International framework of scientific data sharing and open science

Yasuhiro Murayama

Members of Cabinet Office Expert Panel of Open Science Promotion

Research Executive Director,
Big Data Integration Research Center,
National Institute of Information & Communications Technology

My background

Geophysics, Atmospheric/Space Science

Board member of Japan Geoscience Union (2014-2015)

Visiting Professor, Kyoto University, (2013-2014)

Associate member, Science Council of Japan

Research Executive Director, Big Data Integration Res. Ctr.
Natl. Inst. Information & Communications Technology

ICSU-World Data System Scientific Committee, ex officio membe

Member of Cabinet Office Expert Panel of Open Science

Observer Member of EC's High Level Expert Group of European Open Science Cloud

Data Management, Science Policy

Today's Contents

- Introduction
- Data Sharing/Open Research Data
  - Method of Modern Science and Communicatoins
- International Policy Situation
  - EU, OECD, Japan
- Toward Our Best Practice
  - Data publication, data citation
  - Digital socio-technological data/information infrastructure
- Concluding Remarks

Introduction

G8 2013 Science Ministers' Agreement of Open Research Data

G8 Science Ministers Statement London UK, 12

Introduction

We, the G8 Science Ministers met in London on Wednesd
of our respective national science academies, as part of th
this unique meeting we discussed how our nations could le
transparency, coherence and coordination of the global sc
in order to address global challenges and maximise the so
of research.

G8 Open Data Charter will 'increase transparency' and 'fuel innovation'

Five key principles outlines how governments
onomic and social

3. Open Scientific Research Data

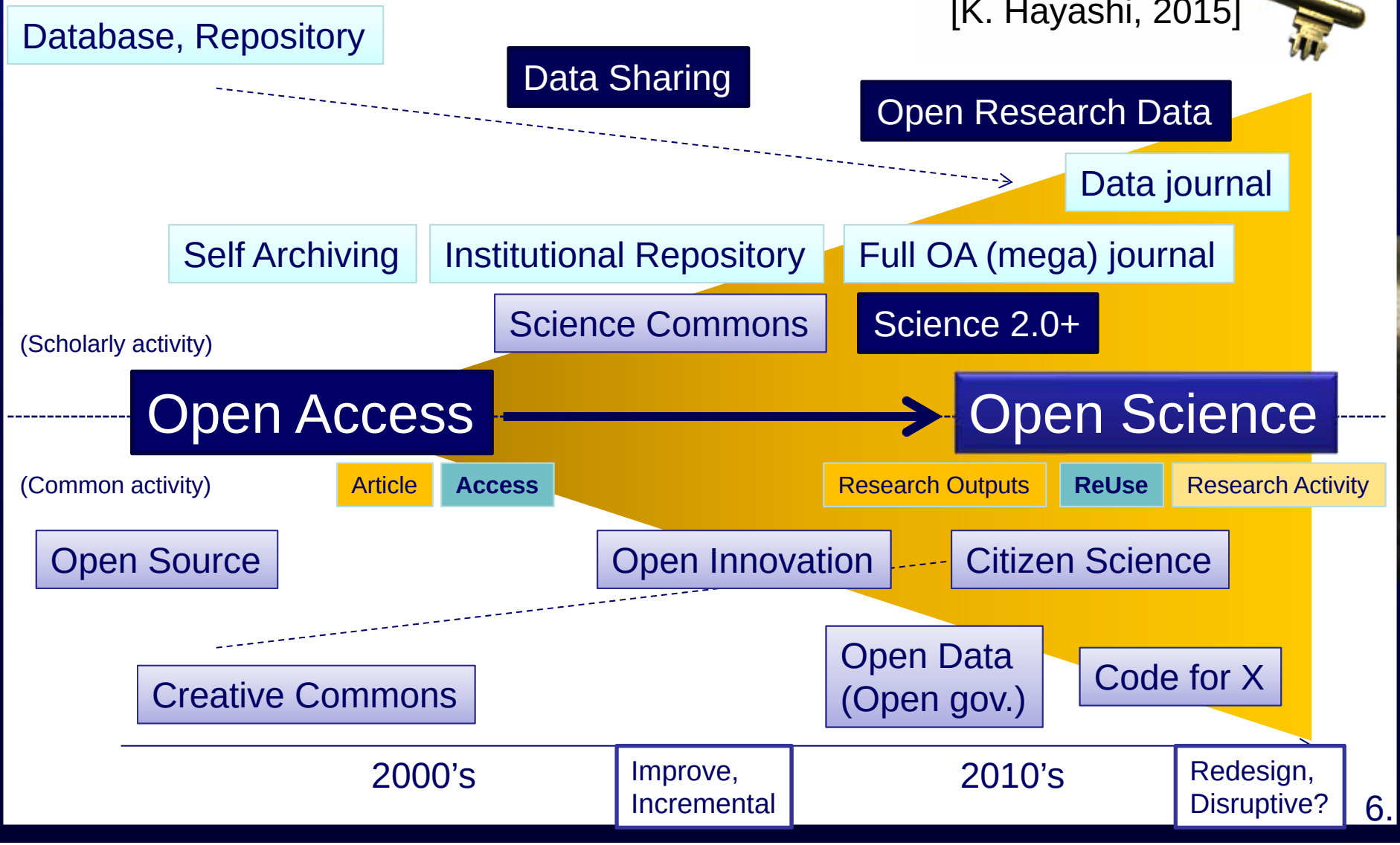
Open enquiry is at the heart of scientific endeavour, and rapid technological change has profound implications for the way that science is both conducted and its results communicated. It can provide society with the necessary information to solve global challenges. We are committed to openness in scientific research data to speed up the progress of scientific discovery, create innovation, ensure that the results of

4. Expanding Access to Scientific Research Results

"Open Government Data"

Open Access to Open Data and Open Science
Overview example

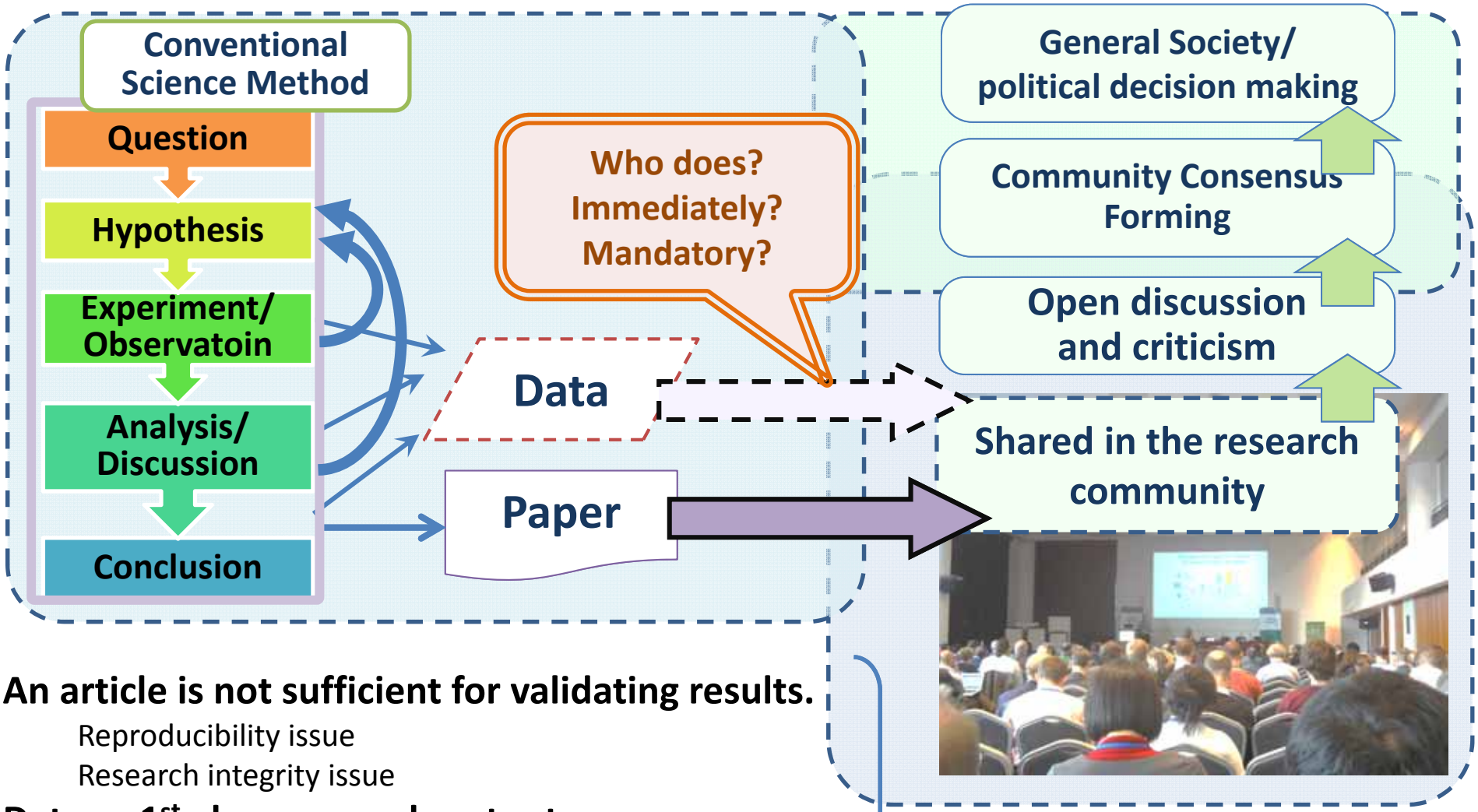
[K. Hayashi, 2015]



Scientific Practice and Data Sharing/Open Research Data:

Changing Scholarly Communications

Society and Science: scholarly papers and data



An article is not sufficient for validating results.

- Reproducibility issue
- Research integrity issue

Data as 1st-class research output

- Social information asset, provided to the general society

Essential for "irreproducible" natural phenomena

- Global change, space, living organisms, health . . .

<u>Mutual trust between Science and Society</u>

Climate Change Knowledge supported by Thousand Scientists

IPCC (Intergovernmental Panel on Climate Change)

WG1 “Physical Science Basis”

19 Headlines
on less than 2 Pages

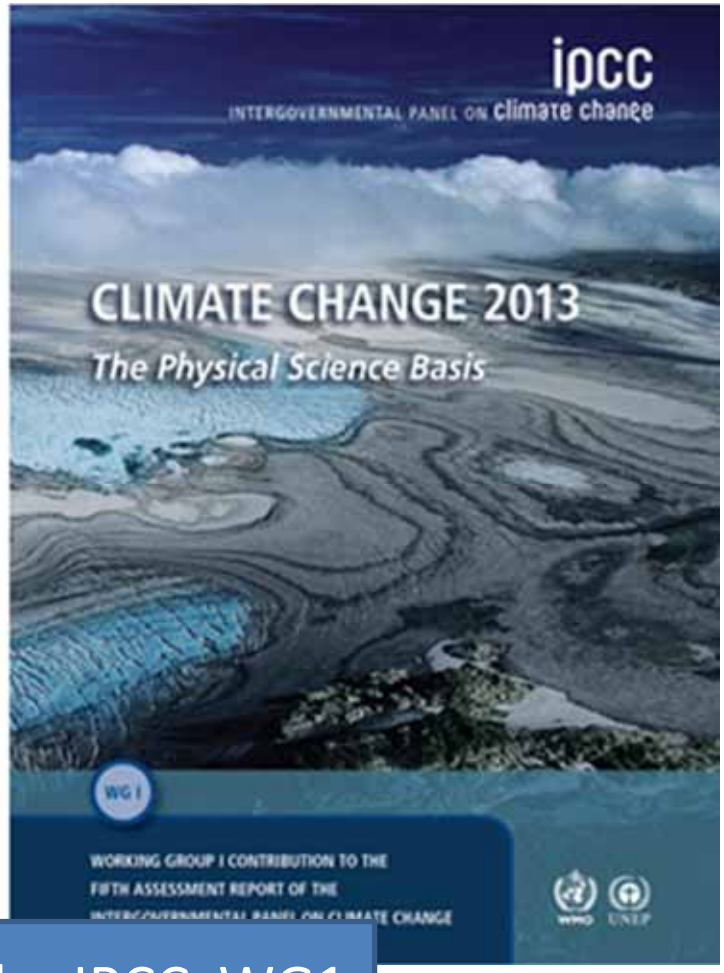
Summary for Policymakers
ca. 14,000 Words

14 Chapters
Atlas of Regional Projections

54,677 Review Comments
by 1089 Experts

2010: 259 Authors Selected

2009: WGI Outline Approved



Approx. 1,300 scientists worked for the IPCC WG1.
(3,000-4,000 scientists for all WG1-3?)

[IPCC, 2013]

[Geoffrey Boulton, CODATA conf., 2012]

A crisis of replicability?

NATURE | VOL 483 | 29 MARCH 2012

REPRODUCIBILITY OF RESEARCH FINDINGS

Preclinical research generates many secondary publications, even when results cannot be reproduced.

Journal impact factor	Number of articles	Mean number of citations of non-reproduced articles*	Mean number of citations of reproduced articles
>20	21	248 (range 3–800)	231 (range 82–519)
5–19	32	169 (range 6–1,909)	13 (range 3–24)

Results from ten-year retrospective analysis of experiments performed prospectively. The term 'non-reproduced' was assigned on the basis of findings not being sufficiently robust to drive a dru velopment programme.
*Source of citations: Google Scholar, May 2011.

10

Replicability of 53 papers on "good" IF journals are examined ;
A paper is cited by a number of secondary works , regardless of its replicability.

[Carthage Smith (OECD Global Science Forum), 2015]

Open science: the (unrealized) potential

- 'Big data' and ICTs open up new scientific opportunities
- Enable collaboration across disciplines
- Increase efficiency, transparency and reproducibility
- Address global challenges more effectively
- Increase knowledge spill-overs for science, innovation and society
- Promote citizen engagement in science

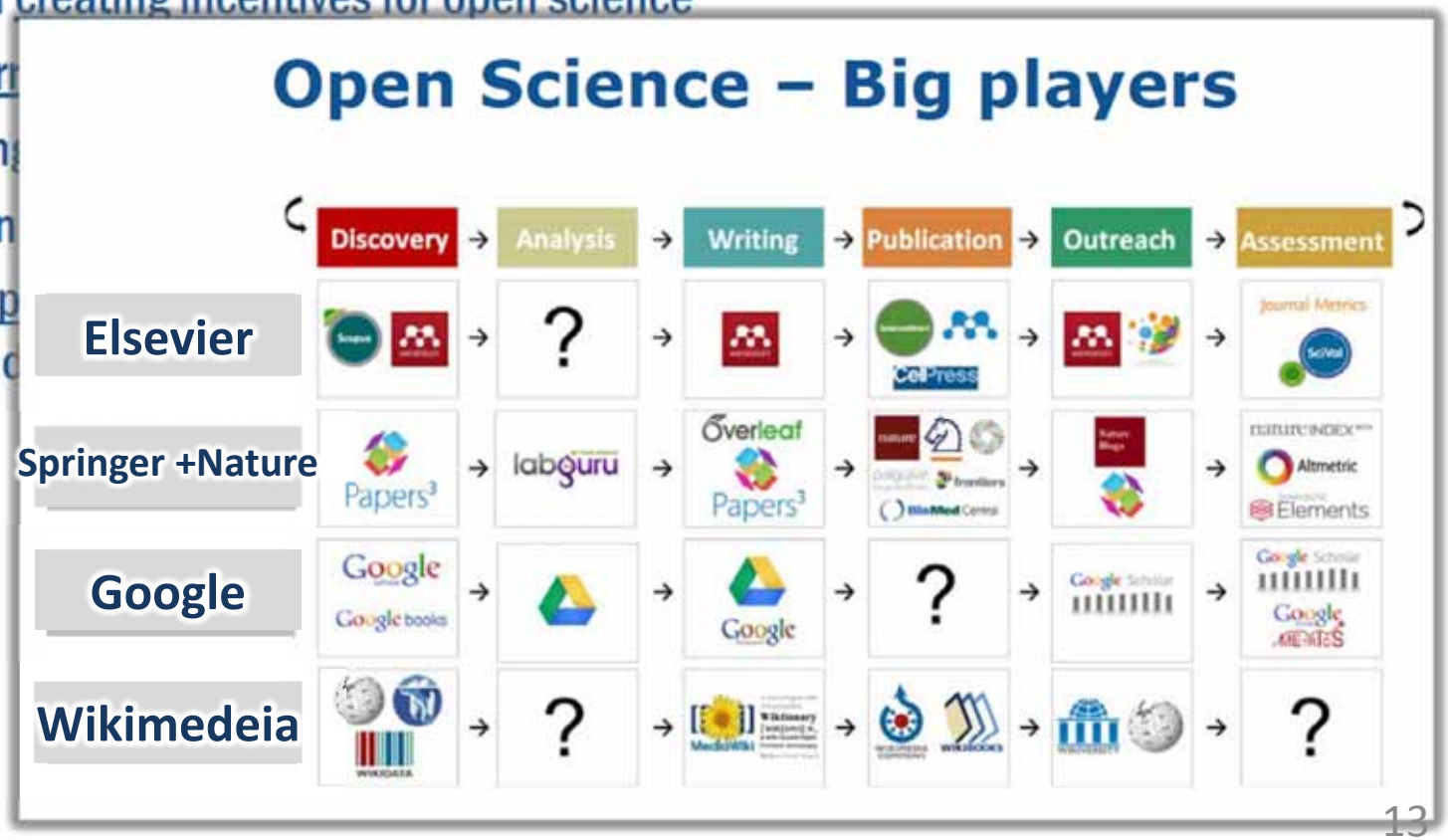
International Policy Situation

[J.C. Burgelman (EC DG-RTD),
Data Sharing Symposium, Tokyo 2015]

European Open Science Agenda

5 broad policy action lines (from public consultation, validated by stakeholders incl. EU Member States):

- ✓ Fostering and creating incentives for open science
- ✓ Removing barr
- ✓ Mainstreamin
- ✓ Developing an
- ✓ Embedding op to societal and



Organization for Economic Co-operation & Development

Defining Open science

[Carthage Smith (OECD Global Science Forum), 2015]

Open science includes:

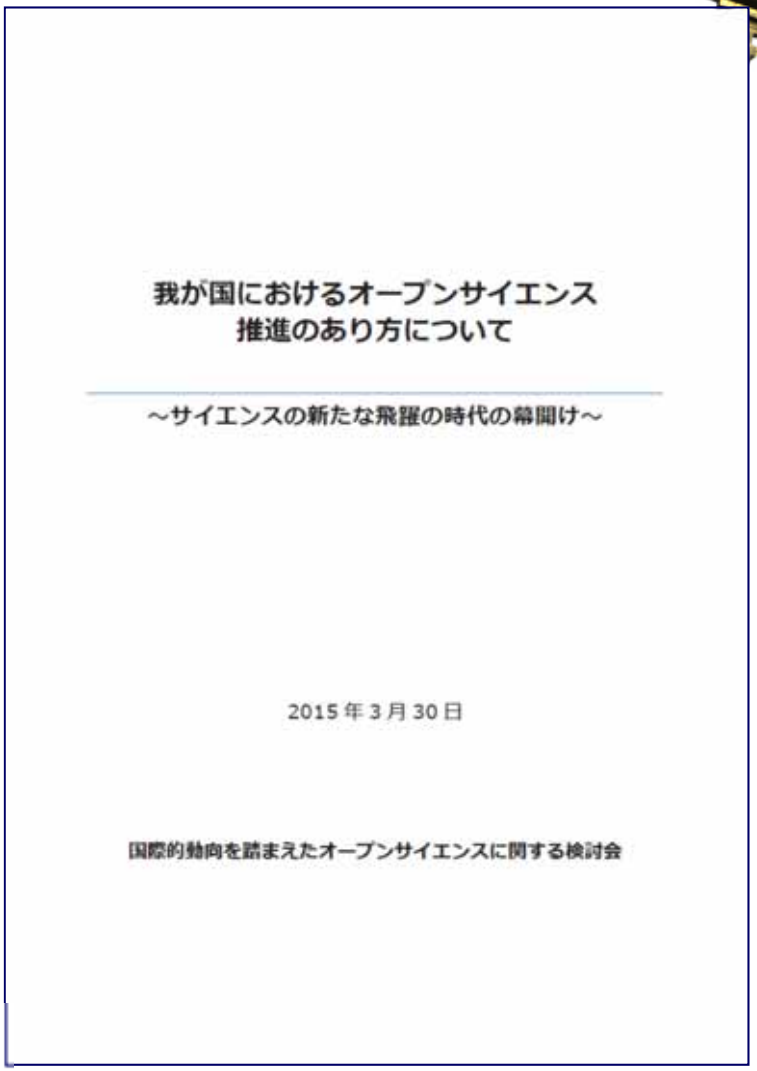
- Open access to scientific publications
- Open and "intelligent" access to research data (and materials)
- Open access to digital applications and source code
- Open access for scientists, the public and commercial companies

➢ re-asserting science as a global public good

"A new scientific paradigm"

- Science is becoming increasingly data-driven

Expert Panel on Open Science based on Global Perspectives (Cabinet Office, Japan)



Promoting Open Science in Japan Opening up a new era for the advancement of science
Report by the Expert Panel on Open Science, based on Global Perspectives Cabinet Office, Government of Japan March 30, 2015

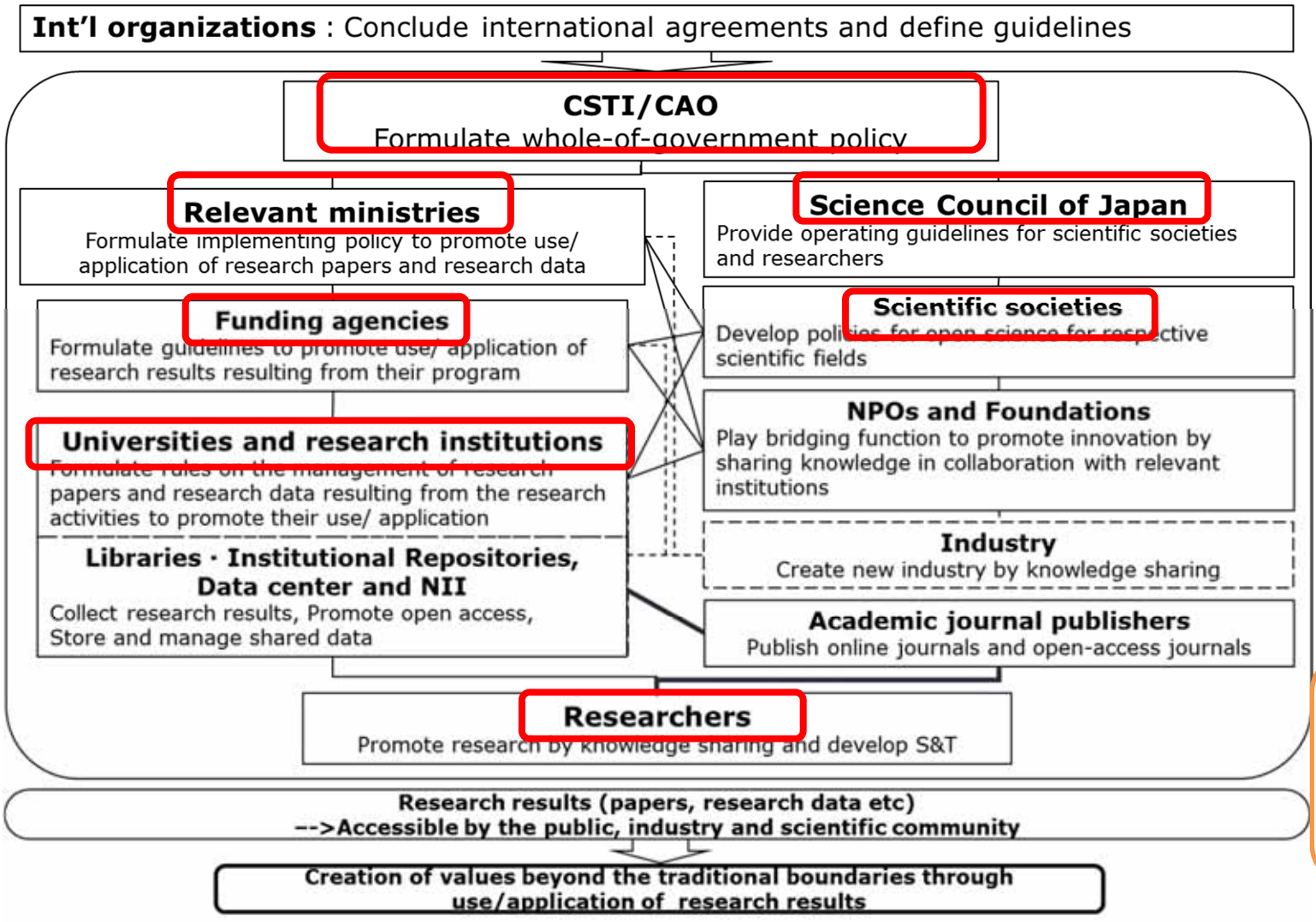
[H. Manago, 2015]

Cabinet Office/CSTI：National Principle of Open Science

Cabinet Office "Expert Panel of Open Science" **(Dec, '14 ---March '15)** [H. Manago, 2015]

http://www8.cao.go.jp/cstp/sonota/openscience/

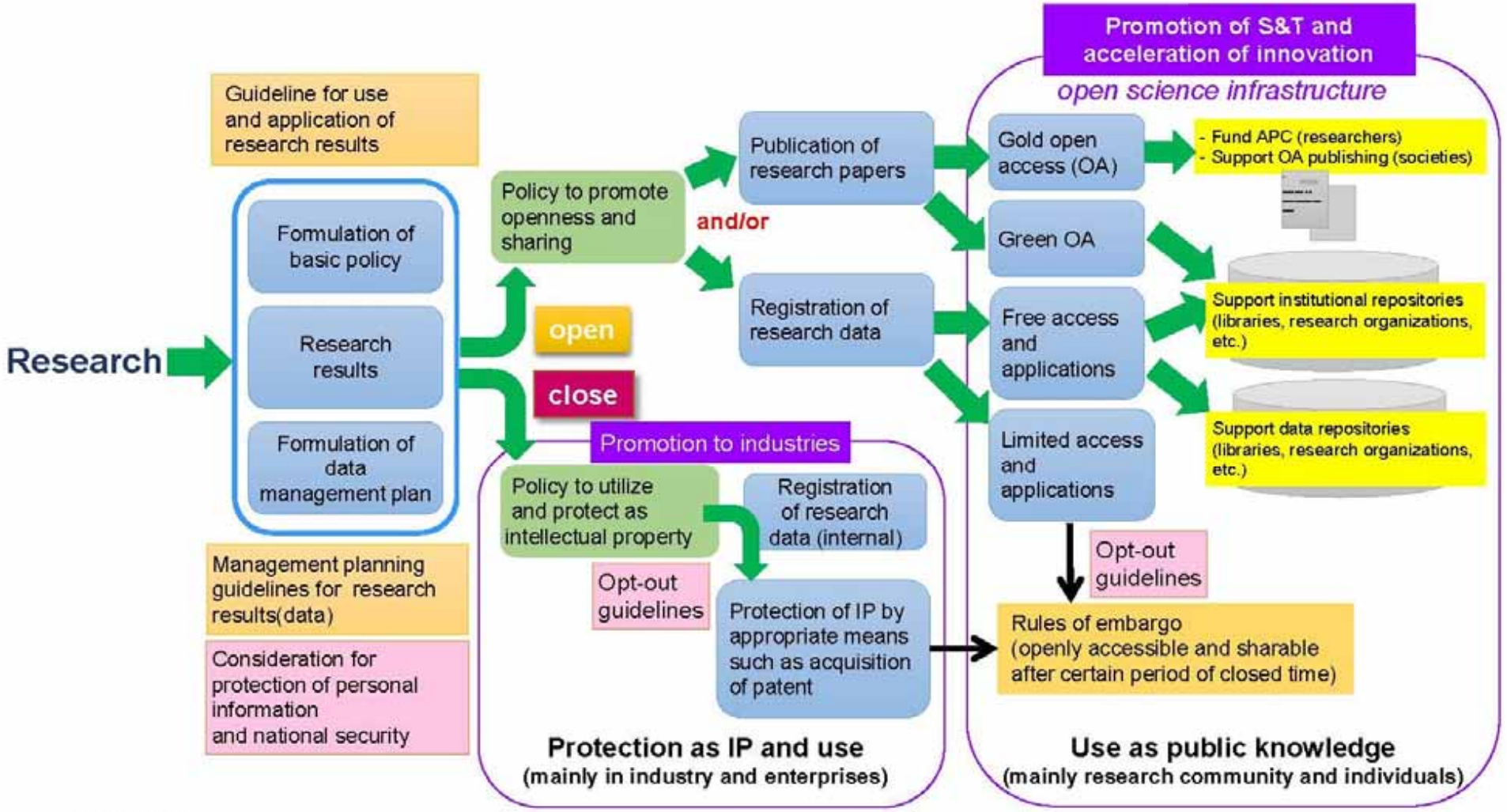
=➔ Final Report was published at the Web site 30 March 2015.



Input to the 5th National Basic S&T Plan

Policy map for Promotion of Open Science

[H. Manago, 2015]



Reference:
Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 Version 1.0 11 December 2013 p.4
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

How to promote “Open Science” in Japan

- The “national principle” is not obligation or mandatory rules, but “guiding principle”.
- The decision by Cabinet Office is now being followed by stakeholders’ discussions (related ministries, scientific societies, universities/ national institutes)
- Every scholars do not accept. Depends on their disciplines and past practice/culture.
- New funding mechanism is also required
  - to encourage researchers, journal editors, publishers, data producers, data infrastructure managers/developers.

G7 2016 Science & Technology Ministers' Meeting
(15-17 May 2016, Tsukuba, Ibaragi, Japan)

MINISTERS' MEETING AGENDA:

1. **Global Health - Health Care and Science and Technology**
2. **Gender and Human Resource Development for STI**
3. **The Future of the Seas and Oceans**
4. **Clean Energy - Developing Innovative Energy Technology**
5. **Inclusive Innovation - Mainstreaming Inclusiveness Among Innovation Policies**
6. **Open Science - Entering into a New Era for Science**

→ **Agreed to establish a new G7 Open Science Working Group**

Photos provided by Cabinet Office of Japan, Ibaragi Prefectural Government, and Tsukuba City Government

Toward Our Best Practice

Print & Electronic Technologies as Social Info. Infrastructures

--- 百年の印刷文化の基礎支えと、成長途中のディジタル･サイエンス

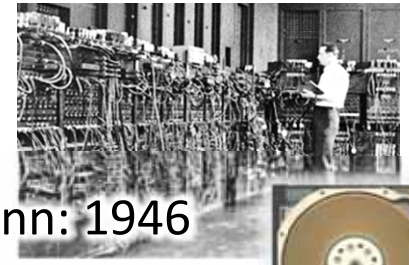
Public library (paper media) : 8c

Printing press/Gutenberg: 1445

First scientific journal: 1665

Intl. Assoc. Academies: 1899

ICSU established: 1931



ENIAC, von Neumann: 1946

Hard Disk Drive: 1956

World Data Center system : 1957

TCP/IP, dial-up (64kbps): 1982

WWW (CERN): 1991

Broadband internet (>1Mbps) : ～2000

351 years

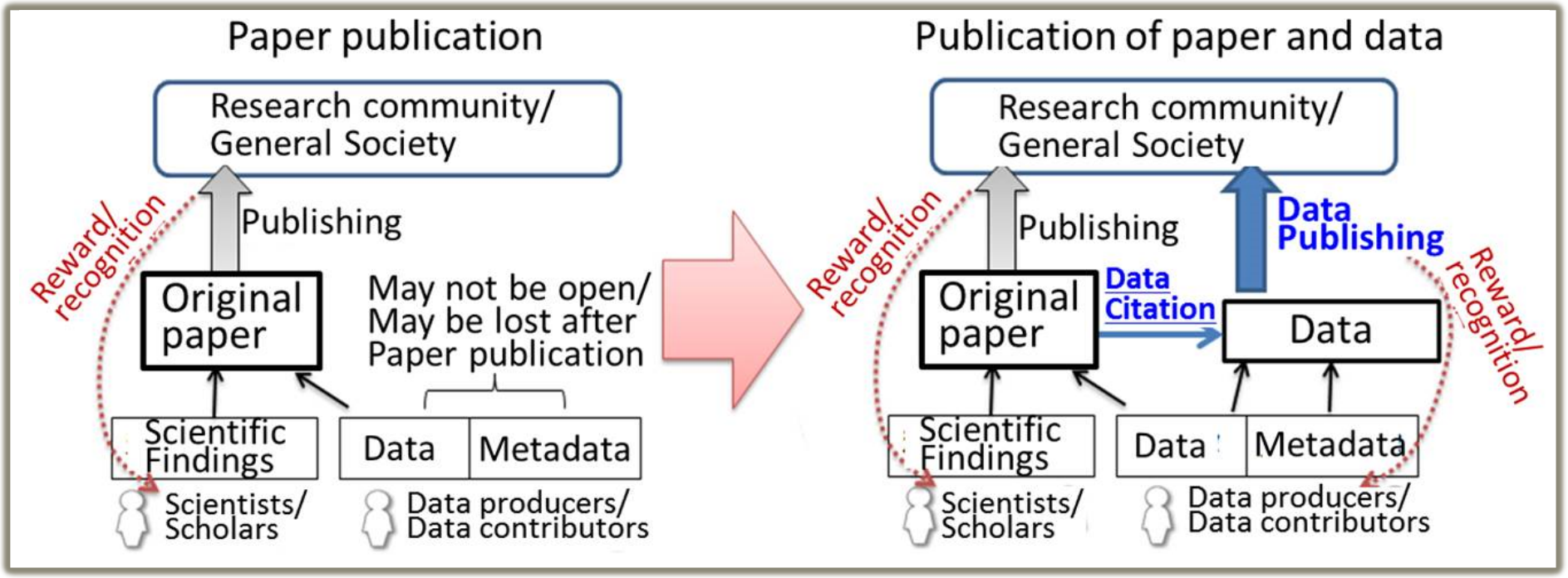
Print Media

Electronic Media

70 years

New global data initiatives: ICSU-WDS, RDA etc. : 2008～2013

“Data Publication” and “Data Citation”



[Society of Geomagnetism, Earth, Planetary and Space Sciences, 2013]

- Data Publications

  cf. journal publication: review, fix (print), publish with DOI…, metrics (citation index etc.)

- Data Citation

  – ID of dataset (“DOI” is OK?), citation standards? metrics?…

- More outputs from scientists to Society

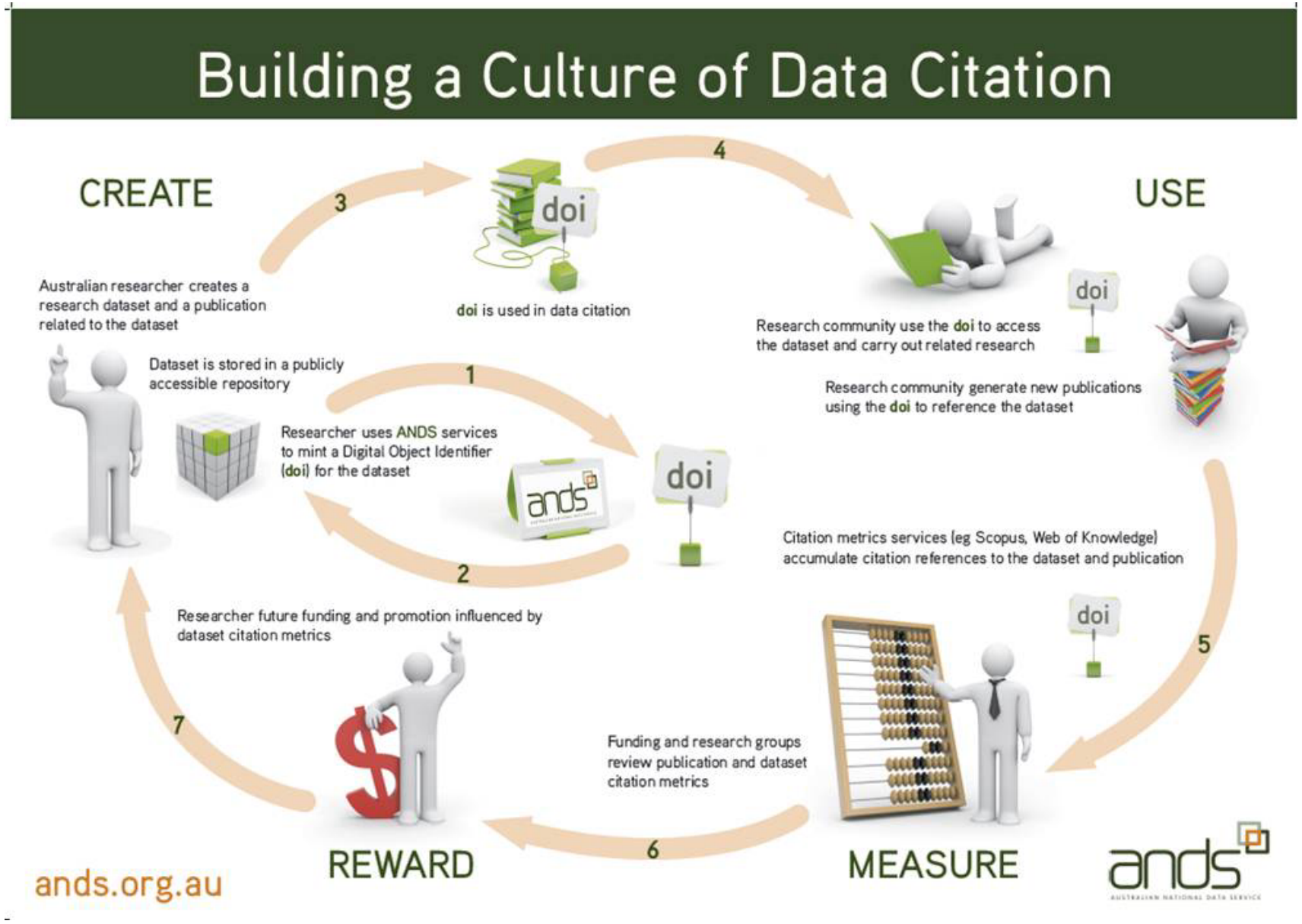


Illustration by Australian National Data Services (ANDS)

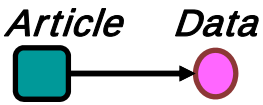
http://www.ands.org.au/cite-data/index.html

Steps by Major (Geophysical) Journals
encouraging data deposition & citation

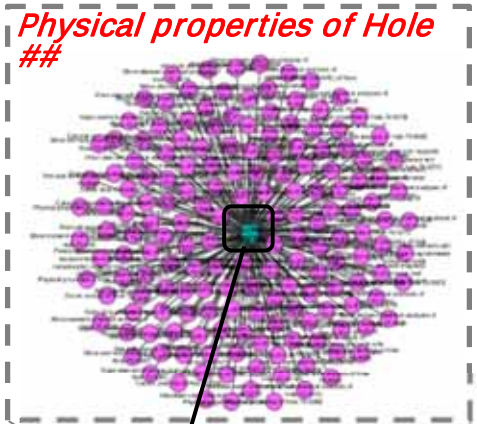
- **Willey/AGU** publication policy:
  "…in AGU's journals, **all data** necessary to understand, evaluate, replicate, and build upon the reported research **must be made available and accessible whenever possible**…"
- **SpringerOpen**/"Earth, Planets and Space", "Geoscience Letters"…
  "…Electronic archiving of data enables readers to replicate, verify and build upon the conclusions published in papers in the journal. **It is recommended that all data** which are not directly attached to a publication as electronic supplementary files **be deposited**…"
- **Elsevier**/JASTP:
  "…Elsevier encourages **authors to deposit raw experimental data sets** underpinning their research publication in data repositories, and to enable interlinking of articles and data…"

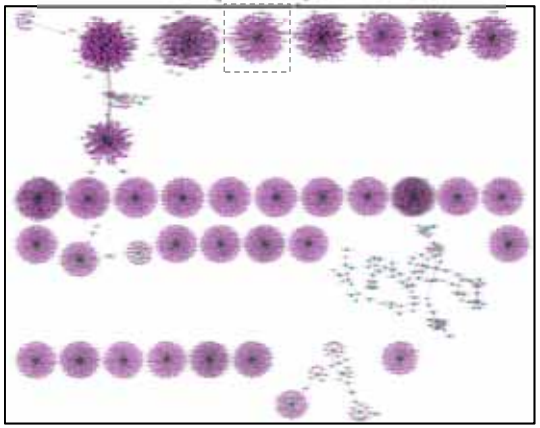
[Zettsu, Murayama, et al., CODATA Conference, 2014]

How datasets are cited by articles



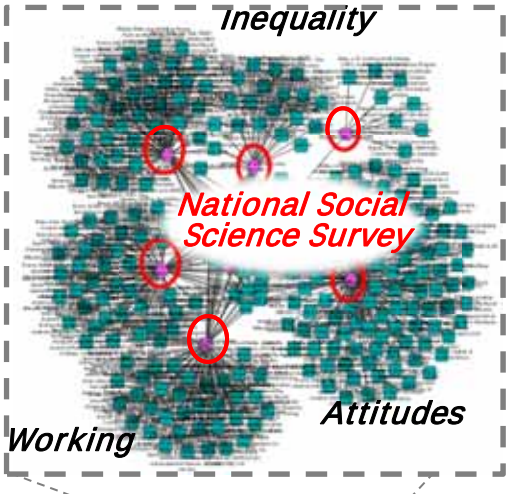
(a) Data collection community

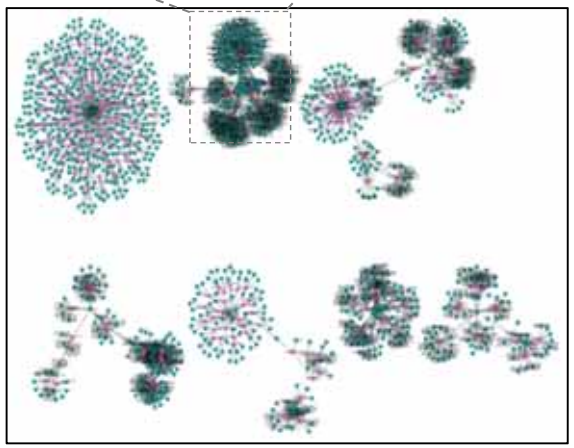




Pangaea
(http://www.pangaea.de/)
384,815 citations from OAI-PMH

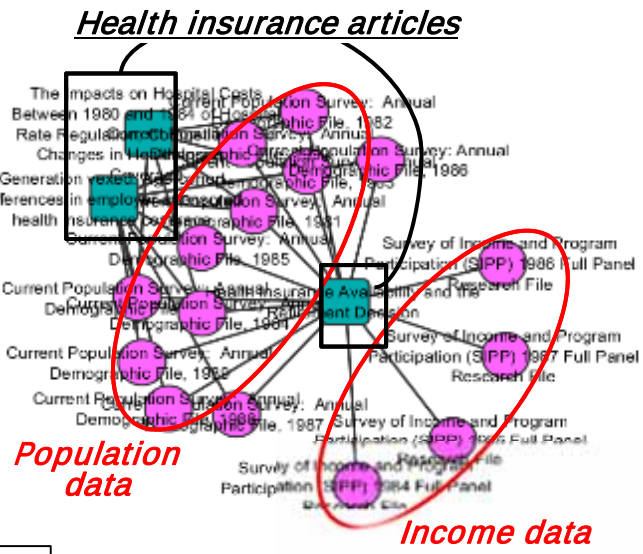
(b) Data sharing community

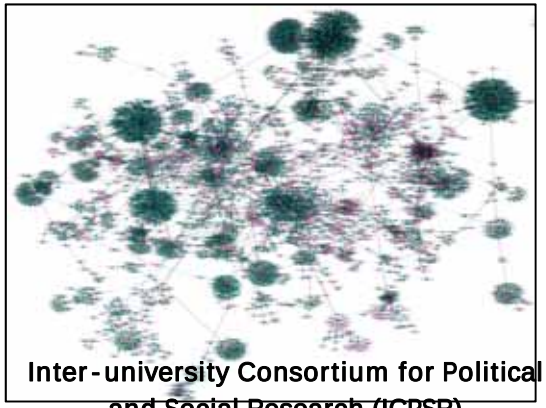




Australian Data Archive (ADA)
(http://www.ada.edu.au/)
16,062 citations from HTML

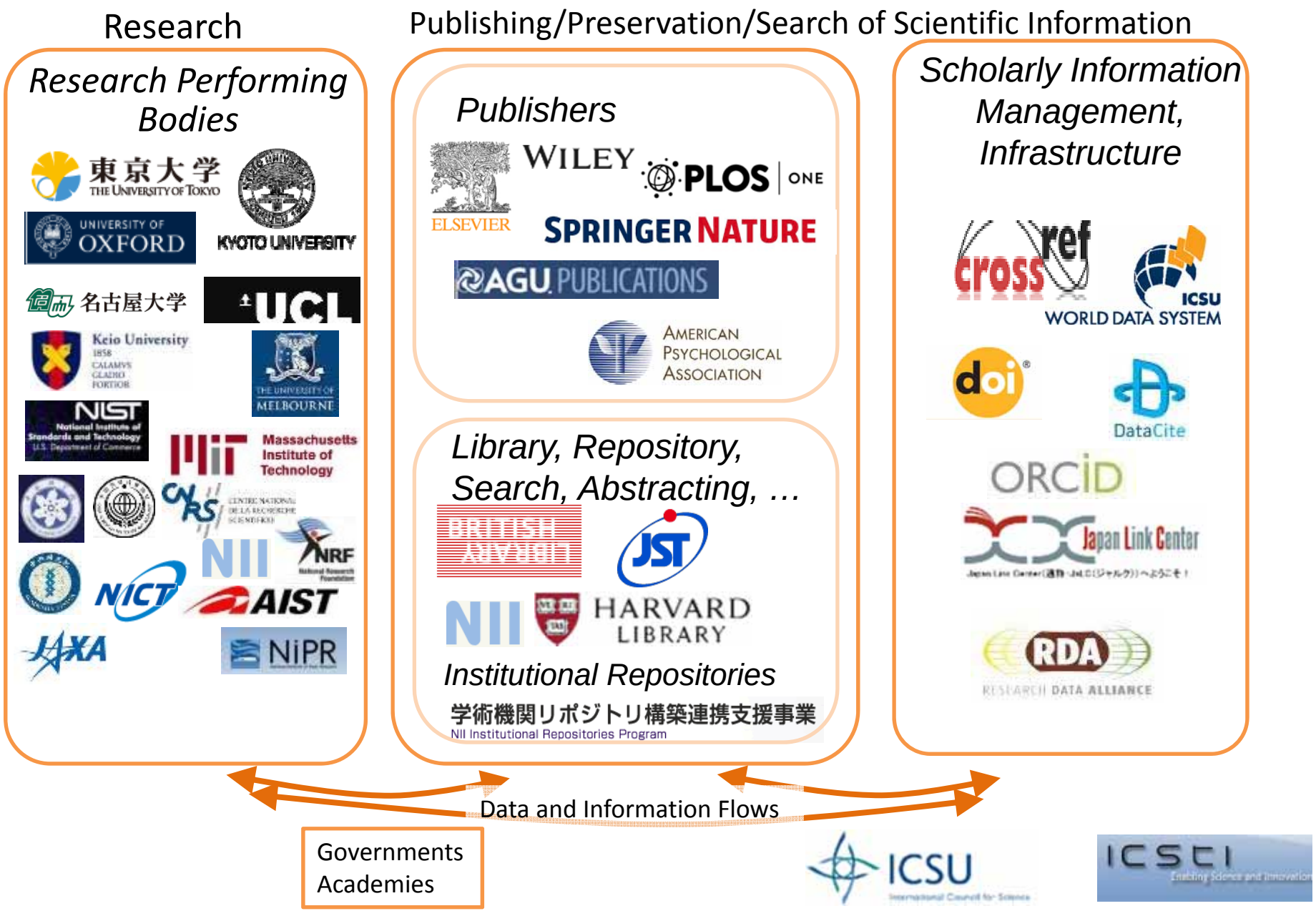
(d) Referential context





Inter-university Consortium for Political and Social Research (ICPSR)
(http://www.icpsr.umich.edu/)
115,154 citations from OAI-PMH

Science as a Social System (with “Print” Publication)



Concluding Remarks

- Open Science is an emerging focus of international Science & Technology policy
- A new mode of scholarly communication is required.
- Scholarly work eco-cycle
  - Create, Use, Measure, Reward/Recognition
  - Librarians, editors, publishers, data managers…
- Journal editors can play an active role for not only publishing articles, but also datasets behind them.

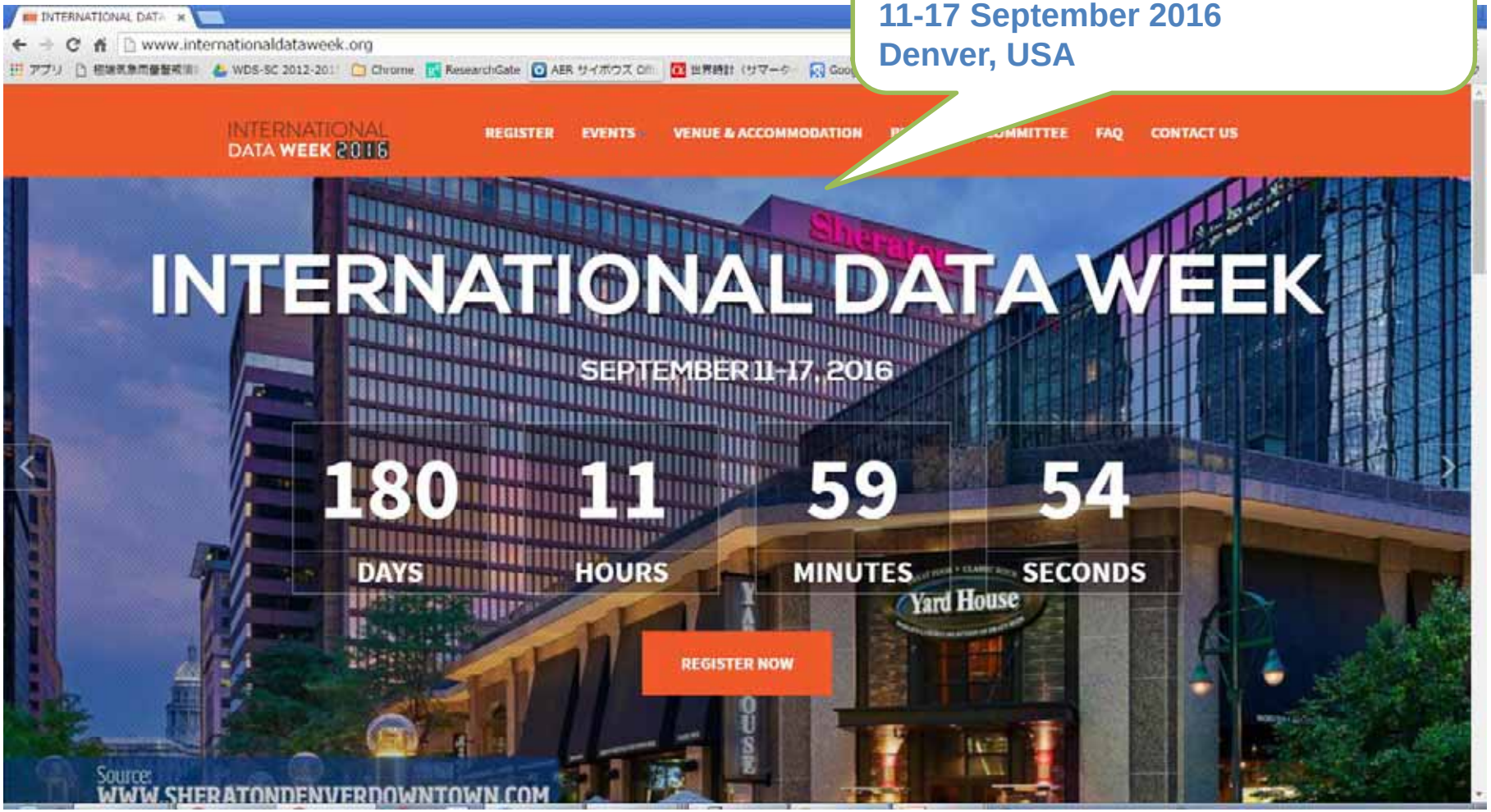
"Science is built of facts the way a house is built of bricks" (Henri Poincare, 1902)

http://www.ypo.co.uk/~/media/DADD106D742F424DBD46557ED5C3D6D0.ashx

http://www.threeland.com/images/travel-guide/notre-dame-cathedral.jpg

http://www.internationaldataweek.org/

RDA Plenary 8
***during* International Data Week 2016**
11-17 September 2016
Denver, USA



…And then, RDA Plenary 9
Date: 5-7 April 2017
Place: Barcelona, Spain