

Institutional Role in Supporting Open Science

Asia Open Access, Dhaka -2019

March 7, 2019

Miho Funamori
National Institute of Informatics

Outline

1. What is Open Science?
2. Drivers and Responses
3. Role of University Libraries in Open Science Era

1. What is Open Science?

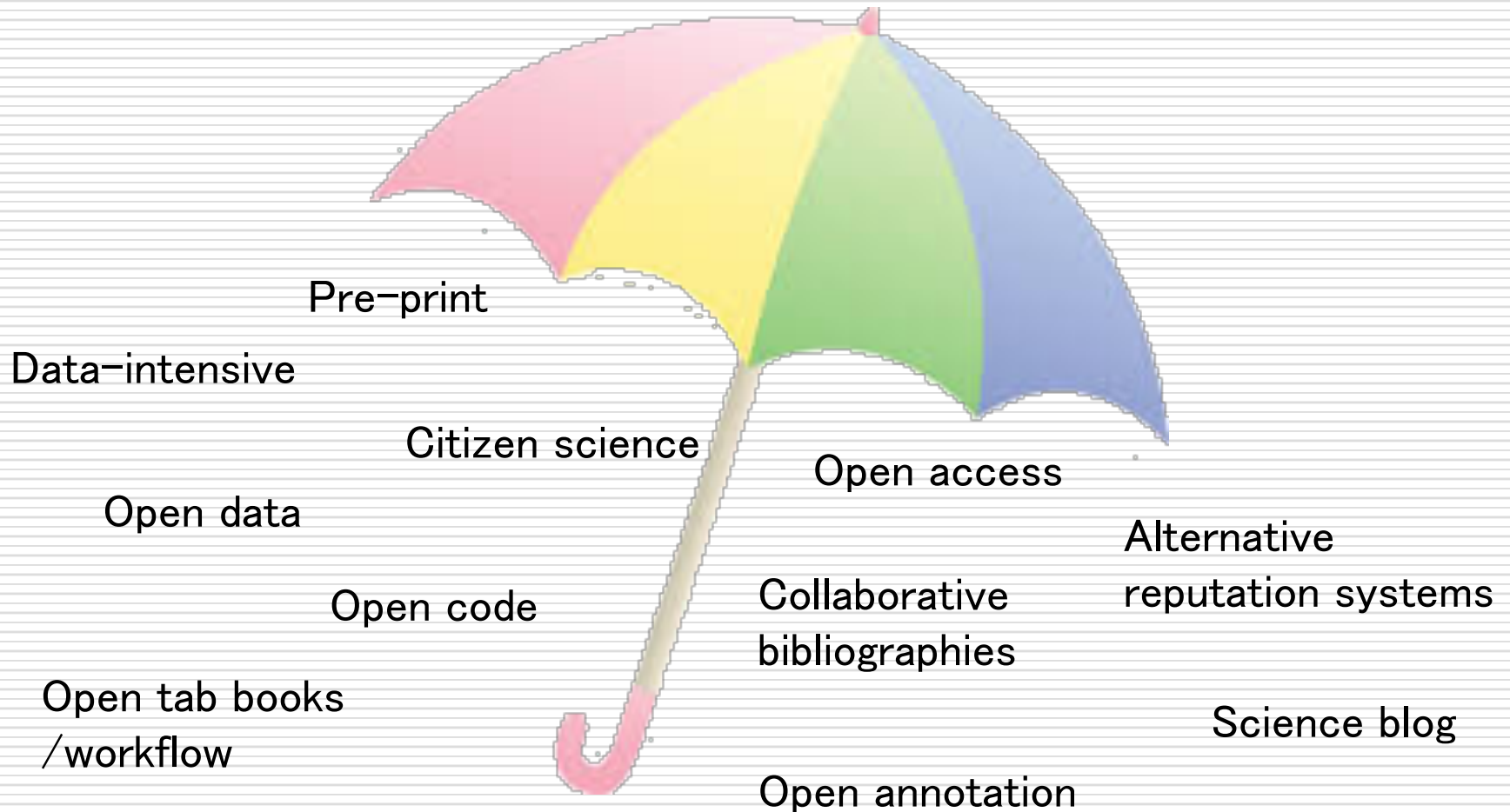
What is Open Science?



... Science has always been open!

Definition: Open Science

...Umbrella Term



Definition: Open Science

- ❑ Said to have no fixed definition
- ❑ General understanding:

- New ways of doing research and organizing science
- Enabled through digital technology
- Reshaping academic value systems

Open Science...European Commission (2014)

...Public Consultation 'Science 2.0': Science in Transition



EUROPEAN COMMISSION

DIRECTORATES-GENERAL FOR RESEARCH AND INNOVATION (RTD) AND
COMMUNICATIONS NETWORKS, CONTENT AND TECHNOLOGY (CONNECT)

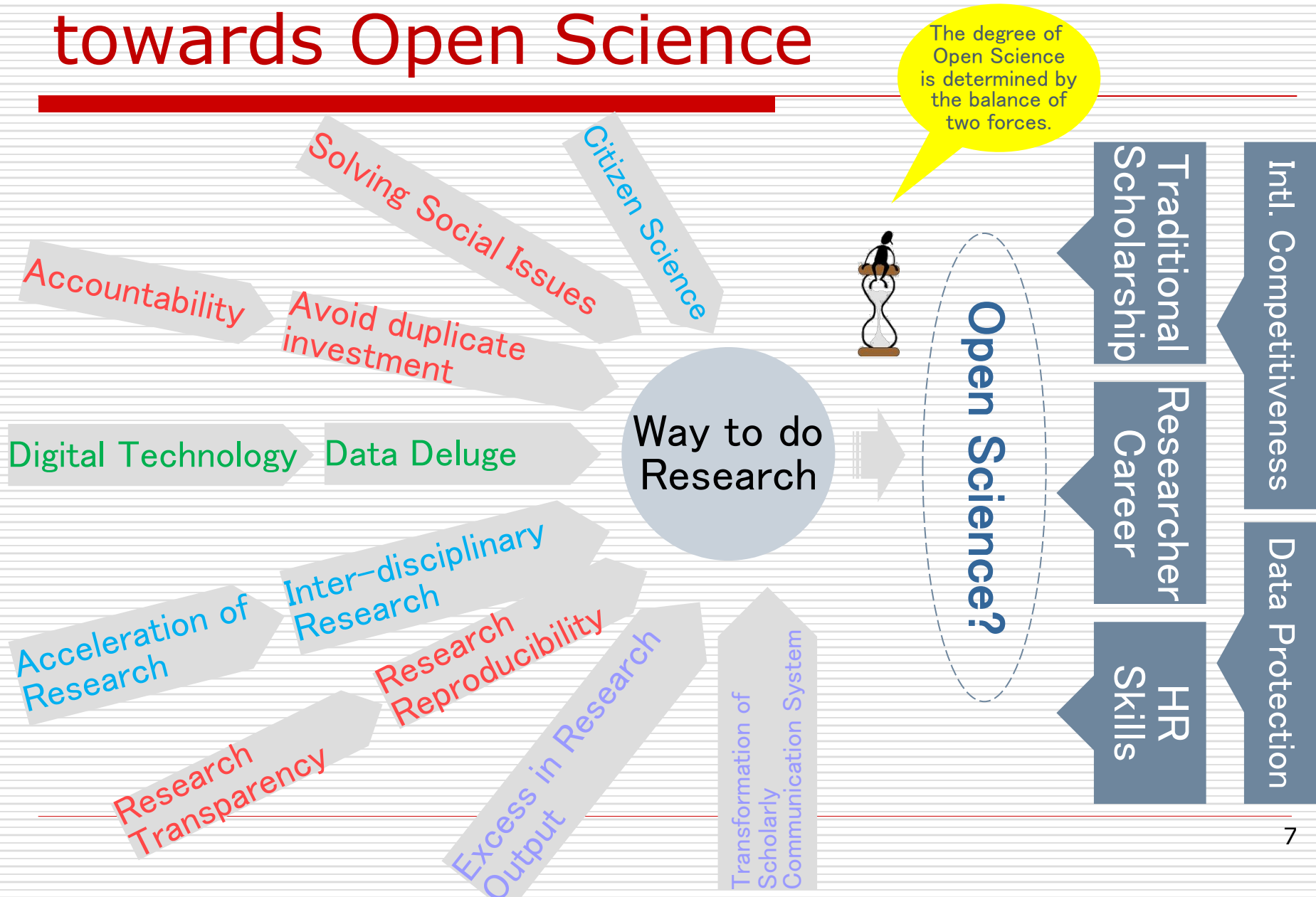
BACKGROUND DOCUMENT

PUBLIC CONSULTATION

'SCIENCE 2.0': SCIENCE IN TRANSITION

'Science 2.0' describes the **on-going evolution in the modus operandi of doing research and organising science**. These changes in the dynamics of science and research are enabled by **digital technologies** and driven by the **globalisation of the scientific community**, as well as the increasing societal demand to address the **Grand Challenges** of our times. They have an **impact on the entire research cycle**, from the inception of research to its publication, as well as on the way in which this cycle is organised.

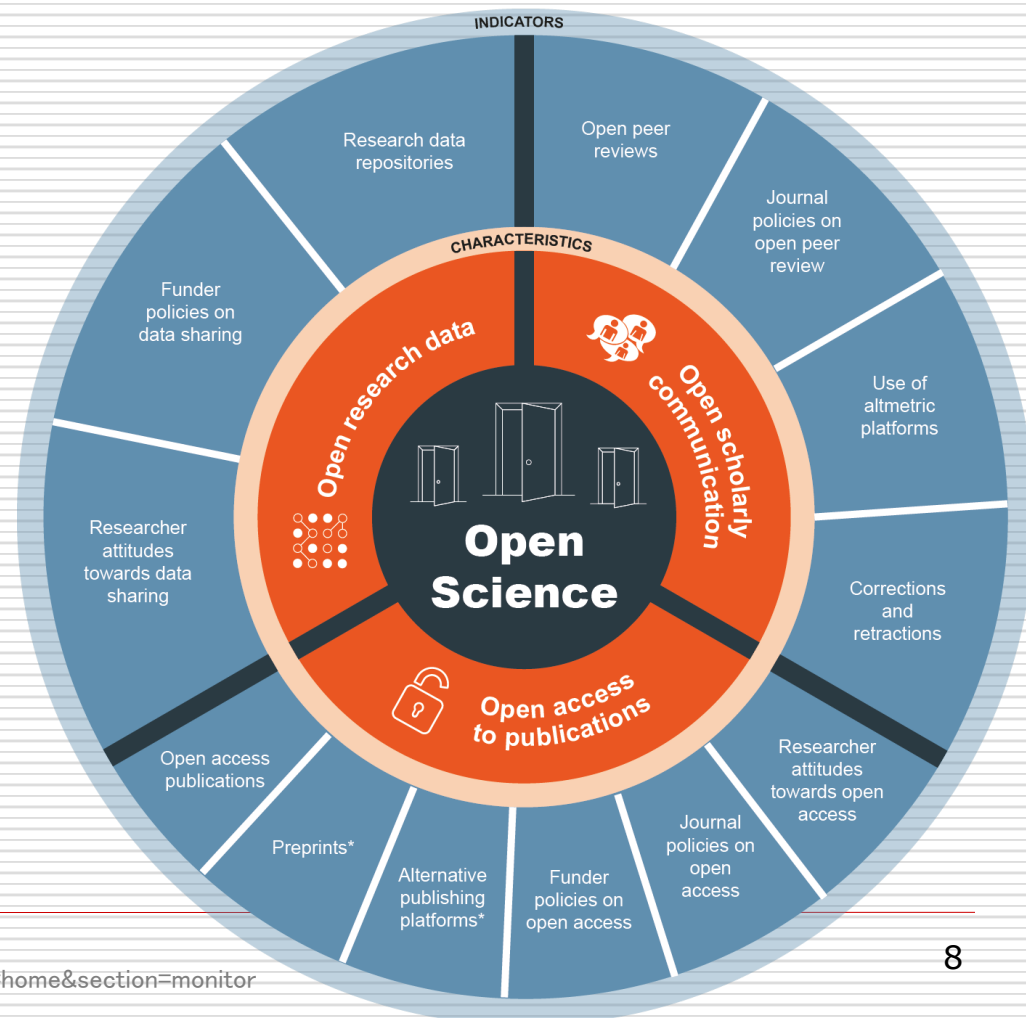
The push and resisting force towards Open Science



Open Science Monitor

- Includes open access to scientific results (publication and data)
- However, it is more than that...!
 - ✓ Researcher Attitude
 - ✓ Open Peer Review
 - ✓ Altmetrics
 - ✓ Correction and Retractions

Open Science Monitor



2-1. Drivers and Responses

Data-Intensive Scientific Discovery

The Fourth Paradigm: Data-Intensive Scientific Discovery

Tony Hey

Corporate Vice President

Microsoft External Research



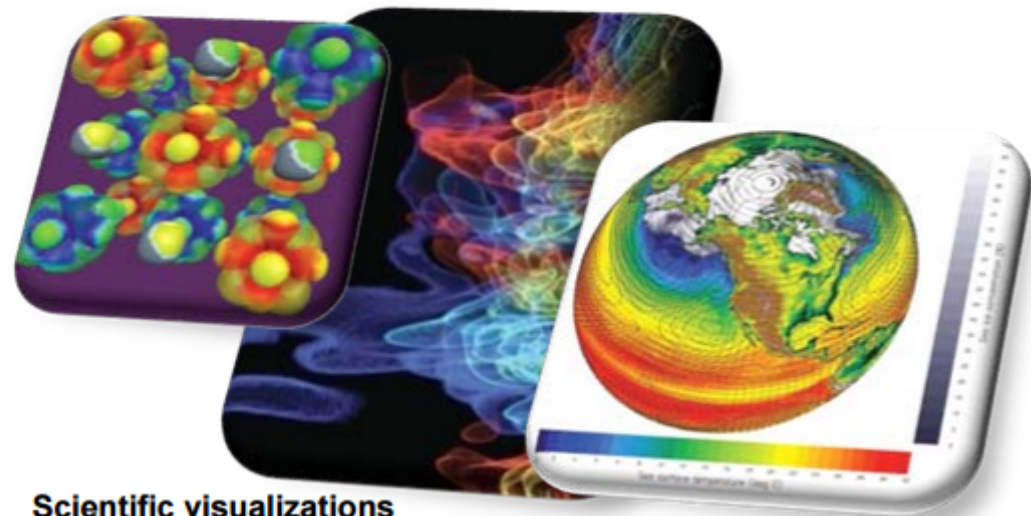
A Digital Data Deluge in Research

- Data collection
 - Sensor networks, satellite surveys, high throughput laboratory instruments, observation devices, supercomputers, LHC ...
- Data processing, analysis, visualization
 - Legacy codes, workflows, data mining, indexing, searching, graphics ...
- Archiving
 - Digital repositories, libraries, preservation, ...



SensorMap

Functionality: Map navigation
Data: sensor-generated temperature, video camera feed, traffic feeds, etc.



Scientific visualizations

NSF Cyberinfrastructure report, March 2007



This work is licensed under a [Creative Commons Attribution 3.0 United States License](https://creativecommons.org/licenses/by/3.0/).

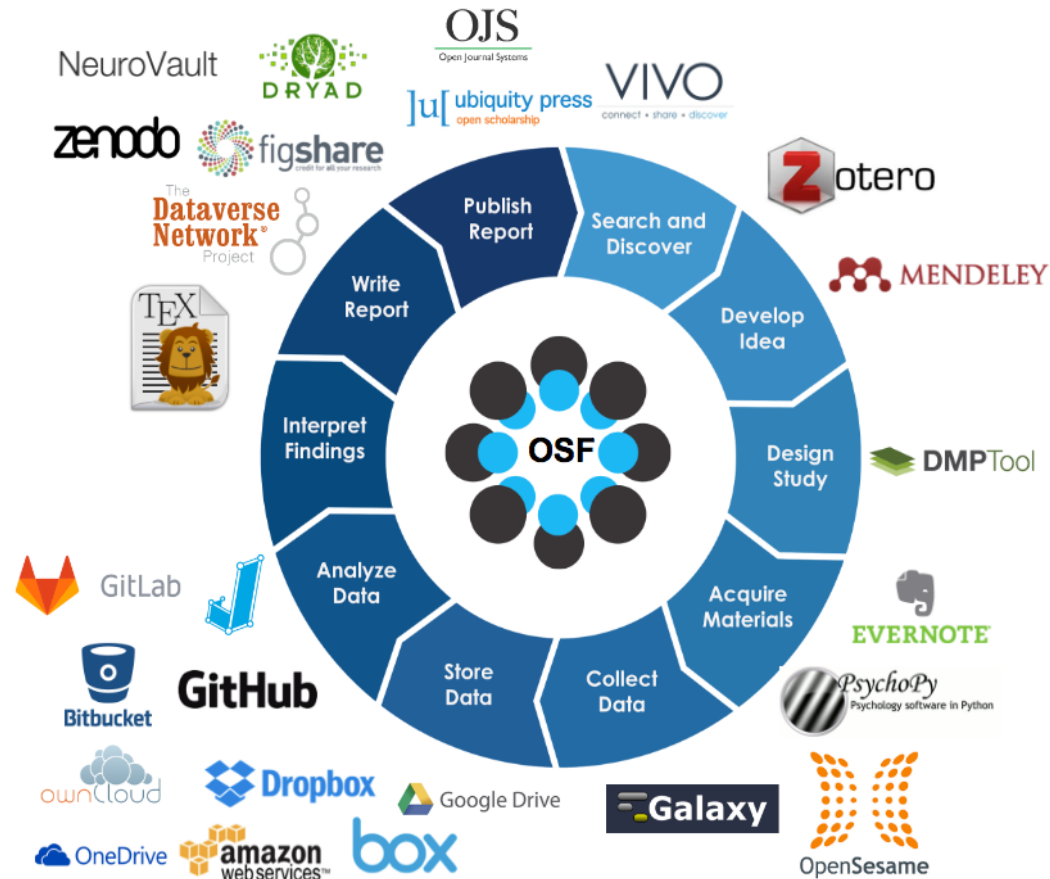


Globalization and Collaboration

□ More and more researchers working on international collaboration projects

■ Need for sharing and storing information

■ Need for online collaboration platform



European Open Science Cloud (EOSC)

- ❑ EOSC aims to accelerate and support the current **transition to more effective Open Science and Open Innovation** in the Digital Single Market.
- ❑ KEY FACTORS:
 - New modes of scholarly communication
 - **Modern reward and recognition practices** need to support data sharing and re-use.
 - **Core data experts** need to be trained and their career perspective significantly improved.
 - A real stimulus of **multi-disciplinary collaboration** requires specific measures in terms of review, funding and infrastructure.
 - The transition **from scientific insights towards innovation needs** a dedicated support policy.



Open access papers 'gain more traffic and citations'

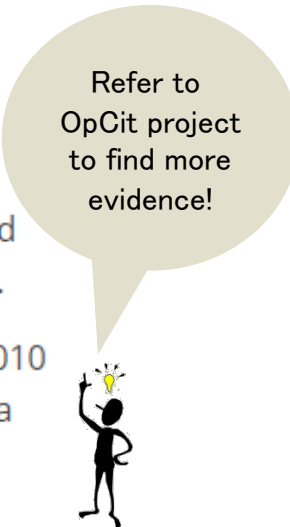
Open access science articles are read and cited more often than articles available only to subscribers, a study has suggested.

July 30, 2014

The Research Information Network analysed the web traffic to more than 700 articles published in hybrid science journal *Nature Communications* in the first six months of 2013.

It found that, after 180 days, articles whose authors had paid for them to be made open access had been viewed more than twice as often as those articles accessible only to the journal's subscribers.

A further analysis of more than 2,000 papers published in *Nature Communications* between April 2010 and June 2013 revealed that open access articles were cited a median of 11 times, compared with a median of seven citations for subscription-only articles. The paper concludes that open access papers enjoy a "small" citation advantage in all disciplines except chemistry.



Alliance of Science Organisations in Germany...Research Data WG

- Principles for the Handling of Research Data (2010.6)
- Research Data at Your Fingertips—A Position Paper (2015.2)
- 'Research Data Vision 2025' –A Discussion Paper (2018.2)
- Research Data Management. A Guide for Researchers (2018.3)



3-2. Drivers and Responses

Call for Research Transparency and
Research Reproducibility

The Retraction Watch Leaderboard

with 21 comments

Who has the most retractions? Here's our unofficial list (see notes on methodology), which we'll update as more information comes to light:

1. [Yoshitaka Fujii](#) (total retractions: 183) Sources: [Final report of investigating committee](#), [our reporting](#)
2. [Joachim Boldt](#) (96) Sources: [Editors in chief statement](#), [additional coverage](#)
3. [Diederik Stapel](#) (58) Source: [Our cataloging](#)
4. [Adrian Maxim](#) (48) Source: [IEEE database](#)
5. [Peter Chen](#) (Chen-Yuan Chen) (43) Source: [SAGE](#), [our cataloging](#)
6. [Hua Zhong](#) (41) Source: [Journal](#)
7. [Shigeaki Kato](#) (39) Source: [Our cataloging](#)
8. [James Hunton](#) (37) Source: [Our cataloging](#)
9. [Hendrik Schön](#) (36) Sources: PubMed and Thomson Scientific
10. [Hyung-In Moon](#) (35) Source: [Our cataloging](#)
11. [Naoki Mori](#) (32) Source: PubMed, [our cataloging](#)
12. [Tao Liu](#): (29) Source: [Journal](#)
13. [Cheng-Wu Chen](#) (28) Source: [our cataloging](#)
14. [Gideon Goldstein](#) (26)
15. [Scott Reuben](#) (25)
16. [Gilson Khang](#) (22) Sources: [WebCitation.org](#), [WebCitation.org](#), [journal](#)
17. [Friedhelm Herrmann](#) (21)
18. [Noel Chia](#) (21)

Subscribe to Blog via Email

Join 13,442 other subscribers

Pages

[Help us: Here's some of what we're working on](#)

[How you can support Retraction Watch](#)

[Meet the Retraction Watch staff](#)

[About Adam Marcus](#)

[About Ivan Oransky](#)

[The Center For Scientific Integrity](#)

[Board of Directors](#)

[The Retraction Watch FAQ, including comments policy](#)

[The Retraction Watch Transparency Index](#)

[The Retraction Watch Leaderboard](#)

The Royal Society: Science as an open enterprise (2012)

AREAS FOR ACTION

- ❑ **Scientists need to be more open** among themselves and with the public and media
- ❑ **Greater recognition** needs to be given to the value of data gathering, analysis and communication
- ❑ **Common standards** for sharing information are required to make it widely usable
- ❑ **Publishing data in a reusable form** to support findings must be mandatory
- ❑ **More experts in managing and supporting** the use of digital data are required
- ❑ **New software tools** need to be developed to analyse the growing amount of data being gathered





CODATA

International Council for Science : Committee on Data for Science and

ABOUT ▾ EVENTS ▾ MEMBERSHIP ▾ INITIATIVES ▾ TASK GROUPS ▾ WORKING GROUPS

Message from President Geoffrey Boulton

- Recent decades have seen an unprecedented explosion in the human capacity to acquire, store and manipulate data and information and to instantaneously communicate them globally, irrespective of location...
- ...Effective exploitation of Big Data depends fundamentally upon an international culture of 'Open Data' that involves sharing of data and their availability for re-use and re-purposing.



ClimateGate Scandal



FOX NEWS
Fair & Balanced
.com

What's Hot [On the Job Hur](#)
Watch Live [STRATEGY RO](#)

HOME U.S. WORLD BUSINESS POLITI

Do E-Mails Reveal Scientist Claims On Climate Change are... **BUNK?**

NASA

Hackers break into servers of a major British climate change research facility and purportedly uncover e-mails urging scientists to 'hide the decline' of temperatures, manipulate data and silence skeptics.



This is the worst scientific scandal of our generation

Telegraph



Science International: Open Data in a Big Data World

- The accord identifies the opportunities and challenges of the **data revolution** as today' s predominant issue for global science policy. It proposes fundamental principles that should be adopted in responding to them. It adds the distinctive voice of the scientific community to those of governments and inter-governmental bodies that have made the case for **open data as a fundamental pre-requisite** in **maintaining the rigour of scientific inquiry** and **maximising public benefit** from the data revolution in both developed and developing countries.



Reproducibility Project: Psychology

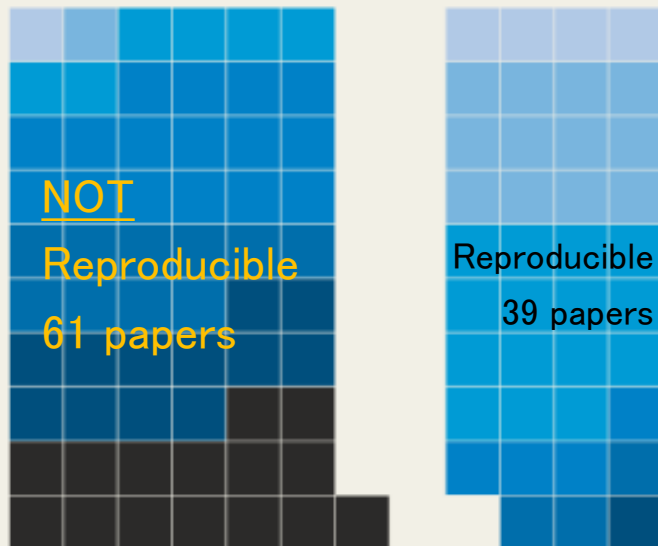
RELIABILITY TEST

An effort to reproduce 100 psychology findings found that only 39 held up.* But some of the 61 non-replications reported similar findings to those of their original papers.

Did replicate match original's results?

NO: 61

YES: 39



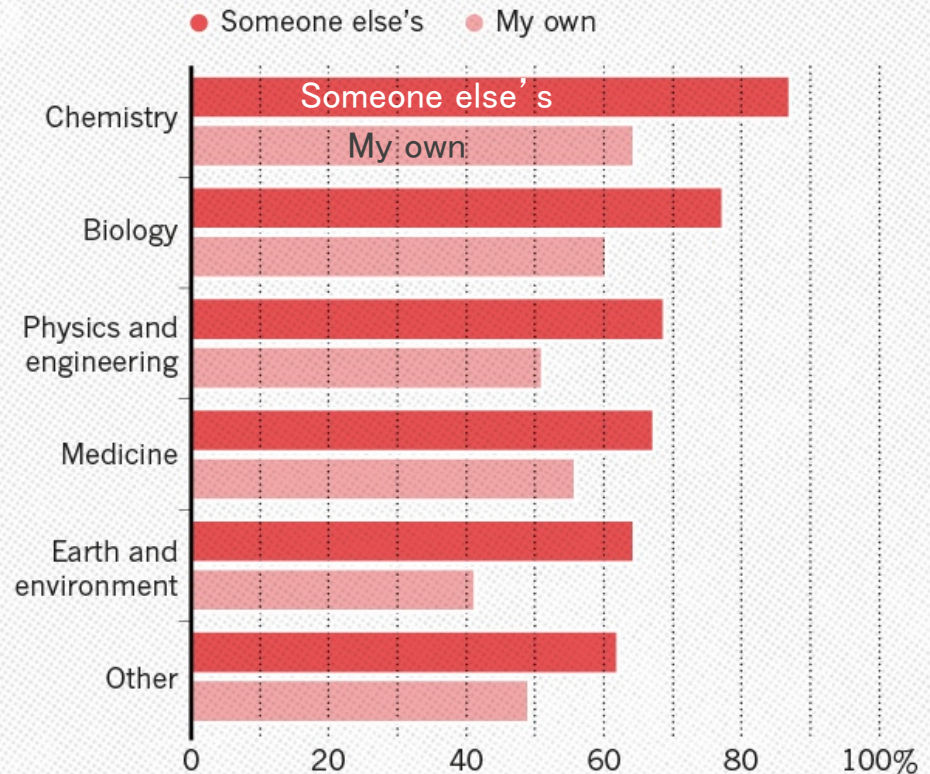
Replicator's opinion: How closely did findings resemble the original study:

- Virtually identical
- Extremely similar
- Very similar
- Moderately similar
- Somewhat similar
- Slightly similar
- Not at all similar

* based on criteria set at the start of each study

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



Source: Nature, "Over half of psychology studies fail reproducibility test" (2015.9.27)

<https://www.nature.com/news/over-half-of-psychology-studies-fail-reproducibility-test-1.18248>

In reality, the major driving force for RDM is scientific misconduct prevention

- MEXT: “Guideline for Dealing with Scientific Misconduct” (2014)
 - 「研究活動における不正行為への対応等に関するガイドライン」
 - Strengthening the guideline in 2006.
 - Holds institutions to be responsible for research transparency and preventing scientific misconduct.

- Science Council of Japan: “Reply: For the Enhancement of Soundness of Scientific Research” (2015)
 - 日本学術会議「(回答)科学研究における健全性の向上について」
 - “Ten-Years Preservation Rule for Research Data”
 - 研究データ10年保存ルール

3-3. Drivers and Responses

Social Demand

Citizen Science...examples

□ Galaxy Zoo

- Crowdsourced astronomy project where people classify galaxies

□ Foldit

- Online puzzle video game about protein folding

□ eBird

- Online database of bird observations

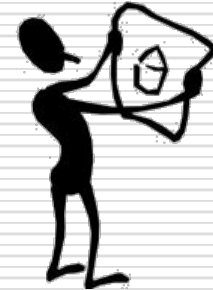
From Access to Research Publications to Access to Research Data

Publicly-
funded
Research

Research
Data



Research
Publications



The push which triggered US government to adopt OA policy

TED Ideas worth spreading WATCH

Share
Add to list
Like
Rate

Sharon Terry | TEDMED 2016

Science didn't understand my kids' rare disease until I decided to study it

15:03

▶

27

The image shows a video player interface for a TED talk. At the top, the TED logo and the tagline 'Ideas worth spreading' are on the left, and the word 'WATCH' is on the right. The main video area shows a woman with short grey hair, Sharon Terry, wearing a red sleeveless dress, standing on a stage. The background features several framed panels with colorful, abstract patterns. A large white play button is centered over the video. On the right side of the video, there is a vertical menu with icons for 'Share', 'Add to list', 'Like', and 'Rate'. At the bottom of the video, there is a red banner with the text 'Sharon Terry | TEDMED 2016' and the title 'Science didn't understand my kids' rare disease until I decided to study it'. Below the banner is a progress bar and a timestamp of '15:03'. At the bottom right of the video player, there are icons for play/pause, volume, full screen, settings, and a share icon. The number '27' is visible in the bottom right corner of the slide.

Data Sharing Policy, NIH (2003-)

In NIH's view, **all data** should be considered for data sharing. Data should be **made as widely and freely available as possible** while safeguarding the privacy of participants, and protecting confidential and proprietary data. To facilitate data sharing, investigators submitting a research application requesting \$500,000 or more of direct costs in any single year to NIH on or after October 1, 2003 are expected to **include a plan for sharing final research data** for research purposes, or state why data sharing is not possible.



Data Management Plan (DMP) required!

Policy Developments in Research Data Sharing

- 2003, NIH, Data Sharing Policy
- 2004, OECD Declaration on Access to Research Data from Public Funding
- 2007, OECD Principles and Guidelines for Access to Research Data from Public Funding
- 2007, Biotechnology and Biological Sciences Research Council (BBSRC-UK), Data Sharing Policy
- 2011, Research Councils UK, Data Sharing Policy
- 2011, NSF, Data Sharing Policy
- 2013, OSTP-US, Increasing Access to the Results of Federally Funded Scientific Research
- 2014-20, Horizon 2020, Open Research Data Pilot

Rationale for making Research Data publicly available

□ Accountability

- Publicly funded research should be transparent

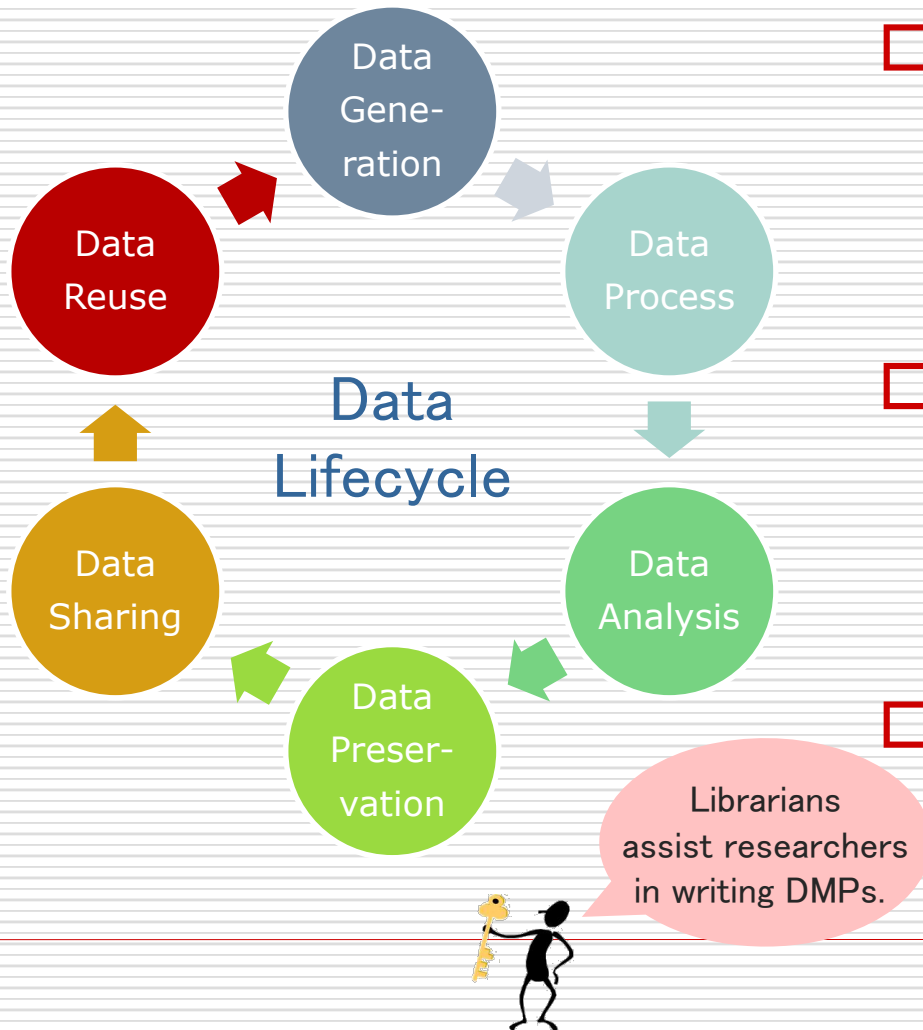
□ Economic Efficiency

- Reuse of data leads to new findings without additional investments

□ Global Challenges Solving and Innovations

- Combining data from multiple discipline leads to solving global challenges
- Industries using data leads to innovations

Data Management Plan (DMP)



- DMP – A plan where researcher state the treatment of data used during research process.
- More and more, research funders are requiring to submit DMP along with research proposal.
- In Japan, JST is asking for DMP since FY2018.

3. Institutional Role in Supporting Open Science

Open Science at Academic Institution Level



1. Hold accountability

- ✓ Meeting mandates by funding agencies
- ✓ Research transparency and reproducibility

2. Promoting research

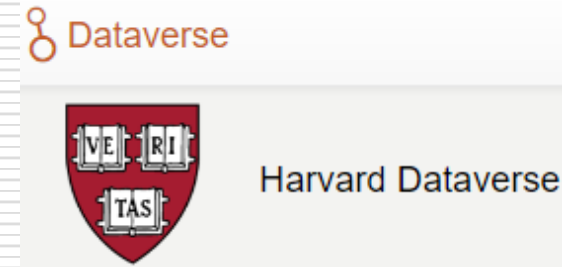
- ✓ Providing proper research environment
- ✓ Accelerating research

3. Disseminating research of the institution

4. Linking academia and society

Data Repositories

□ General



□ Disciplinary Data Repositories

➤ Numerous

How we classify our tools and services

Data Management Support

Data Management Planning

Active Data Infrastructure

Data Stewardship

Before research

During research

After research



Providing Training for Research Data Management

MANTRA
Research Data Management Training

MANTRA is a free online course for those who manage digital data as part of their research project.

Learning Units: Select one to start

- Research data explained >
- File formats & transformation >
- Data management plans >
- Documentation, metadata, citation >
- Organising data >
- Storage & security >

Research Data Management and Sharing

About this course: This course will provide learners with an introduction to research data management and sharing. After completing this course, learners will understand the diversity of data and their management needs across the research data lifecycle, be able to identify the components of good data management plans, and be familiar with best practices for working with data including the organization, > More

ga088: オープンサイエンス時代の研究データ管理

ga088を受講登録する (無料)

EDINA The University of Edinburgh
GACCO Certified by JMACC
The University of North Carolina at Chapel Hill, The University of Edinburgh

<http://datalib.edina.ac.uk/mantra/>

<https://www.coursera.org/learn/data-management>

https://lms.gacco.org/courses/course-v1:gacco+ga088+2017_11/about

Tangible actions in Open Science at Academic Institutions



President

1. Hold accountability for research

- ✓ Research office caring for mandates and transparency
- ✓ Providing infrastructure for data storage

2. Disseminating research

- ✓ Make research discoverable and reusable
- ✓ Provide institutional repositories

University library related activities highlighted!

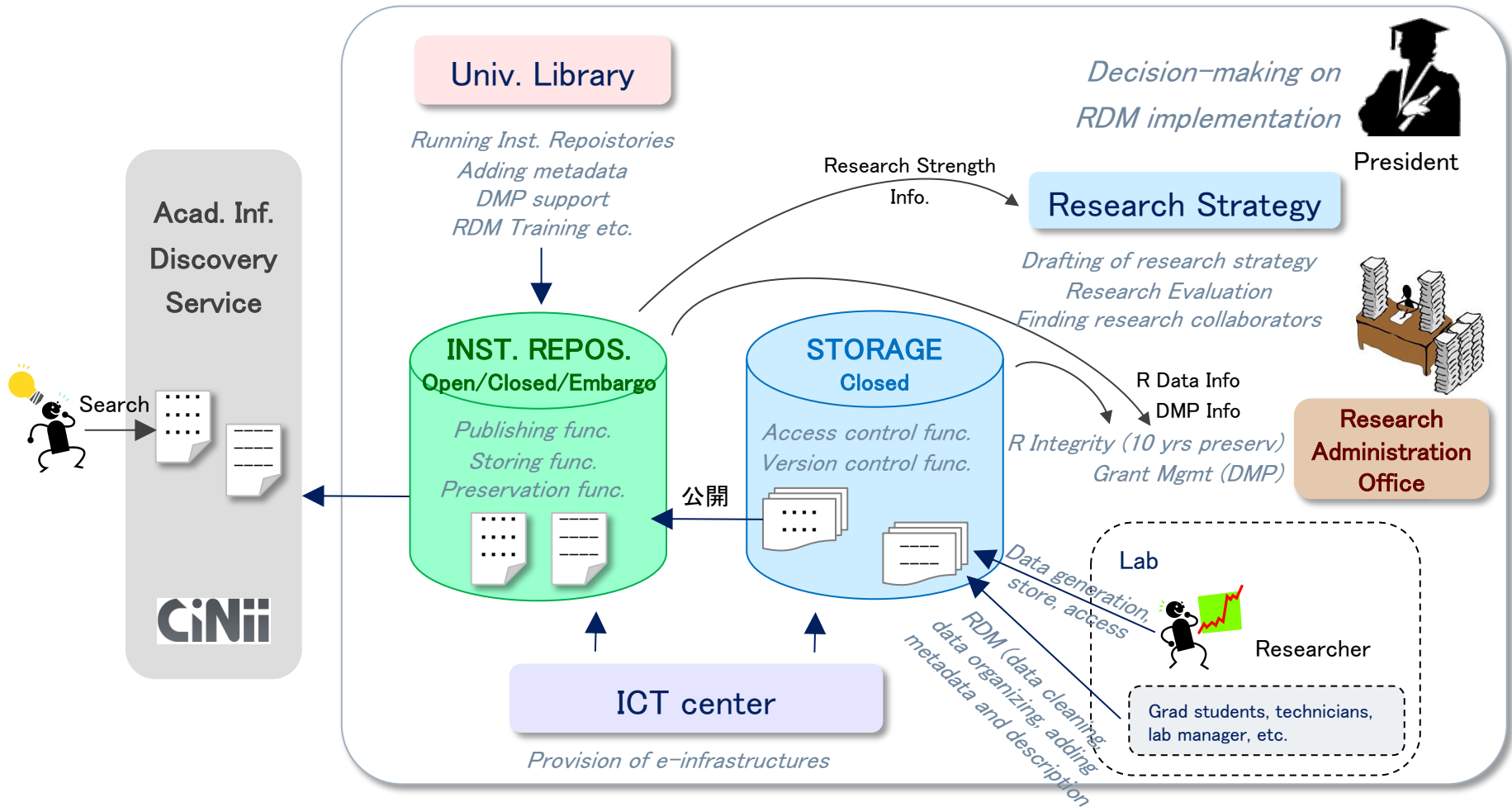
3. Promoting research

- ✓ Provide access to scholarly contents, promoting OA
- ✓ Promoting new research paradigm, data-intensive science
- ✓ Advocating researchers for new research paradigm
- ✓ Evaluate research within institution

4. Linking academia and society

- ✓ Matching the needs of society with the seeds of researchers

RDM implementation in an academic institute

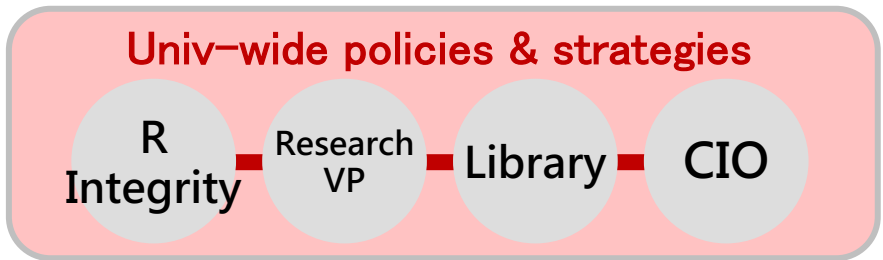


Multi-stakeholder Approach needed to implement RDM at universities

Multi Stakeholder Approach

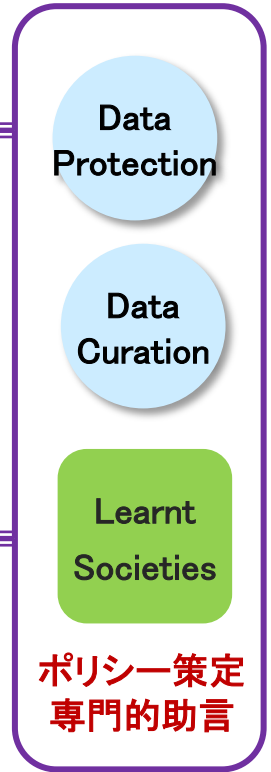
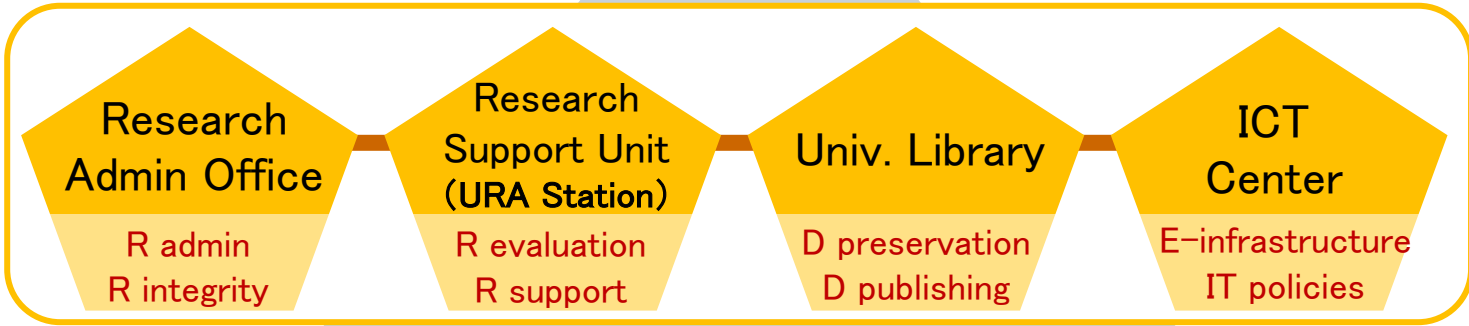
I want to make the university research competitive!

Administration

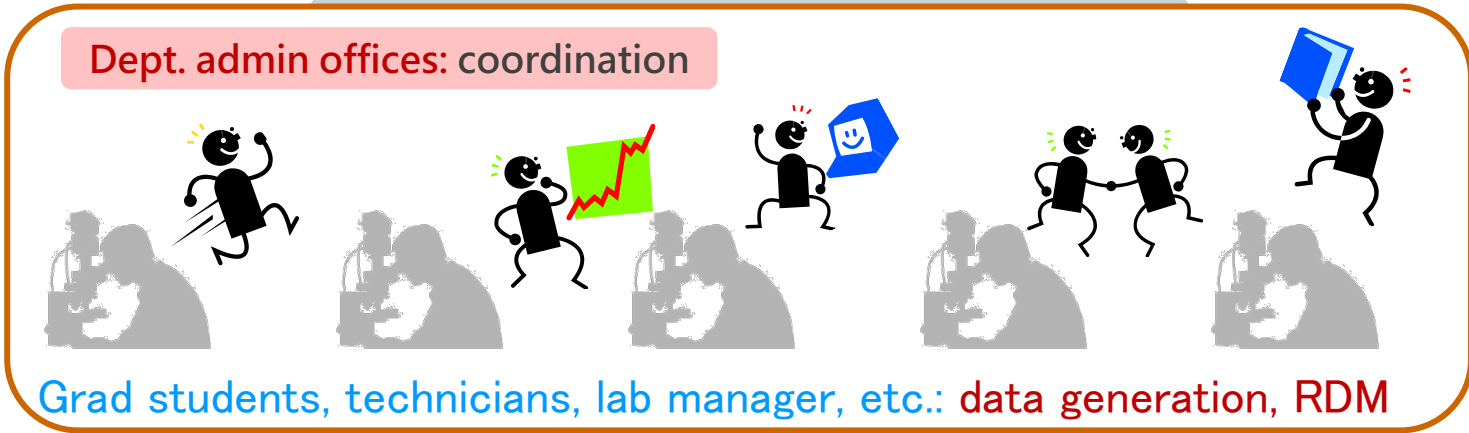


Professional Assoc.

Univ-wide Service Units



Departments, Labs



ポリシー策定
専門的助言

Why an RDM Charter?

- ❑ Participants at **AXIES-RDM** session started to claim,
 - *“We need a charter in order to convince the university administration and to get the researchers and staff engaged!”*
- ✓ **AXIES**
 - Academic eXchange for Information Environment and Strategy
 - Community of CIOs and ICT centers of universities in Japan.
 - Counterpart to EDUCAUSE in the US

“RDM Charter for Academic Institutions”

□ RDM Charter

- Not for researchers, but
- **For academic institutions!**



Researcher

Don't dare to
tell me how to
manage my data!
I know what
I'm doing!

□ Purpose of RDM Charter

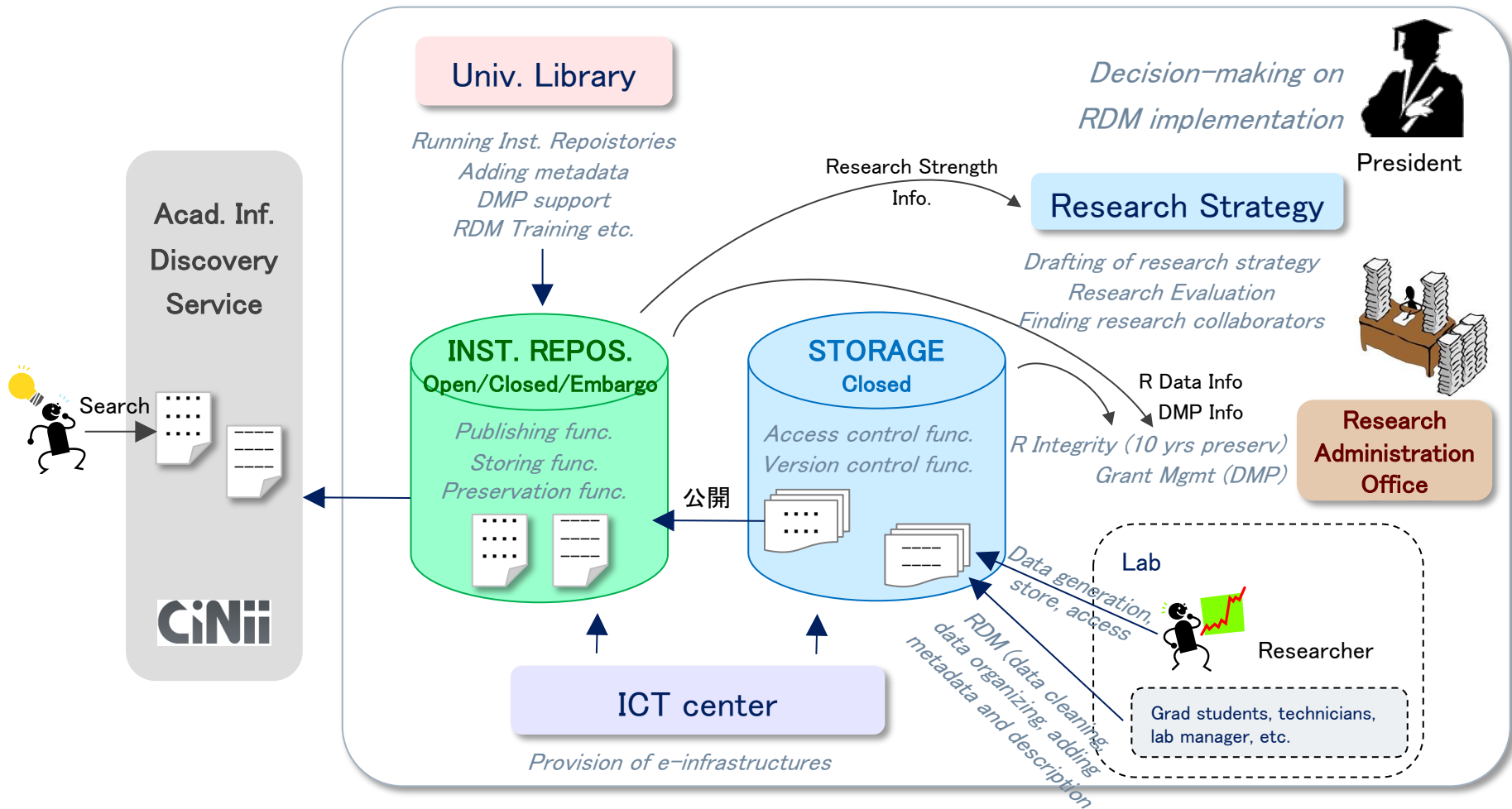
- Give university administration ideas and options to implement RDM in respective institutions.

“RDM Charter for Academic Institutions”

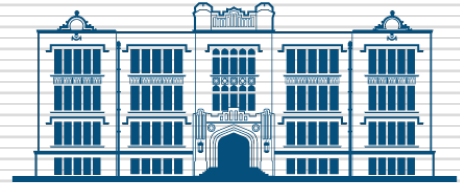
...Composition

- The Charter (3 pages)
 - Addresses the viewpoints why academic institutions needs to take RDM seriously.
 - Viewpoints in bullet points:
 - I. Role of academic institutions in RDM
 - II. Policies and organizations for RDM needed in acad. Institutions
 - III. RDM procedures in acad. Institutions
 - IV. RDM Purpose options in acad. Institutions
 - V. Digital platform functions needed for RDM in acad. Institutions
 - VI. Human resources development for RDM in acad. Institutions
 - VII. Reuse and service options of research data in acad. institutions
- Appendix (12 pages)
- Glossary (12 terms, 3 pages)
- References (2 pages)

RDM implementation in an academic institute



Role of University Libraries in promoting Open Science



1. Providing scholarly contents within institution

- ✓ Acquiring and locating scholarly contents (books, e-journals, other)
- ✓ Promoting open access

2. Stewarding scholarship within institution

- ✓ Provide storage for active data and long-term preservation
- ✓ Provide DMP tool

3. Disseminating scholarship of institution

- ✓ Provide institutional repository for publishing
- ✓ Adding proper metadata and curating data

4. Advocating for good scholarship practices, i.e. Open Science

- ✓ Working on Knowledge Graph (linking publication, data, researcher, grant, etc)
- ✓ Provide RDM Training

University Libraries in Open Science Era



- ❑ Expanding the scope of scholarly contents
 - ✓ Books, journals, gray-literature, *research data*
- ❑ Develop new services for new contents
 - ✓ DMP tools, RDM training
 - ✓ Data curation, facilitating reuse of scholarship
 - ✓ Building knowledge graph
- ❑ Advocating of good scholarship practices in the Open Science era!
 - ✓ Promoting open access and RDM
 - ✓ Disseminating and promoting reuse of scholarship
 - ✓ Caring for research transparency and reproducibility
 - ✓ Advocating for new research paradigm

*As the steward of scholarship of the institution,
be the change agent for new research paradigm!*

FORCE11

The Future of Research Communications
and e-Scholarship

FORCE11 is a community of researchers, students, librarians, publishers, funding officers, and support service providers that individually and collectively work together to facilitate a change in the way research is communicated. Together this community has made significant strides towards changing the way research communication is conducted, playing a role in a number of recent major initiatives including:

- The FAIR Principles
- The Data Citation Principle
- The Scholarly Commons
- RRDs (Research Resource Identifiers)
- Software Citation Principles
- Data Citation Principles Implementation Project
- Annotating All Knowledge Coalition

Join the FORCE11 community at www.force11.org

The FORCE11 community is active all year round, but comes together for its two flagship events:

FSCI

FORCE 2019

The FORCE11 Scholarly Communication Institute (FSCI) is a 5-day program of intensive courses, group activities and hands-on training focused on the latest trends in scholarly research communication. This 'summer school' is intended for researchers, administrators, librarians, students and others to navigate the new ways required for open research communication.

In 2019, FSCI will be held in collaboration with the UCLA Library, Los Angeles, CA.

August 5 - 9. Registrations open March 4.

www.force11.org/fsci/2019
#FSCI

The annual FORCE11 conference heads to Edinburgh, Scotland for its 2019 incarnation.

By bringing together stakeholders from across the information community for an open discussion, the conference aims to unlock exciting new ideas and potential collaborations.

Details of the conference schedule will be announced later in 2019.

Preconference workshops October 15.
Main conference October 16 and 17.

www.force11.org/meetings/force2019

Register at FORCE11!

The Future of Research Communications
and e-Scholarship

- Registration is **free!**
- **Newsletter** will be sent.
- **Travel fellowship** will be available for participants from developing countries for below two events!

FORCE11 events:

□ FSCI

- August 5-9 @ UCLA
- FORCE11 Scholarly Communication Institute—a **5-day intensive program!**

□ FORCE2019

- October 15-17 @ Edinburgh
- Annual FORCE11 conference