Open Science Landscape and Data Citation: working with International CODATA-ICSTI Data Citation Task Group
科学技術データ･学術データの保全事業とデータ引用

Dr. Yasuhiro Murayama

村山泰啓

ICSU-WDS Scientific Committee ex officio

Member of Cabinet Office Expert Panel of Open Science (2014)

Associate member, Science Council of Japan

Director, Integrated Science Data Systems Research Lab.
Natl. Inst. Information & Communications Technology

International Programme Office Hosted by NICT
Based in Tokyo, Japan

Toady's Talk

- Open Research Data
  (Science and International Policy/Society)
- "Data and Science" and Society
- CODATA-ICSTI data citation task group

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.

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

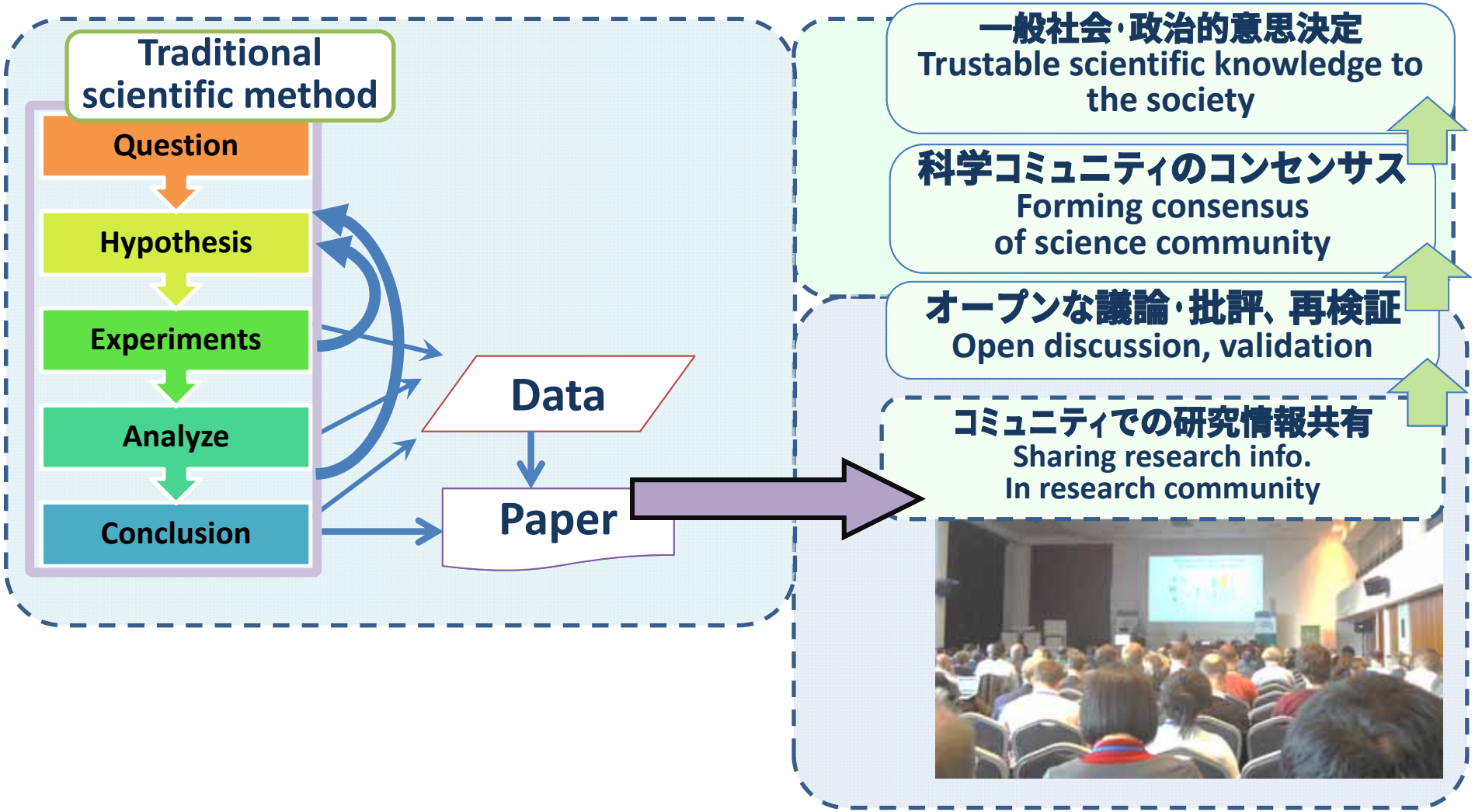
4. Expanding Access to Scientific Research Results

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

Five key principles outlines how governments should release datasets for economic and social benefits

"Open Government Data"

ある類型：サイエンスのあり方と社会との関わり方



1000人規模の科学者が気候変動の知識を作る

IPCC (Intergovernmental Panel on Climate Change)

WG1 “Physical Science Basis”

19 Headlines

on less than 2 Pages

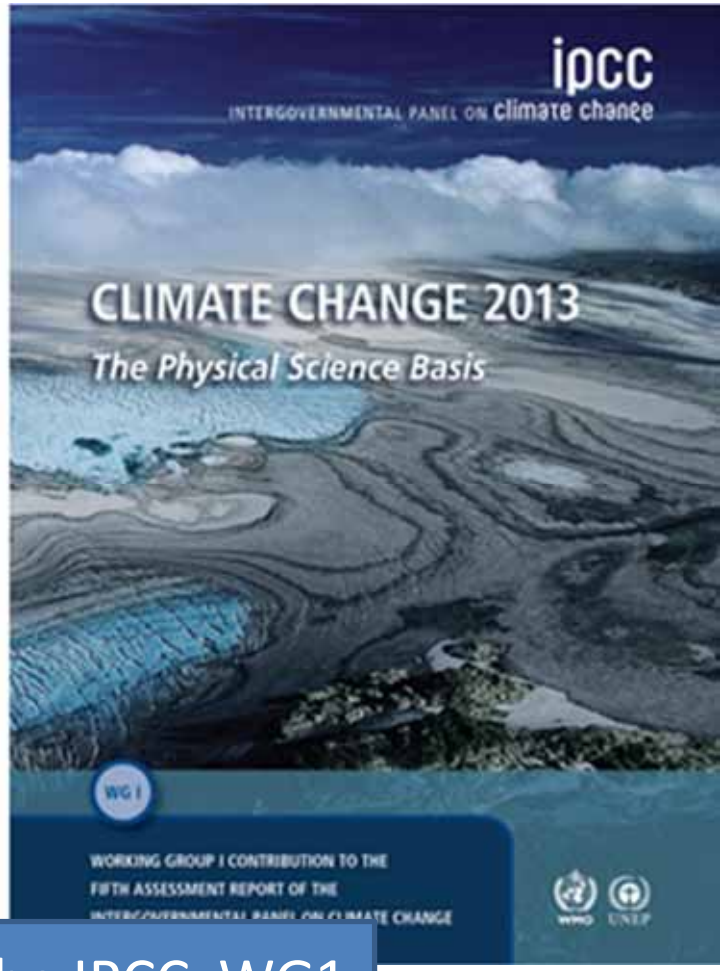
Summary for Policymakers
ca. 14,000 Words

14 Chapters
Atlas of Regional Projections

54,677 Review Comments
by 1089 Experts

2010: 259 Authors Selected

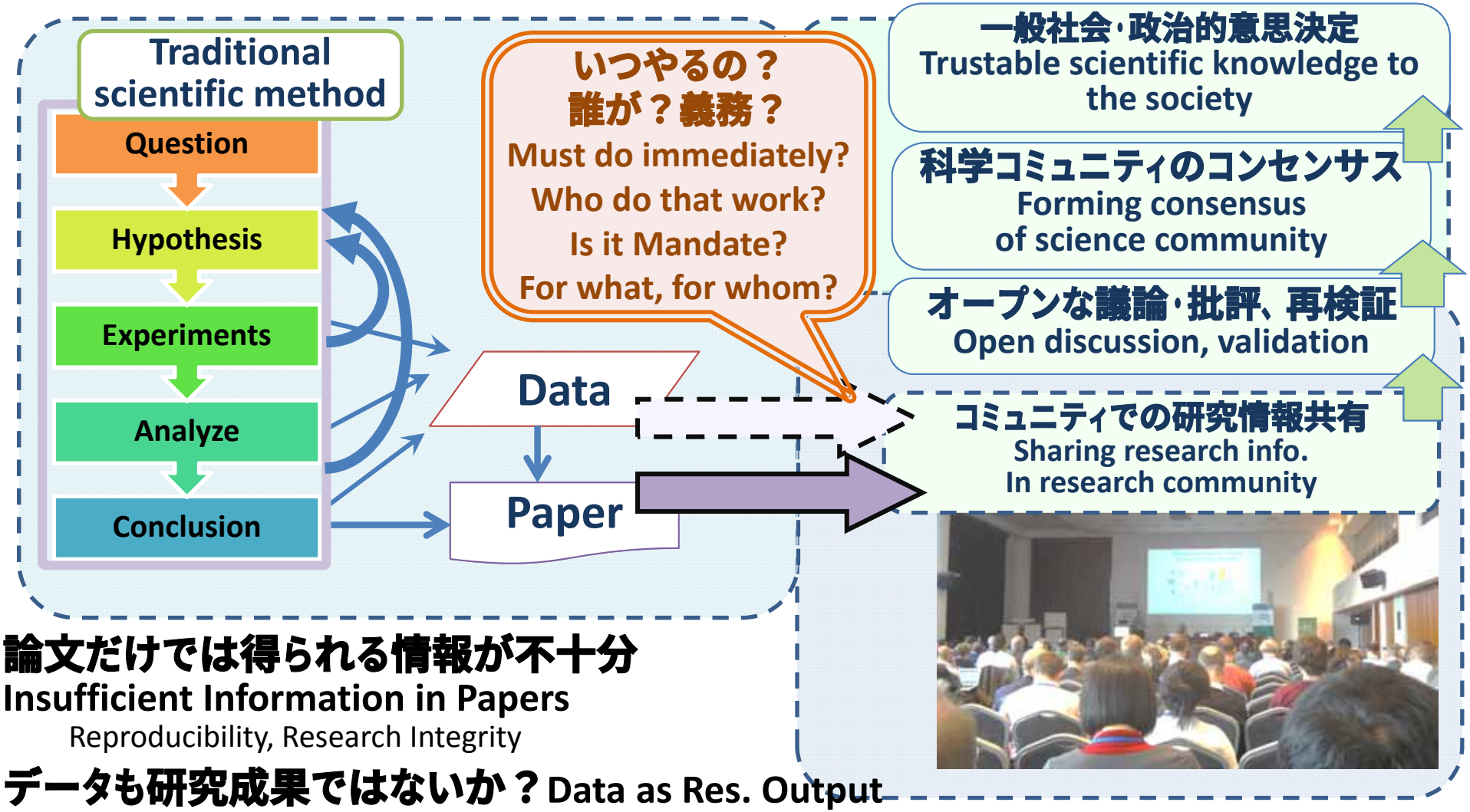
2009: WGI Outline Approved



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

[IPCC, 2013]

ある類型：サイエンスのあり方と社会との関わり方



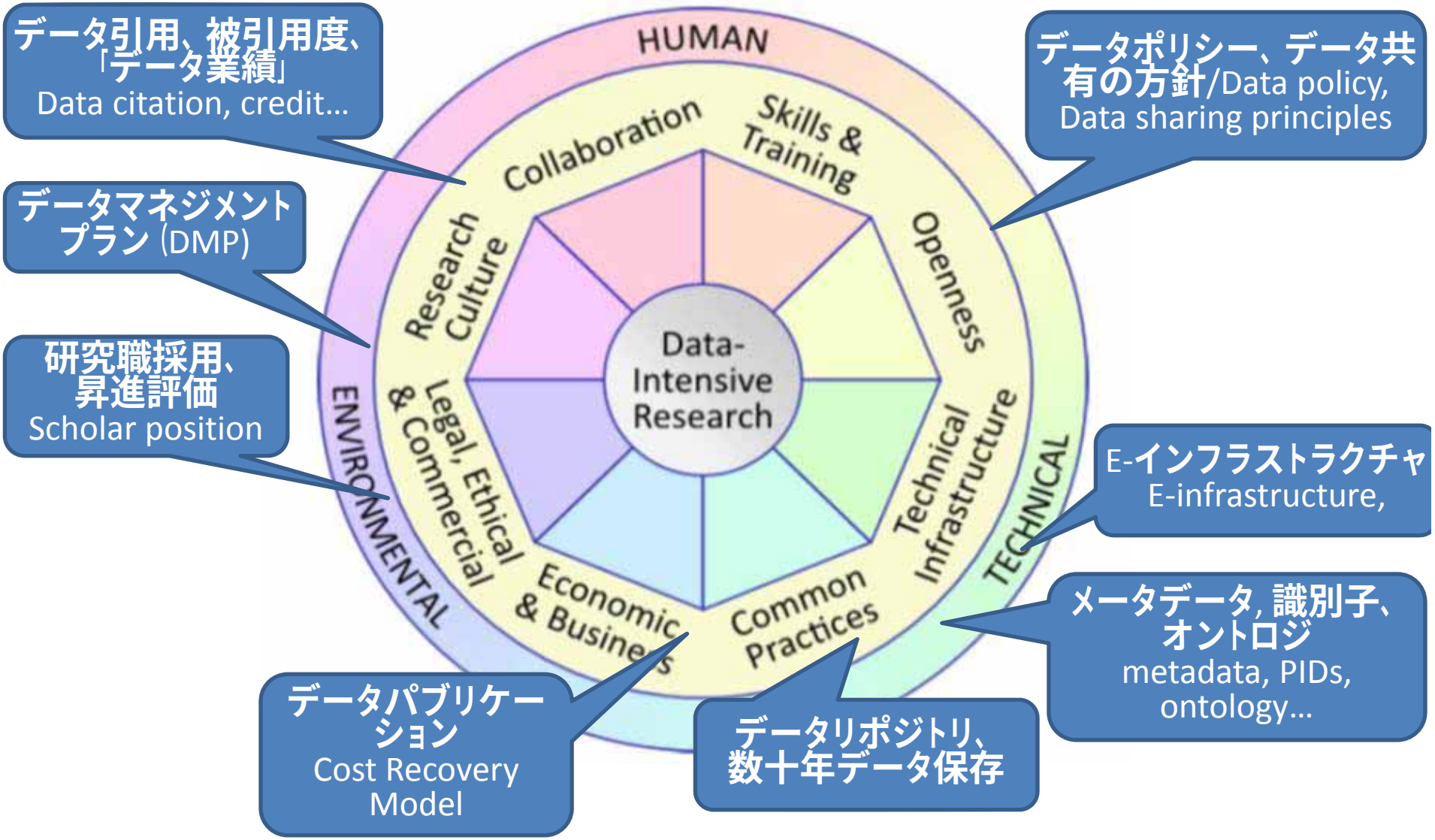
論文だけでは得られる情報が不十分
Insufficient Information in Papers
Reproducibility, Research Integrity

データも研究成果ではないか？ Data as Res. Output
Important Intellectual Asset, to be Shared with Society

再現できない事象の検証
Irreproducible Phenomena
e.g., Environment, Earth/Space, Living Organisms…

→ Data Management is also Question of Mutual Trust of Science and Society
データ問題⇔科学と社会の相互信頼

“Data and Science”



↑https://www.rd-alliance.org/filedepot_download/383/230

- RDA Community Capability Model Interest Group
  - Secretary: Univ. of Bath & Microsoft Research Connections

Toward "Global Community of Trusted Science Data Repository/Services for Long Term Preservation"

Creation of ICSU-World Data System

on ICSU 29th General Assembly decision (October 28, 2008):

supported by

PAST (since 1950's)	WDC (World Data Center) : 50 WDSs at max. FAGS (Federation of Astronomical and Geophysical Data Analysis Services)
PRESENT (2008~)	ICSU International Scientific Unions data bodies ICSU National Members data bodies ICSU Interdisciplinary Bodies data activities

92 Members (June 2015)

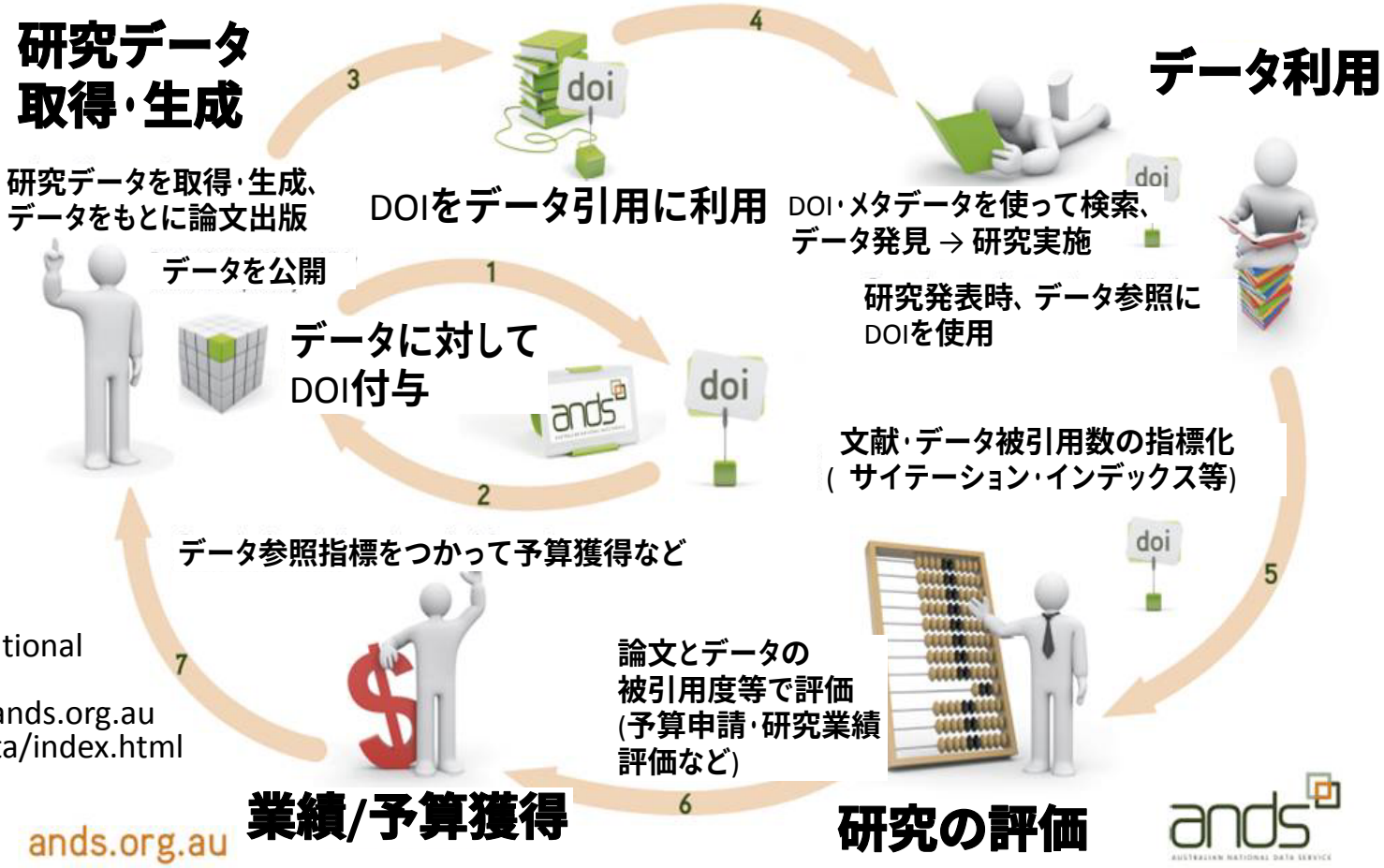
60 Regular	Data curation & data analysis services
10 Network	Networks of Regular Members & umbrella organizations
4 Partner	Do not deal directly with data stewardship, but support to ICSU-WDS
18 Associate	Organizations interested in the WDS endeavour

Research Data Alliance

- 2013年～
- Kicked by G8+O6 data infrastructure WG.
- Informatics, computer eng., librarian, lib. science, domain science (ocean, bio, astronomy, …)
- Platform for various communities/ stakeholders discussion

豪政府機関 Australian National Data Service による取組み例

データ・サイテーション（データ参照・引用）文化の形成へむけて



Australian National
Data Service
http://www.ands.org.au
/cite-data/index.html

（和訳は能勢（京大、2014）を参考にした）

○DOI（Digital Object Identifier）を論文だけでなく、データにも付与し、論文・文献で引用する取り組み
➔データ公開者・機関の活動評価（論文と同等に）とクレジット

Other Organizations Working on Data Citation

- **International Council for Scientific and Technical Information (ICSTI)**
- **DataCite**
- **The Dataverse Network**
- **National Information Standards Organization (NISO)**
- **Creative Commons and Science Commons**
- **CENDI –** U.S. interagency group
- **Global Biodiversity Information Facility (GBIF)**
- **World Data System (WDS)**
- **STM-Association**
- **Digital Curation Center, UK**
- **Research Data Alliance (RDA)**

(as of Sept. 2013)
[modified from Jan Brase (2013)]

… and many more

http://www.codata.org/news/59/62/International-Series-of-Workshops-to-Implement-Data-Citation-Principles

CODATA

International Council for Science . Committee on Data for Science and Technology

ABOUT · EVENTS · MEMBERSHIP · COMMITTEES · TASK GROUPS · WORKING GROUPS · PU

CODATA-ICSTI Data Citation Standards and Practices

The growth of electronic publishing of literature has created new challenges, such as the need for mechanisms for citing online references in ways that can assure discoverability and retrieval for many years into the future. The increase in online datasets presents related, yet more complex challenges. Data citation standards and good practices can also form the basis for increased incentives, recognition, and rewards for scientific data activities that in many cases are currently lacking in all fields of research.

Between 2011 and 2014, the task group published two major reports, *For Attribution: Developing Data Attribution and Citation Practices and Standards* (2012); and *Out of Cite Out Of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data* (2013). These reports laid out the landscape of research data attribution and citation issues, practices, and policies. The group was then instrumental in convening a synthesis group of organizations that cooperated in developing the **Joint Data Citation Principles**.

DC1

Data Citation Principles

Objectives

The main objective in the next period of activity, April 2015-October 2016, is to promote the implementation of the data citation principles in the research policy and funding communities throughout the world. We plan to hold at least 11 national and regional workshops dedicated to this focused objective and to write a synthesis paper that integrates the findings.

We have received verbal commitments to hold such a meeting in 2015 from representatives affiliated with our group, who plan to organize one-day workshops for the science policy and funding community in their respective countries, and following a similar model: **China, Taiwan, Japan, India, Australia, South Africa, Israel, Brazil, EU, France and USA**

Christine Borgman (Co-Chair)

Distinguished Professor & Presidential Chair in Information Studies,
University of California, Los Angeles
Department of Information Studies,
Graduate School of Education and Information Studies
235 GSE&IS Bldg, Box 951520, Los Angeles, CA 90095-1520, USA
Phone: (310) 825-6164 / Fax: (310) 206-4460
christine.borgman (at) ucla.edu

Dr. Jan Brase (Co-Chair, ICSTI representative)

Head of Research and Development
Georg-August-University Göttingen
Göttingen State and University Library
D-37070 Göttingen, Germany
Papendiek 14 (Historical Building, room 2.213)
Tel: +49 551 39-33878
brase (at) sub.uni-goettingen.de

Martie J. van Deventer (Co-chair)

Portfolio Manager, Information Services
CSIR South Africa
mvandeve (at) csir.co.za

データ引用基本方針に関する国際合同宣言

Joint Declaration of Data Citation Principles

(Data Citation Synthesis Group, hosted by FORCE11, March 2014)

https://www.force11.org/group/joint-declaration-data-citation-principles-final

1. Importance
2. Credit and Attribution
3. Evidence
4. Unique Identification
5. Access
6. Persistence
7. Specificity and Verifiability
8. Interoperability and Flexibility

[modified from Jan Brase (2013)]

CODATA-ICSTI Task Group organizes <u>This</u> International Series of Workshops to Implement Data Citation Principles

<u>Objective:</u>

- Implementation of DC principles in the research policy and funding communities (研究予算・政策関係コミュニティでのデータ引用方針の実践へ向けて)
- To establish how principles may best be turned into practice in country （原則から実践へ）

Each workshop invites stakeholders including government（政府）, funders（予算機関）, research institutions（研究機関）, research data archives（データ保存機関）, learned societies（学会）, publishers（出版事業）, journal editors（ジャーナルのエディター）etc.

International Series of Workshops to Implement Data Citation Principles

- **China,** Lan Zhou, after the Second National Scientific Data Conference, **25 August 2015** (including remote introductory presentation from Christine Borgman)
- **Australia, 28 October 2015** (including introductory presentation in person from Simon Hodson).
- **Japan, Tokyo, 29 October 2015** (includng keynote, in person, from Christine Borgman).
- **India, New Delhi, 5-6 November 2015** (including keynote, in person, from Jan Brase).
- **South Africa, Pretoria, December** (TBC, including remote introductory presentation from Jan Brase).
- **USA, possibility of March/April 2016** at Oak Ridge.

. . .

ではデータ引用を日本で進めるには？

Then, how to promote Data Citation in JP?

- **政府、予算機関の科学成果把握の現状**
  current status of government, funding agency
- **研究者、論文文献を扱う側の現状 (学会?)**
  those who do research/science, those who handle literature (library, publisher…)
  (professional societies' consensus?)
- **データ引用・データマネジメント、プラットフォームの実験、試行**
  trials/experiments of data citation and management/platform
- **国際展望と将来の研究評価、成果創出**
  Intl. perspectives and Japan, future research styles, research outcomes

一方で：
オープンサイエンスは絶対正義か？

- 理想像と現実
- 成果があがらない取り組みはしてもしょうがない？
- 「全体最適」と「局所最適」の両方が大事

長期的に大事なこと

今現在大事なこと

- 遠い理想を心に抱きつつ、今できるわずかなことを行う

その戦略を最適化したい

- 研究者が、自分の研究・学会・社会のためにベストだと思うやり方を探してみよう・考えておこう
- 自分たちの後輩にも「理想像」を伝えたい。

PLEASE KEEP IN MIND:

- DON'T **discourage/decelerate research.**
  - Any rules or regulations should not discourage scientists/researchers.
    - Which data should be open/close, embargo period, use condition, designing services for users…
- **DON'T regulate what we don't yet understand.** ("The Data Harvest", RDA Europe, December 2014)

RDA

RESEARCH DATA ALLIANCE

Research Data Alliance
Plenary Meeting 7
1- 3 March 2016, Tokyo, Japan

research data sharing without barriers

rd-alliance.org

シンポジウム：

科学の発展への起爆剤

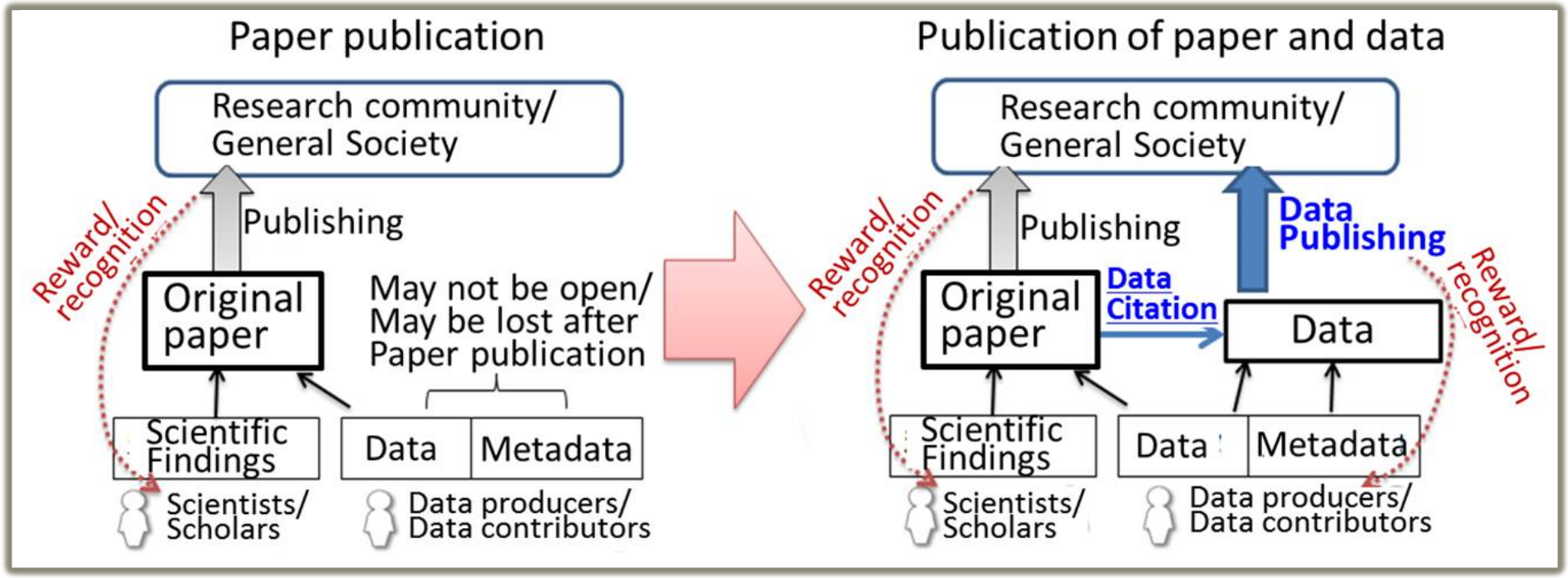
～データ駆動型科学の推進に向けて～

内閣府
文部科学省
国立研究開発法人 科学技術振興機構
国立研究開発法人 産業技術総合研究所
国立研究開発法人 情報通信研究機構
国立研究開発法人 物質・材料研究機構
大学共同利用機関法人 国立情報学研究所

2016年2月29日
一橋講堂

FIN.

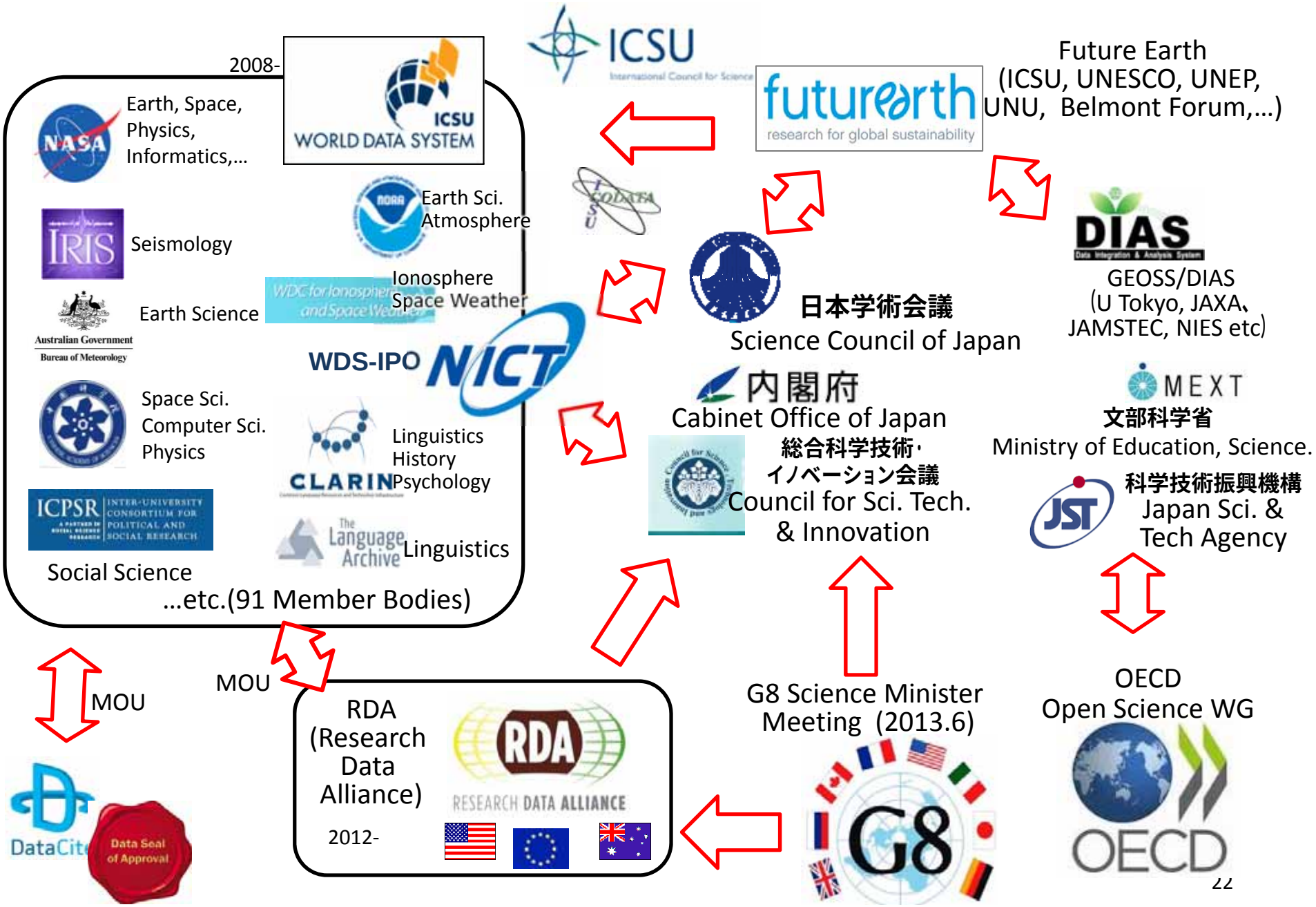
“Data Publication” and “Data Citation”



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

- Data Publications
  - Data to be counted as research product/outcome (cf. journal publication)
- Data Citation
  - Recognition of intellectual work by data creators/contributors/research institutes (cf. research paper citation)
- More outputs from scientists to Society

Landscape of WDS and Open Science/Open Data (from my viewpoint)



Inauguration of WDS Intl. Programme Office

ICSU (Intl. Council for Science) has established ICSU-WDS (October 2008).

Science Council of Japan (SCJ) cooperated hosting WDS Intl. Programme Office at NICT, Japan.

WDS-IPO Executive Director: Dr. Mustapha Mokrane

Scientific Committee

[Y. Murayama, 2012]
[F. Kasuga, 2013]

Example of DOI-minting to Earth Science database

- Mesospheric wind velocity data (30min. mean) observed with MF radar at Poker Flat, Alaska

First Data-DOI Registration by Japanese Platform

doi:10.17591/55838dbd6c0ad

http://www2.nict.go.jp/isd/doi-landingpage/wds/10.17591__55838dbd6c0ad.html

Mesospheric wind veloc ×

www2.nict.go.jp/isd/doi-landingpage/wd

Mesospheric wind velocity data (30min. mean) observed with MF radar at Poker Flat, Alaska

Horizontal wind velocity in the altitude range of approx. 60-90 km is observed with Poker Flat MF (medium frequency) radar, using the radar wave at 2.43 MHz. The radar receives weak radio echo signals returned from the weakly ionized atmosphere (ionospheric D-region) at the target altitudes, to deduce horizontal air motions (Murayama, Y., K. Igarashi, D. D. Rice, B. J. Watkins, R. L. Collins, K. Mizutani, Y. Saito, and S. Kainuma, Medium Frequency Radars in Japan and Alaska for Upper Atmosphere Observations, IEICE Trans., E83-B, pp.1996-2003, 2000). Poker Flat MF radar has been constructed as part of Japan-US joint research program of Arctic middle & upper atmosphere ("Alaska Project") in collaboration between National Institute of Information and Communication Technology, Japan (formerly Communications Research Laboratory), and Geophysical Institute, University of Alaska Fairbanks.

Landing Page

Data Citation

Citation: Alaska Project of NICT (CRL)-GI/UAF, Mesospheric wind velocity data (30min. mean) observed with MF radar at Poker Flat, Alaska, doi:10.17591/55838dbd6c0ad

General Characteristics

Parameters: Mesospheric horizontal wind velocity

Processing level:

Latitude: 65.1

Longitude: -147.5

Temporal resolution: 30 minutes

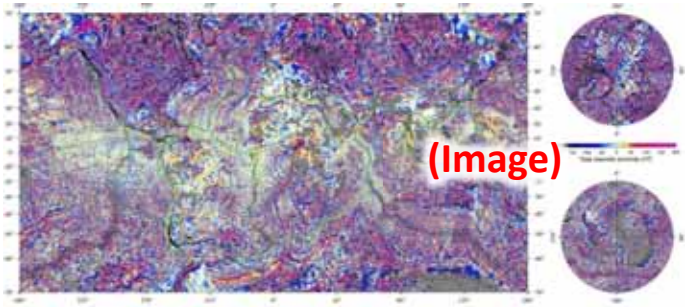
Data description, Data format, Link to data, etc.

Digital data

-55.033333	-89.900000	-56.134989
-55.000000	-89.900000	-56.127400
-54.966667	-89.900000	-56.119808
-54.933333	-89.900000	-56.112213
-54.900000	-89.900000	-56.104616
-54.866667	-89.900000	-56.097016
-54.833333	-89.9[illegible]	[illegible]413
-54.800000	-89.9[illegible]	[illegible]806
-54.766667	-89.900000	-56.074197
-54.733333	-89.900000	-56.066584
-54.700000	-89.900000	-56.058968
-54.666667	-89.900000	-56.051348
-54.633333	-89.900000	-56.043725
-54.600000	-89.900000	-56.036097
-54.566667	-89.900000	-56.028466
-54.533333	-89.900000	-56.020832
-54.500000	-89.900000	-56.013193
-54.466667	-89.900000	-56.004623

(Image)

Data plot



(Image)

data citation

Cited by Kinoshita, T., Y. Murayama, and S. Kawamura (2015), Tidal modulations of mesospheric gravity wave kinetic energy observed with MF radar at Poker Flat Research Range, Alaska, J. Geophys. Res., 120, doi:10.1002/2014JD022647