high-galactic latitude ($|b| > 20$ deg) *XMM-Newton* fields mainly in the southern hemisphere. Owing to the large parameter space in redshift ($0 \leq z \leq 4.25$) and flux ($10^{-15} \leq F_{0.5-4.5\,\mathrm{keV}} \leq 10^{-12}$ erg s^{-1} cm^{-2}) covered by the XWAS this work provides an excellent resource for the further study of subsamples and particular cases. The overall properties of the extragalactic objects are presented in this paper. These include the redshift and luminosity distributions, optical and X-ray colours and X-ray-to-optical flux ratios.

Key words. X-ray: general – Surveys – X-rays: galaxies – Galaxies: active

http://adsabs.harvard.edu/abs/2013arXiv1302.5329E

Research data example – level 4b

data paper describing and linking to output dataset(s)

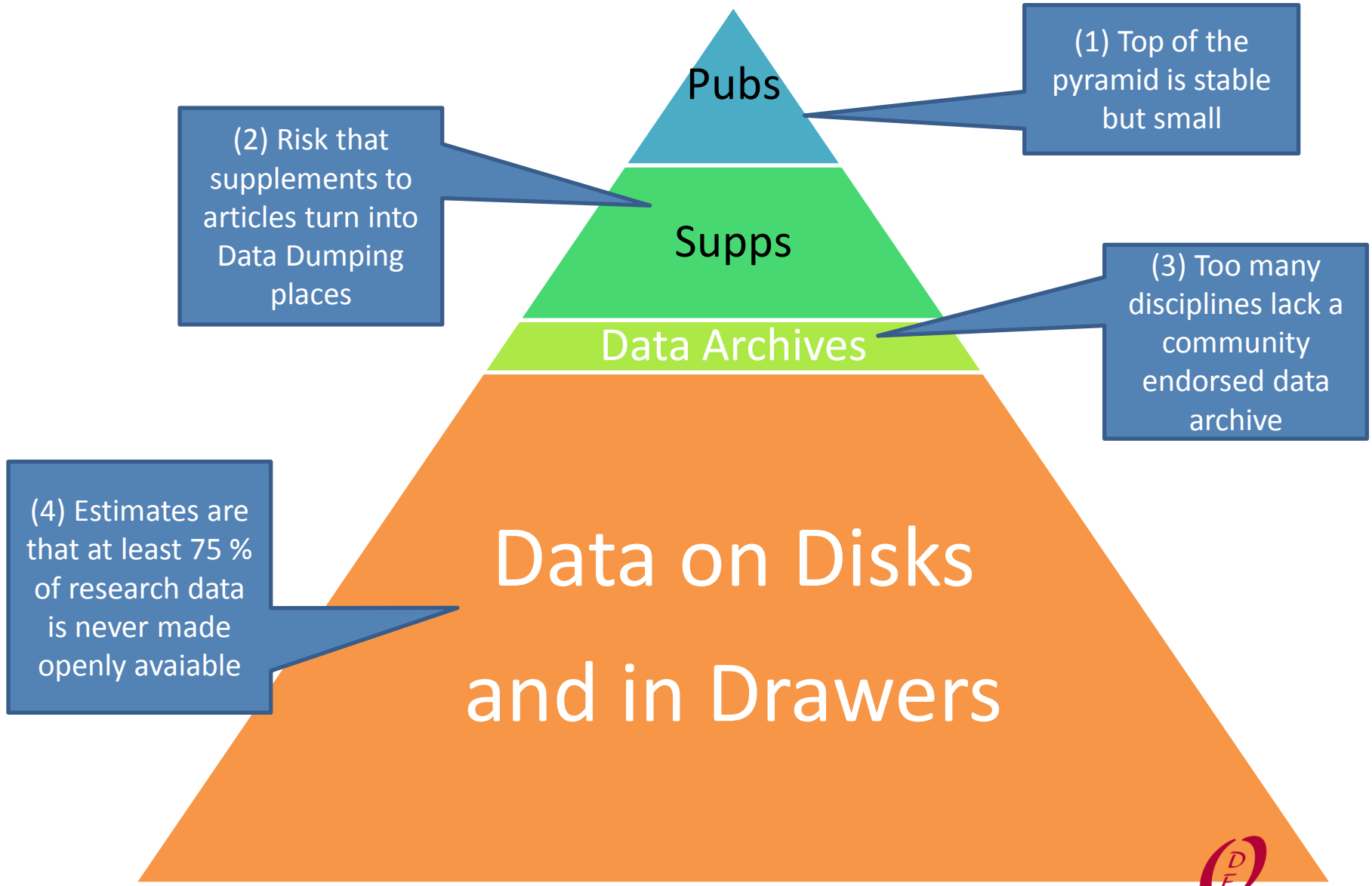


Live Data paper!

Dataset citation is first thing in the paper and is also included in reference list (to take advantage of citation count systems)

DOI: 10.1002/gdj3.2

Can't publish it all?
ODE Data Publication Pyramid:



Opportunities for Data Exchange

From Mayernik et al (submitted)

Most cited Bulletin of the American Meteorological Society (BAMS) articles. Data from Web of Science, Jun 11, 2013

Article	Data paper?	Citations	Article details
1	Yes	10,113	Kalnay, E; et al. The NCEP/NCAR 40-year reanalysis project, 1996.
2	No	3,201	Torrence, C; Compo, GP. A practical guide to wavelet analysis, 1998.
3	No	2,367	Mantua, NJ; et al. A Pacific interdecadal climate oscillation with impacts on salmon production, 1997.
4	Yes	1,987	Kistler, R; et al. The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, 2001.
5	Yes	1,791	Xie, PP; Arkin, PA. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, 1997.
6	Yes	1,448	Kanamitsu, M; et al. NCEP-DOE AMIP-II reanalysis (R-2), 2002.
7	No	1,014	Baldocchi, D; et al. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, 2001.
8	Yes	902	Rossow, WB; Schiffer, RA. Advances in understanding clouds from ISCCP, 1999.
9	Yes	900	Rossow, WB; Schiffer, RA. ISCCP cloud data products, 1991.
10	No	877	Hess, M; Koepke, P; Schult, I. Optical properties of aerosols and clouds: The software package OPAC, 1998.
11	No	815	Willmott, CJ. Some comments on the evaluation of model performance, 1982.
12	No	815	Trenberth, KE. The definition of El Nino, 1997.
13	Yes	785	Woodruff, SD; Slutz, RJ; et al. A comprehensive ocean-atmosphere data set, 1987.
14	Yes	776	Meehl, G.A.; et al. The WCRP CMIP3 multimodel dataset - A new era in climate change research, 2007.

PREPARDE: Peer REview for Publication & Accreditation of Research Data in the Earth sciences

Jonathan Tedds (Leicester), Sarah Callaghan (BADC), Fiona Murphy (Wiley), Rebecca Lawrence (F1000R), Matthew Mayernik (NCAR), John Kunze, Carly Strasser (CDL), Angus Whyte (DCC), Becca Wilson (Leicester), Simon Hodson (Jisc/CODATA) and #PREPARDE project team

+ Geraldine Clement Stoneham (MRC), Elizabeth Newbold, Rachel Kotarski (BL) on data peer review

http://www.le.ac.uk/projects/preparde

JISC University of Leicester British Atmospheric Data Centre NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE NATURAL ENVIRONMENT RESEARCH COUNCIL WILEY-BLACKWELL University of Reading DCC University of California CDL California Digital Library

F1000 FACULTY of 1000 POST-PUBLICATION PEER REVIEW NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

PREPARDE: Peer REview for Publication & Accreditation of Research Data in the Earth sciences http://www.le.ac.uk/projects/preparde

- **capture the processes and procedures required to publish a scientific dataset**
  - ingestion into a data repository
  - formal publication in a data journal
- address **key issues in data publication**
  - how to peer-review a dataset?
  - what criteria are needed for a repository to be considered objectively trustworthy?
  - how can datasets and journal publications be effectively cross-linked for the benefit of the wider research community?
- PREPARDE team includes **key expertise** in
  - Research
  - academic publishing
  - data management
- **Earth Sciences focus** but produce general guidelines applicable to a wide range of scientific disciplines and data publication types incl life sciences (F1000R)

JISC University of Leicester British Atmospheric Data Centre NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE NATURAL ENVIRONMENT RESEARCH COUNCIL WILEY-BLACKWELL University of Reading DCC University of California CDL California Digital Library

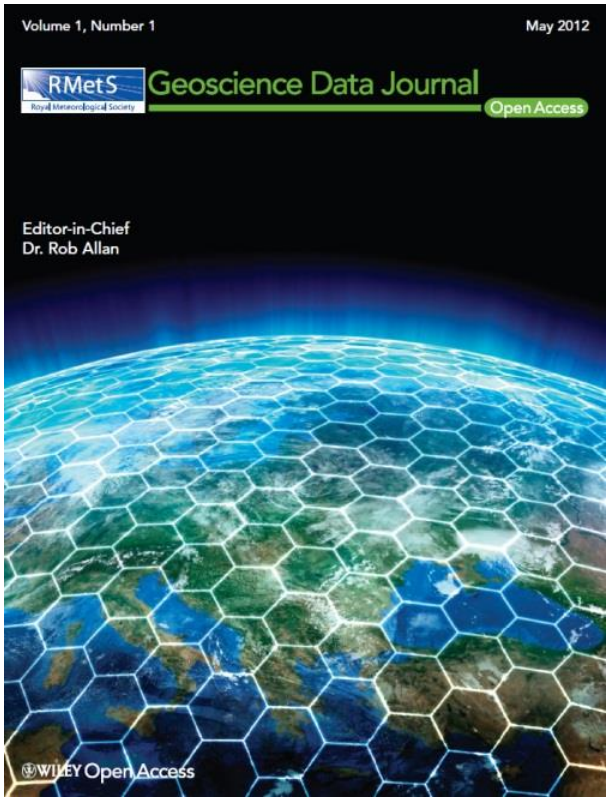
F1000 FACULTY of 1000 POST-PUBLICATION PEER REVIEW NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

PREPARDE key use case: *Geoscience Data Journal*, Wiley-Blackwell and the Royal Meteorological Society

- Partnership formed between Royal Meteorological Society and academic publishers Wiley Blackwell to develop a mechanism for the formal publication of data in the Open Access *Geoscience Data Journal*

- GDJ publishes short data articles cross-linked to, and citing, datasets that have been deposited in approved data centres and awarded DOIs (or other permanent identifier).

- A data article describes a dataset, giving details of its collection, processing, software, file formats, etc., without the requirement of novel analyses or ground breaking conclusions.
  - the when, how and why data was collected and what the data-product is.

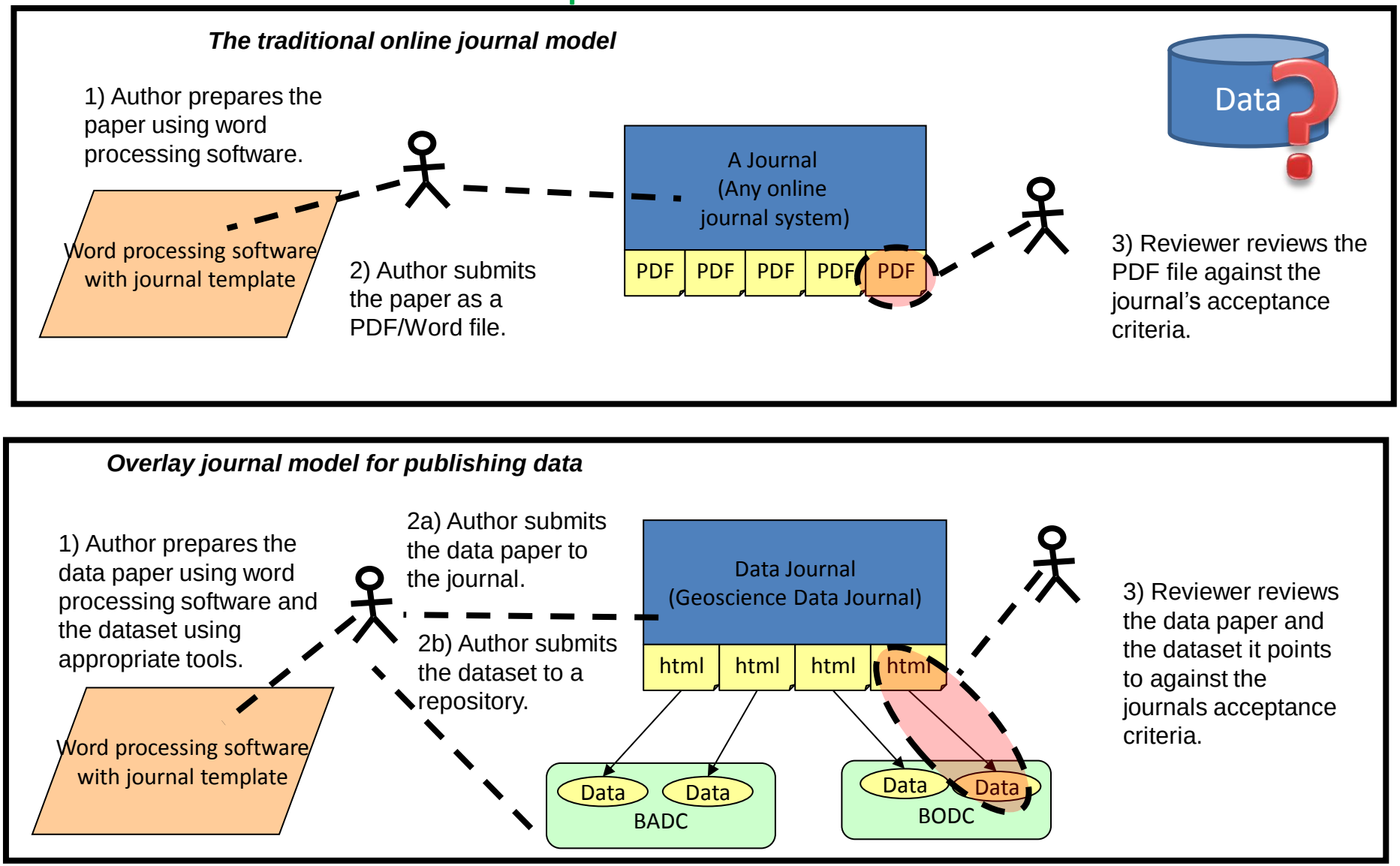


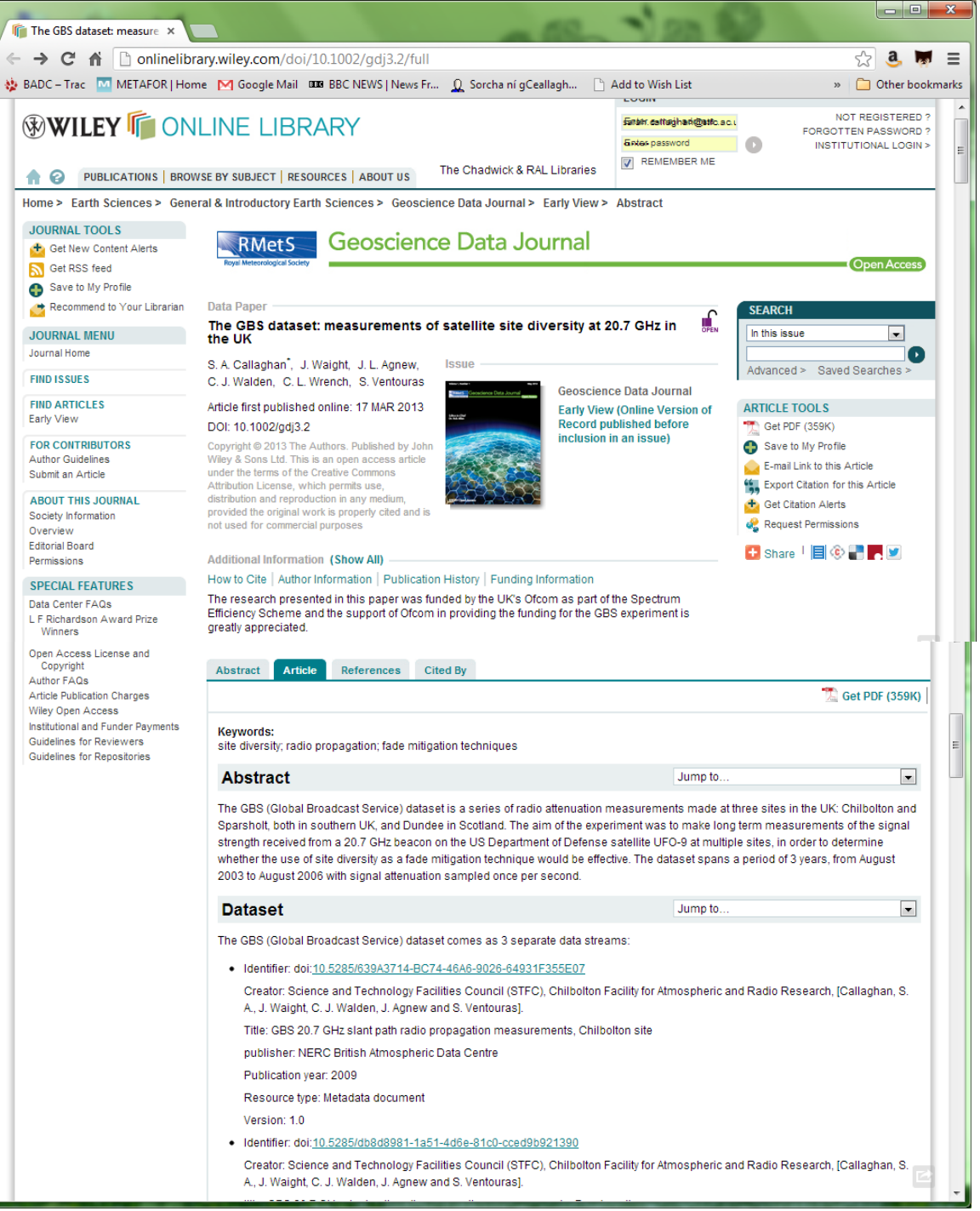
RMetS
Royal Meteorological Society

WILEY-BLACKWELL

http://www.geosciencedata.com/

How: to publish data in GDJ

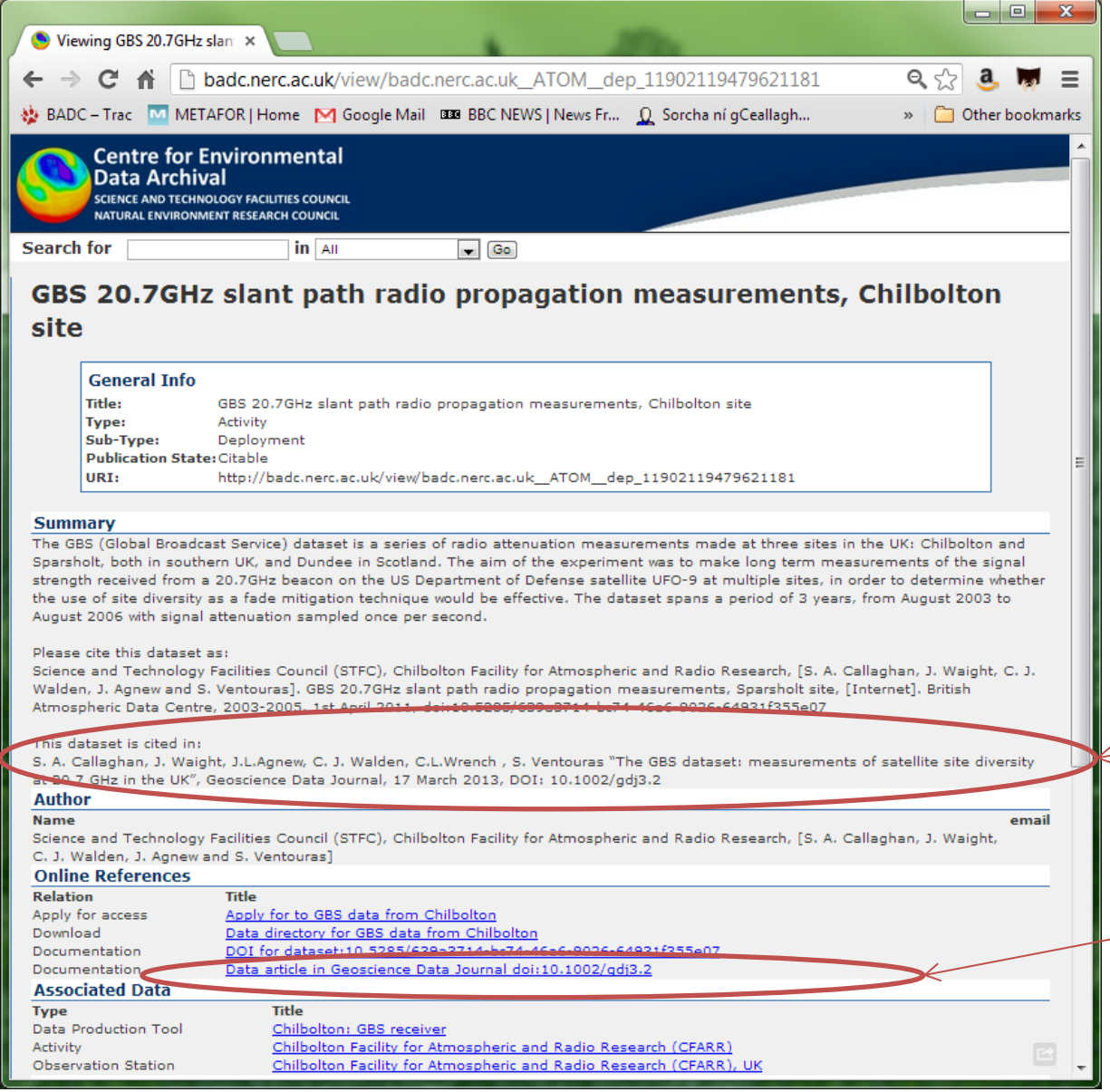




Live Data paper!

Dataset citation is first thing in the paper and is also included in reference list (to take advantage of citation count systems)

DOI: 10.1002/gdj3.2



Dataset catalogue page (and DOI landing page)

Reference to Data Article

Clickable link to Data Article

JISC
University of Leicester
British Atmospheric Data Centre
NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL
WILEY-BLACKWELL
University of Reading
DCC
University of California CDL California Digital Library
F1000Research
NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

UPmetajournals

open health data

ubiquity press open access

Open Health Data is a peer-reviewed journal publishing papers that describe openly-accessible datasets with high reuse potential. We work with trusted data repositories to ensure that associated datasets are professionally archived, preserved, and openly available.

Open Health Data was previously published as the *Journal of Open Public Health Data*. The title and scope of the journal were amended

Most recent articles

Nagulapalli
Data from "Social determinants of unmet hospitalisation need...
04 Nov 2013

Most popular articles

Alexander, Wint
Projected Population Proximity Indices (30km) for 2005, 2030 & 2050
11 Jul 2013

It’s a long road….

What do researchers need to make this all possible?

- Incentives - citations, promotion, support - long way to go
- Institutional and funder policy framework - mostly there now
- Appropriate discipline specific community centres of expertise - rare, mostly limited to big science niches or very broad but poorly sustained
- Institutional support services for the basics - pilots to date
- Software tools that are open and can be adapted - on the way

So institutions alone aren't enough – we need:

RDA
RESEARCH DATA ALLIANCE

Research Data Sharing without barriers

Home | About | Organisation | Working and Interest Groups | News & Events | Plenary Meetings | RDA in the Press

Home » Interest Groups » Publishing Data IG

Publishing Data IG

Status: Recognised & Endorsed

The Publishing Data Interest Group brings together all stakeholders involved in publishing research data including researchers, discipline specific and institutional data repositories, academic publishers, funders and service providers. Every effort will be made to get a good representation from related international programmes, their working groups and other private or institutional activities involved in this area. We will build on existing resources, reports and other shared experiences from the different stakeholders and will nurture more specific and targeted working groups addressing practical aspects in publishing research data. As such, the Publishing Data Interest Group can be regarded as a broad and inclusive forum for interested individuals to contribute to and test, validate and promote the findings of the Working Groups. In particular we plan to address the implementation of workflows for publishing data and therefore help establish appropriate supporting infrastructure.

Download the Publishing Data Interest Group Charter

Group visibility

Public - accessible to all site users

Group menu

- Group News Archive
- Group Wiki
- Old forum (Read-only)
- Public File Repository
- Public Discussion Area
- RDA Mailing lists

https://rd-alliance.org/internal-groups/publishing-data-ig.html

Accepted Research Data Alliance **Interest Group**

Publishing Data

- **Close coordination with ICSU-WDS working group, CODATA, FORCE11, DataCite and other ongoing initiatives** in data publication
  - WDS under International Council of Science, RDA wider
  - Avoid duplication within related RDA and WDS WGs – join up
  - For WDS partnerships between publishers and data centres key
- scope the territory – gap analysis & test implementations
- Use RDA Forum and build on http://jiscmail.ac.uk/data-publication 350+ list
- Take findings from RDA / WDS group(s) and trial in other communities / disciplines / institutional repositories
- Four proposed RDA-WDS Working Groups:
  - https://rd-alliance.org/data-publishing-2020-four-case-statements.html

RDA/WDS: Publishing Data Interest Group

Co-chairs: Michael Diepenbroek[1], Eefke Smit[2], Jonathan Tedds[3]

Email: jat26@le.ac.uk. [1]PANGAEA, University of Bremen, Germany [2]STM Publishers, The Netherlands [3]University of Leicester, UK

Introduction

In the empirical sciences, data has traditionally been an integral part of scholarly publishing. In recent decades rapid technical developments, such as digital data and high-throughput techniques, have dramatically altered the scholarly publishing paradigm. This requires new approaches in order to ensure the availability and usability of research data.

Various technical solutions in use or proposed to date offer promise but do not yet provide sufficient benefit and incentives for the data producers themselves and so take up among researchers is still relatively low. The concept of Publishing Data is undergoing a renaissance as part of scholarly communication and on the base of new and proven technologies such as establishing persistent identifiers for datasets. Publishing data offers a strong incentive for researchers to share their data and benefits the wider community through a focus on data quality.

The impact on citation rates is beginning to be demonstrated through bibliometric studies of research articles that include underlying data or are based on secondary reuse of existing datasets such as in astronomy.

The Publishing Data Interest Group brings together all stakeholders involved in publishing research data including researchers, discipline specific and institutional data repositories, academic publishers, funders and service providers. The following 4 initial Working Groups are being developed through the RDA in patnership with the ICSU-WDS and are currently developing Case Statements under the umbrella of the Publishing Data Interest Group. New Working Groups can be formed or join the Interest group as it develops.

Workflows for Archiving and Publishing Data

Jonathan Tedds, Kim Finney, John Helly, Hylke Koers, Fiona Murphy, Amy Nurnberger, Lisa Raymond, Mary Vardigan, Eva Zanzerkia

- Investigate current workflows for archiving and publishing data
- The role of QA/QC and peer-review in the publication process
- The role of science publishers/journals in the data publication process
- Barriers to implementation

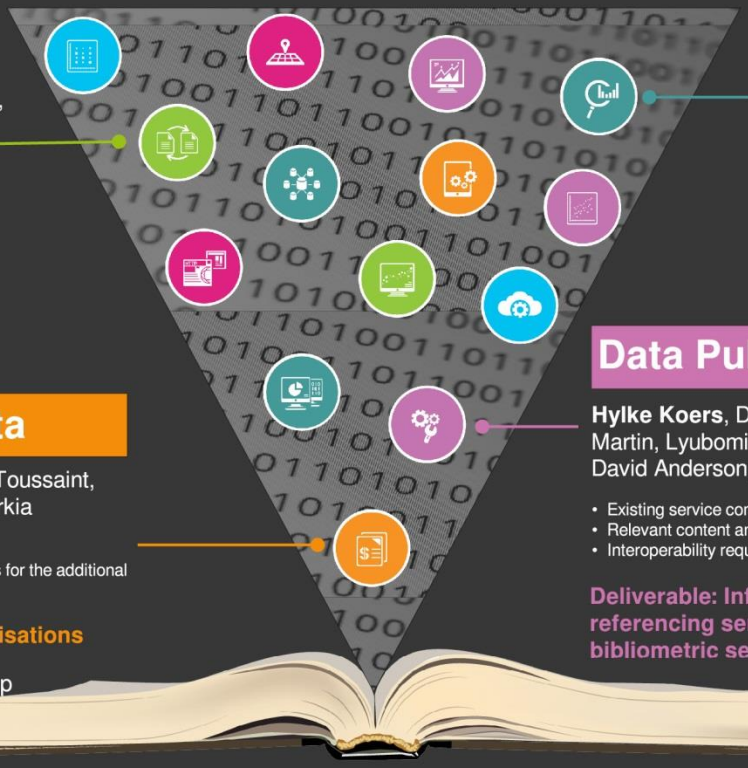
Deliverable: Provide a range of generic and discipline specific workflows for data publication identifying roles, resources and stakeholders

The Costs of Publishing Data

Ingrid Dillo, Simon Hodson, Barbara Sierman, Frank Toussaint, Mark Thorley, Kim Finney, Anita de Waard, Eva Zanzerkia

- Investigate current cost strutures for archiving and publishing data
- Elaborate a business model based on open access which compensates for the additional costs due to data publication

Deliverable: Recommendations for funding organisations



Bibliometrics Including Published Data

Kerstin Lehnert, Euan Adie, Jan Brase, Ross Cameron, Cyndy Chandler, Ingeborg Meijer, Fiona Murphy, Lyubomir Penev, Fiona Nielsen, Nigel Robinson, Mary Vardigan

- General requirements for citability of scientific data (granularity, citation information and persistent identifiction)
- Current citation practice in data centres and literature

Deliverable: Recomendations for data publishers and academic publishers

Data Publication Services

Hylke Koers, David Carlson, John Helly, Francisco Hernandez, Caroline Martin, Lyubomir Penev, Nigel Robinson, Johanna Schwarz, Eva Zanzerkia David Anderson, Juanle Wang

- Existing service components to be used as building blocks
- Relevant content and interoprability standards
- Interoperability requirements for data centres (registration, metadata and data services)

Deliverable: Infrastructure and organisation for a one-for-all cross referencing service for academic publishers and providers of bibliometric services

Scan to visit the Publishing Data Interest Group website: https://rd-alliance.org/internal-groups/publishing-data-ig.html

Scan to join the DATA-PUBLICATION mailing list: https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=DATA-PUBLICATION