Open Science in Japan and Trials for Building an Open Research Information System at NISTEP

HAYASHI, Kazuhiro
NISTEP Research Unit for Data Application
Jan. 23th, 2025

The 133rd GIST Seminar The Barcelona Declaration: Building an Open Research Information System by Europe and Japan

Introduction and Items

Introduction

1. **Open Science in Japan**
   - ✓ Policy:
   - ✓ Data Platform:
   - ✓ Monitoring Surveys:
2. **Open Research Information System at NISTEP**
   - ✓ utilizing Open Data for EBPM
   - ✓ Development of new survey research and data analysis methods
3. **Summary (and one more thing)**

Introduction

As a digital transformer for Open Science （Kaz Hayashi）

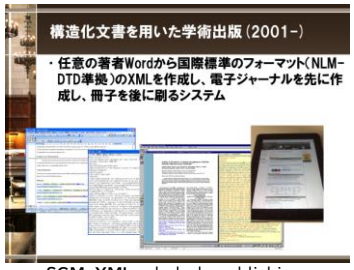
Mosaic
Amazo
W3C
Google
Facebook
Youtube
Twitter
LINE
AlphaGo
ChatGPT
AGI

1990
1995
2000
2005
2010
2015
2020

Contribution

- Organic Syntheses
- **Digital Peer-Review Tracking System**
- Online journals
- **XML Publishing**
- DOI and CrossRef
- EJ-business
- **Open Access 2005**
- Advocacy for policy makers
- ORCID
- altmetrics
- RDA Tokyo 2016
- **G7 OSWG**
- OECD
- **Citizen Science**
- blockchain
- Virtual Learned Society

DX for journals with XML publishing



SGM, XML scholarly publishing

Advocacy for Science Council of Japan

To Open Access and sharing Data

International Contribution For Open Science Policy

Japan Open Science Summit

DX for Learned Society

Blockchain taskforce for Chemistry

Re-Openness of Science (citizen science, virtual society, DeSci)

Theme

Digitalization of journals

DX for society and media

DX for research outputs focusing on Data and multidimensional Impacts

DX for PTA

Citizen Science（NHK）

DX for research and community

Decentralized Science **blockchain**

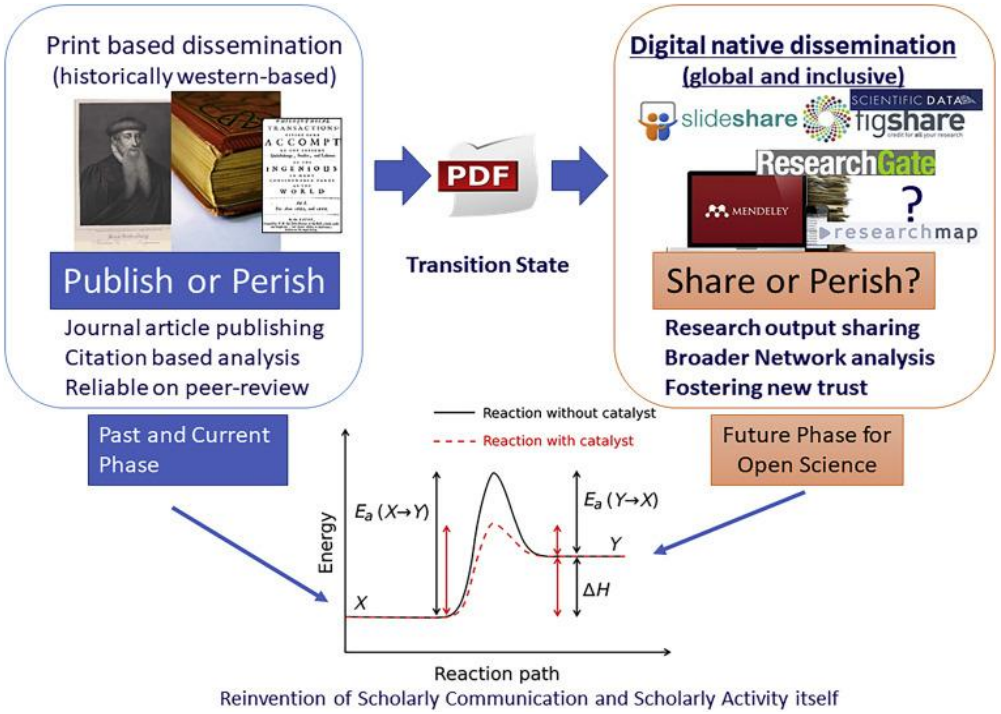
Scholarly communication on **Metaverse**

Wish to know how science and society would be changed eventually

TEDxKioicho

Vision: Reliable and practical Phase Transfer

Phase Transfer

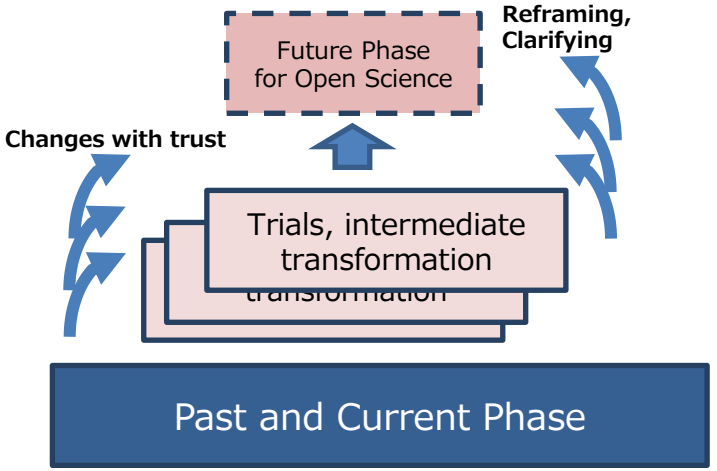


Reinvention of Scholarly Communication and Scholarly Activity itself

Additive Transformation

(in the early stage)

To ambiguous but promising phase



From established and reliable Phase

https://doi.org/10.1016/j.patter.2020.100191

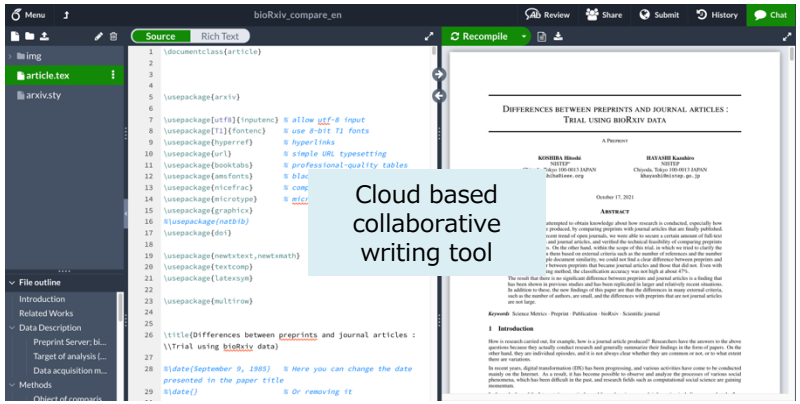
Steady dialogue among stakeholders for **behavior changes**
(with exploiting AI and other technologies on the other hand)

NISTEP Seminar
(Selected excerpts for Open Science)

- 2013 Mark Hahnel (figshare)
- 2013 Jason Priem (ImpactStory)
- 2013 Timo Hannay (Digital Science)
- 2017 Laurel Haak (ORCID)
- 2018 Rebecca Thomas (F1000 Research)
- 2019 Shelley Stall (AGU)
- 2019 Hilary Hanahoe (RDA)
- 2023 Reslie Macintosh(Ripeta)
- 2023 John Willinsky (PKP)
- 2023 Daniel Hook (Digital Science)

Research Unit for Data Application (RUDA)

- **Brand new unit for DX**
  - Launched on April, 2021
  - Light-weight
    - Director 1, researcher (data scientist) 1, other staffs 2
      - Affiliated Fellow 10 as virtual staffs
  - Online friendly since before COVID-19



〈2021.4-〉

Director General

Deputy Director General

Senior Fellow

〈Research Sector〉

1st Theory-Oriented Research Group
Theoretical and empirical studies and research on effects of S&T on the economy and society

2nd Theory-Oriented Research Group
Statistical studies and research on R&D promotion systems of S&T

1st Policy-Oriented Research Group
Empirical studies and research on conditions for promotion of S&T, such as personnel, relationship of S&T with people and society

2nd Policy-Oriented Research Group
Empirical studies and research on industry-academia collaboration, academic start-ups, regional innovation

Center for S&T Foresight and Indicators
Theoretical / Empirical studies and research on S&T trends and foresight
Theoretical / Empirical analysis and research on S&T activities of S&T policies

Research Unit for Data Application
Studies and research for Data Application on S&T policies

〈Research Support Sector〉

General Affairs Division
Personnel, welfare, archives, accounting, supplies, and maintenance

Planning Division
General coordination and planning of surveys and research, PR

Research Unit for Data Application (RUDA)

Mission

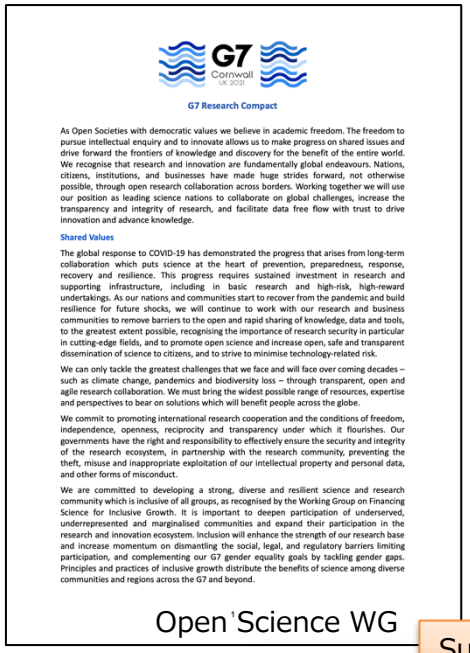
- Surveys and research that capture new research trends focusing on Open Science and AI
  - **Leading Open Science Policy and its monitoring**
  - **Development of new survey research and data analysis methods based on Data Science, AI-related technology**
  - Building a space for dialogue that promotes the transformation of science, society, and science and society oriented by open science
- DX for NISTEP
  - Support other research Groups and Units to exploit the potential of Data Science, AI and Open Science

Small unit as Enhancer or Catalyzer for transformation

1. Open Science in Japan: Policy, Data Platform, Monitor

Open Science as the global agenda for the future

G7 Research Compact



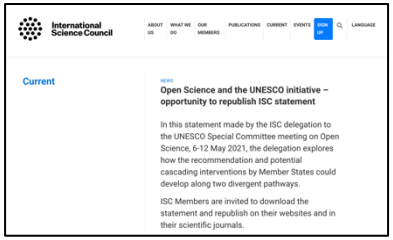
Open Science WG

UNESCO

Supporting as an expert

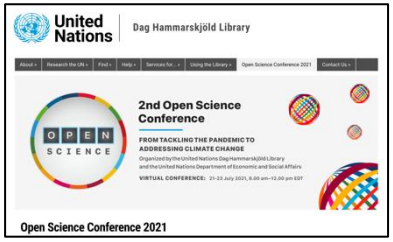
ISC
(International Science Council)

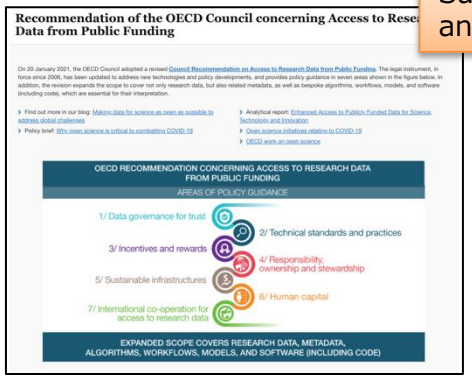
support



UN

support





OECD

https://www.mofa.go.jp/mofaj/files/100200013.pdf
https://en.unesco.org/science-sustainable-future/open-science/recommendation
https://council.science/current/news/open-science-and-the-unesco-initiative/
https://www.un.org/en/library/OS21
https://www.oecd.org/sti/recommendation-access-to-research-data-from-public-funding.htm

Entering an era of Open Science as a Default

The 6th Science, Technology and Innovation Basic Plan

Open Science as a significant key driver for STI policy(Society 5.0)

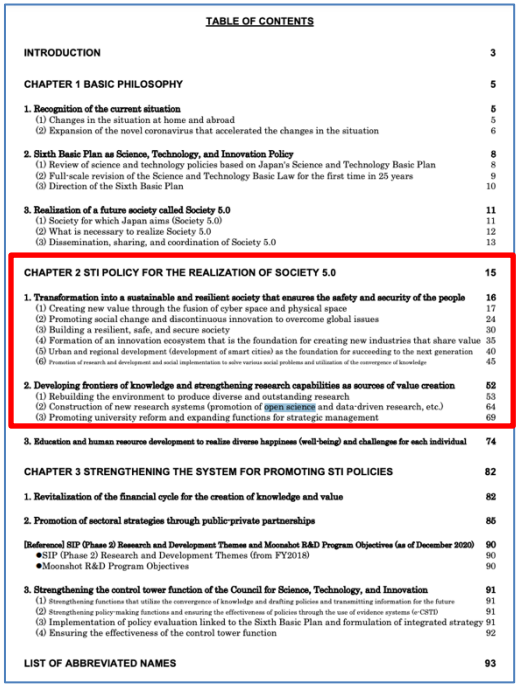


TABLE OF CONTENTS

CHAPTER 2 STI POLICY FOR THE REALIZATION OF SOCIETY 5.0 15

1. Transformation into a sustainable and resilient society that ensures the safety and security of the people 16
(1) Creating new value through the fusion of cyber space and physical space 17
(2) Promoting social change and discontinuous innovation to overcome global issues 24
(3) Building a resilient, safe, and secure society 30
(4) Formation of an innovation ecosystem that is the foundation for creating new industries that share value 35
(5) Urban and regional development (development of smart cities) as the foundation for succeeding to the next generation 40
(6) Promotion of research and development and social implementation to solve various social problems and utilization of the convergence of knowledge 45

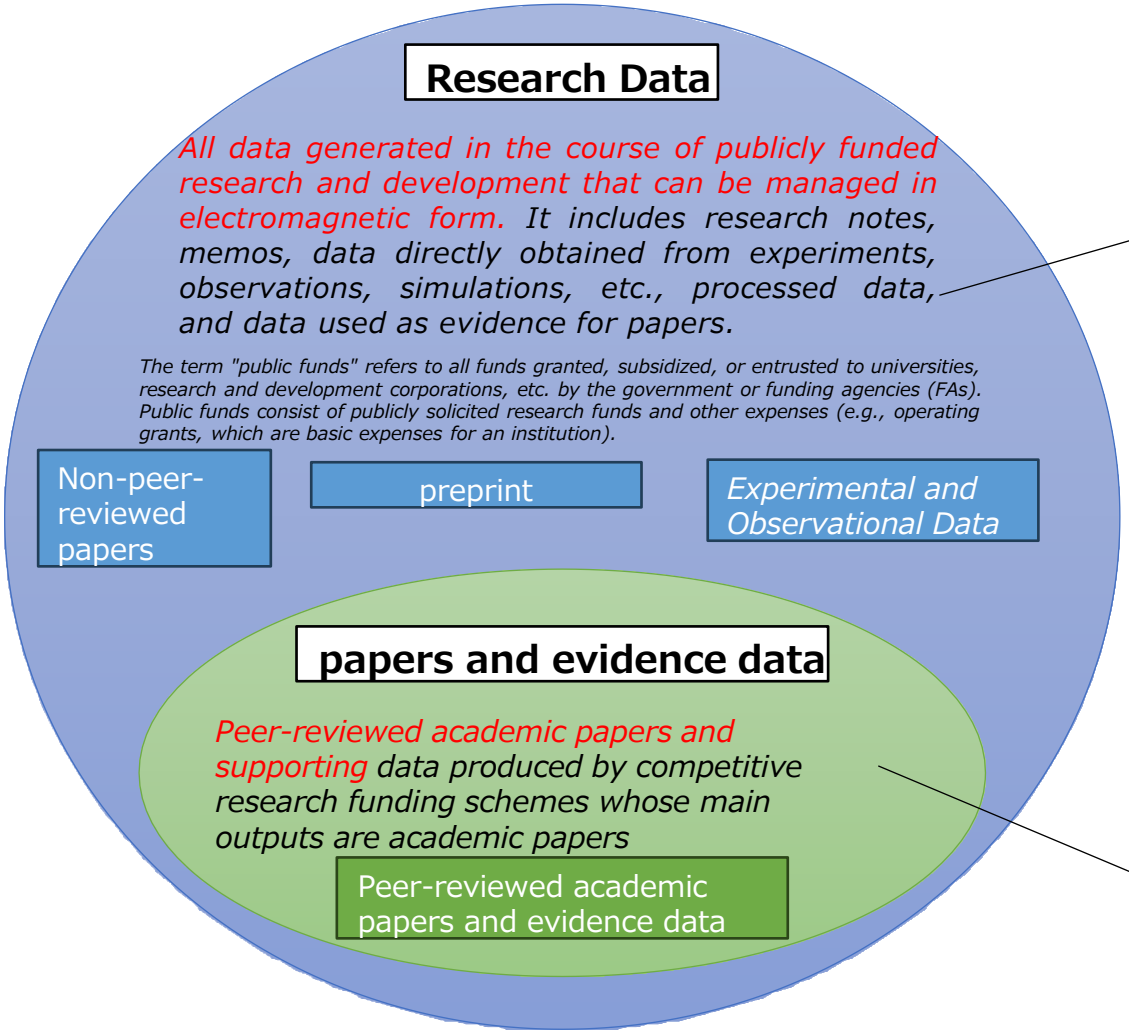
2. Developing frontiers of knowledge and strengthening research capabilities as sources of value creation 52
(1) Rebuilding the environment to produce diverse and outstanding research 53
(2) Construction of new research systems (promotion of open science and data-driven research, etc.) 64
(3) Promoting university reform and expanding functions for strategic management 69

3. Education and human resource development to realize diverse happiness (well-being) and challenges for each individual 74

Beyond Open Access, rather directing disruptive transformation of scholarly communication, which drives innovation

Relationship between open access to publicly funded academic papers, etc. and the management and utilization of research data

Tentative translation



Basic Approach to the Management and Utilization of Publicly Funded Research Data" (decided by the Council for the Promotion of Integrated Innovation Strategies on April 27, 2021)

- ✓ Position NII RDC as a core platform to make metadata searchable
- ✓ Open/shared/unshared settings based on open/closed strategy
- ✓ Researchers identify data to be managed and grant metadata
- ✓ Introduction of a system for granting metadata to publicly solicited research funds
- ✓ Establishment of data policies at universities and other institutions, etc.

Basic Policy for the Realization of Immediate Open Access to Academic Papers and Other Documents" (decided by the Council for the Promotion of the Integrated Innovation Strategy on February 16, 2024)

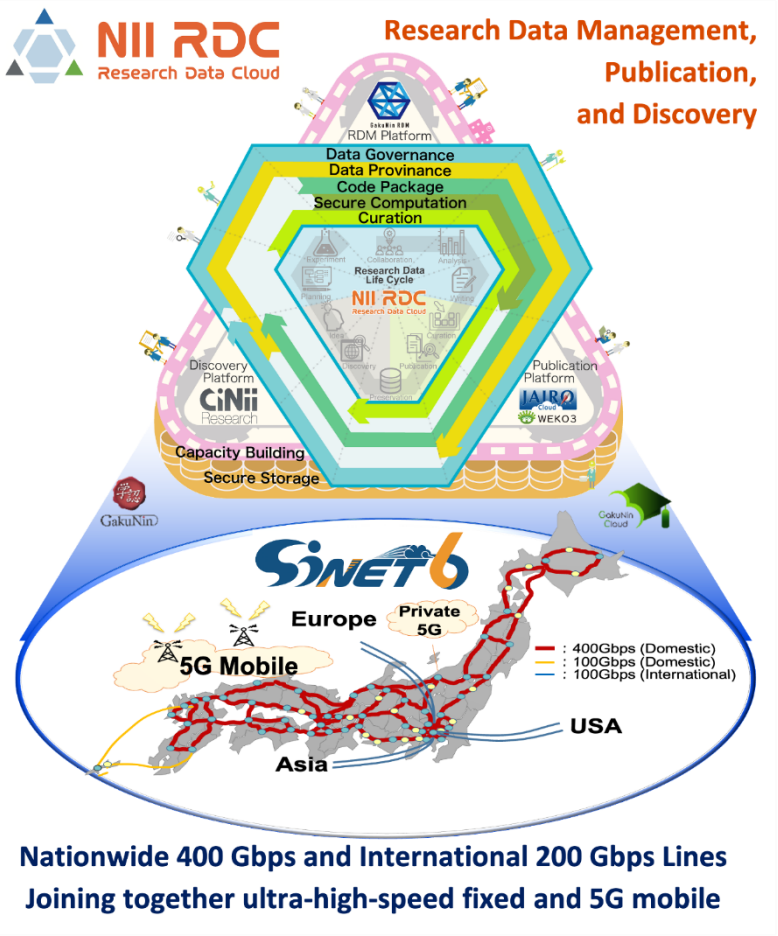
- ✓ *Require recipients of competitive research funds (including corporations) whose main results are academic papers to post academic papers and evidence data in an information infrastructure* such as an institutional repository immediately after publication in an academic journal.*
- ✓ *The aim is to enable everyone to freely utilize the results of research through the posting of academic papers and evidence data in information infrastructures such as institutional repositories.*
- ✓ *Support the development and enhancement of platforms for managing and utilizing research results, such as the research data infrastructure system (NII Research Data Cloud), other preprints, and academic papers, as a means of disseminating research results so that they can be freely utilized by everyone. etc.*

*The information infrastructure, such as institutional repositories, is the research data infrastructure system (NII Research Data Cloud), which is positioned as "Japan's core platform for the management and utilization of research data" in the Sixth Science, Technology and Innovation Basic Plan (approved by the Cabinet on March 26, 2021). (2) Academic papers and evidence data shall be searchable on the NII Research Data Cloud.

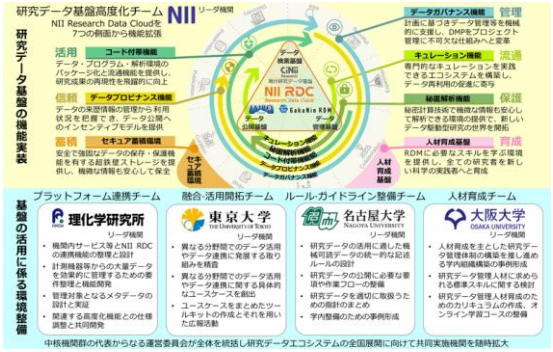
https://joss.rcos.nii.ac.jp/session/data/JOSS2024_B3_2.pdf

Research Data Platform towards Open Science Paradigm

Infrastructure



Application



Promoting the use of AI, etc.Research Data Ecosystem Development Project

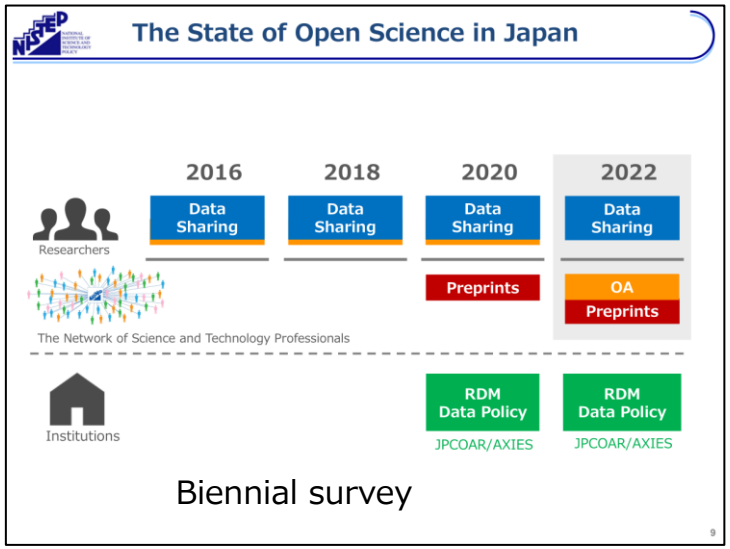
- RIKEN: Platform Coordination
- Univ of Tokyo: Integration and Utilization
- Nagoya Univ: Rules and guideline
- Osaka Univ: Human Resource Development

https://ccrd.nii.ac.jp/sc22/

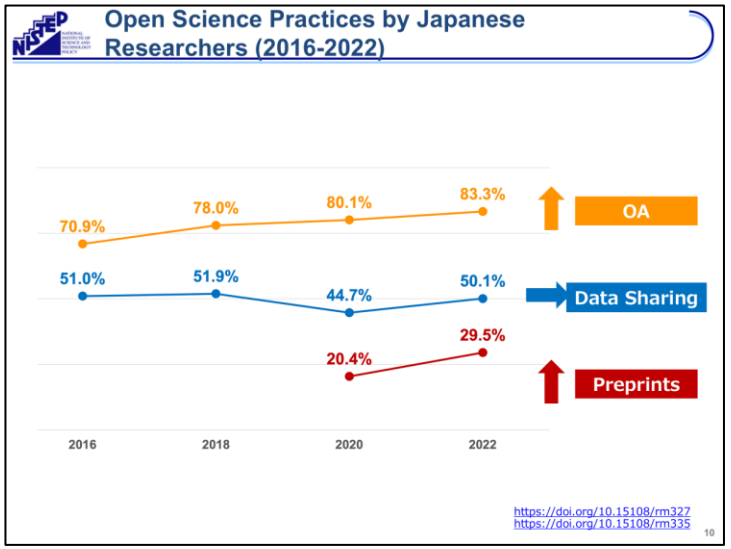
https://www.nii.ac.jp/creded/project.html

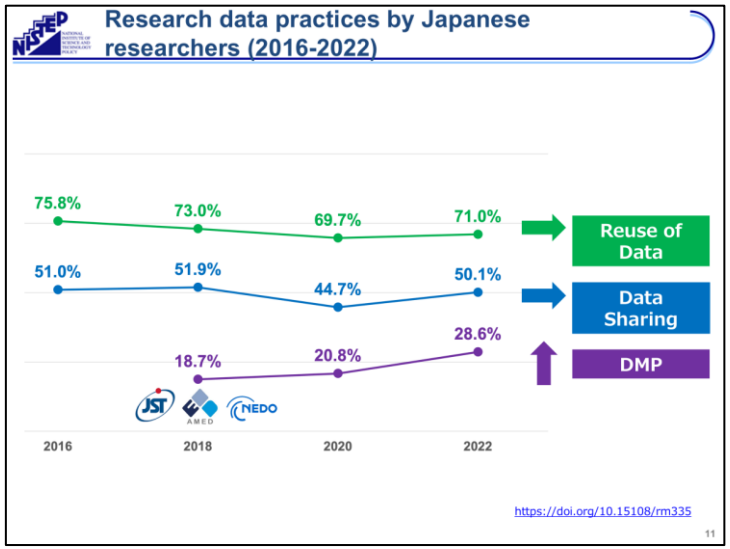
Monitoring Survey since 2016

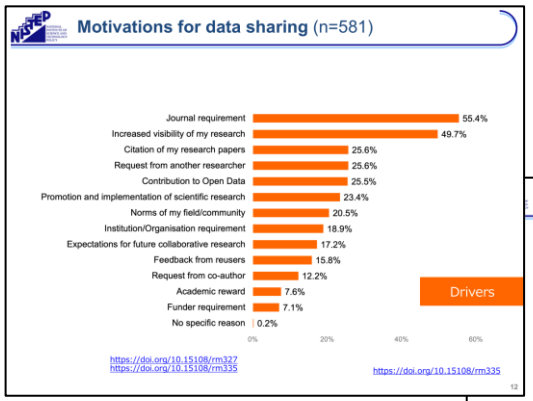
Survey for Semi-cohort of S&T experts (ca 2000) in NISTEP

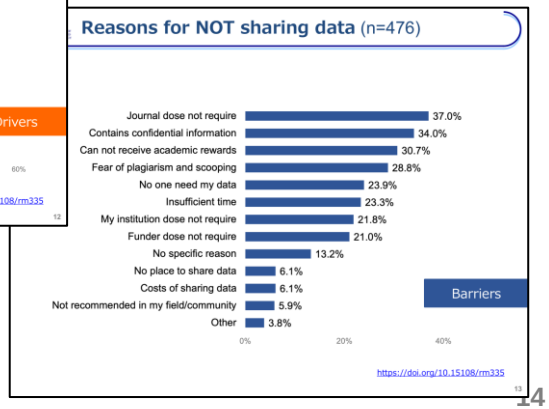


Biennial survey









Collaboration with and contribution by **Ui Ikeuchi**, NISTEP

2. Open Research Information System at NISTEP

Collaboration with and contribution by
Hitoshi Koshiba, NISTEP

262 人 工 知 能 31 巻 2 号（2016 年 3 月）

特 集「人工知能研究のベンチマークとは―標準問題・データセット・評価手法―」

オープンデータによるオープンサイエンスの推進

Open Data for Open Science

小柴 等 Hitoshi Koshiba — 文部科学省科学技術・学術政策研究所 National Institute of Science and Technology Policy（NISTEP）. hitoshi.koshiba@nistep.go.jp. http://researchmap.jp/hkoshiba/

林 和弘 Kazuhiro Hayashi（同 上） khayashi@nistep.go.jp. http://orcid.org/0000-0003-1996-4259

本稿では，オープンサイエンス，オープンデータの概要と動向を紹介し，人工知能研究への有用性と，人工知能研究がそれらにもたらすと思われるインパクトについて述べる．

社会と科学が距離を縮め，人工知能に対しても実用への期待と需要が高まる中で，科学界において提案される各種の技術や手法がビジネスや行政に関わる実際的な問題を解くことができるか，もしくは解くための支援ができるか，といったことも重要な評価指標の一つとなっている．

ここで機械学習を中心とした場合の人工知能では，技

眼から提案する．

以下では，オープンサイエンスやオープンデータそのものについて説明したうえで，関連するいくつかの事例について紹介しながら，これらの提案について説明する．

なお，意見・提言などに相当する部分は個人の見解である．

2. オープンサイエンスとオープンデータ

2·1 オープンサイエンス

Open Data for Open Science(<Special Issue>Benchmarks in AI Research: Standard Problems, Data Sets, Evaluation Methods)

Journal of the Japanese Society for Artificial Intelligence 2016
https://doi.org/10.11517/jjsai.31.2_262

Backgrounds and Issues

※ Digital Transformation

- **Progress in DX* and openness**
  - Progress in the release and utilization of open data on science and technology and research activities, including public funding, patents, and preprints.
  - By collecting these data and analyzing them together, it will be possible to quickly grasp the whole picture of the trends.

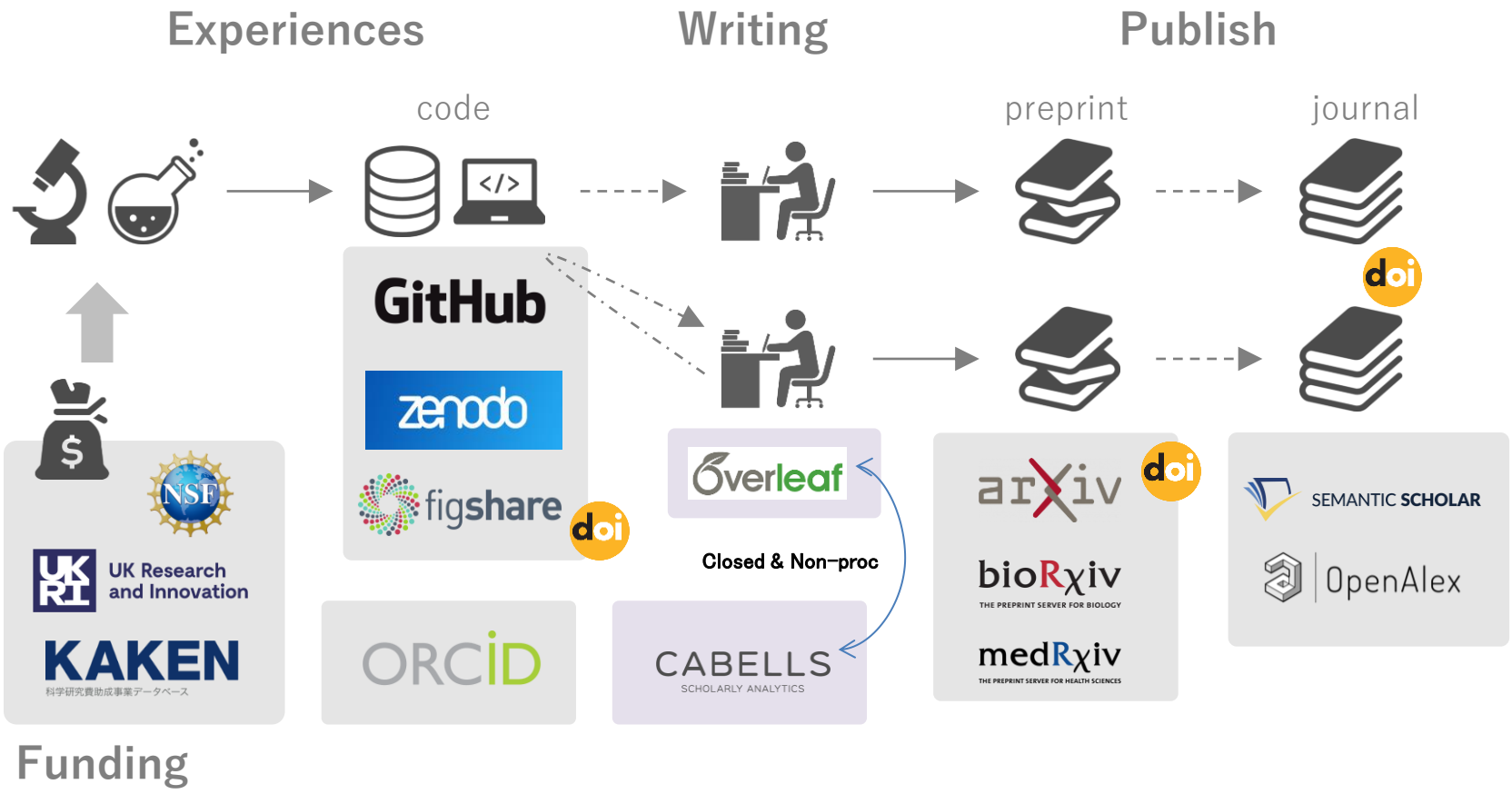
▼

- **Challenges: lack of uniform classification, lack of explicit citation relationships, etc.**

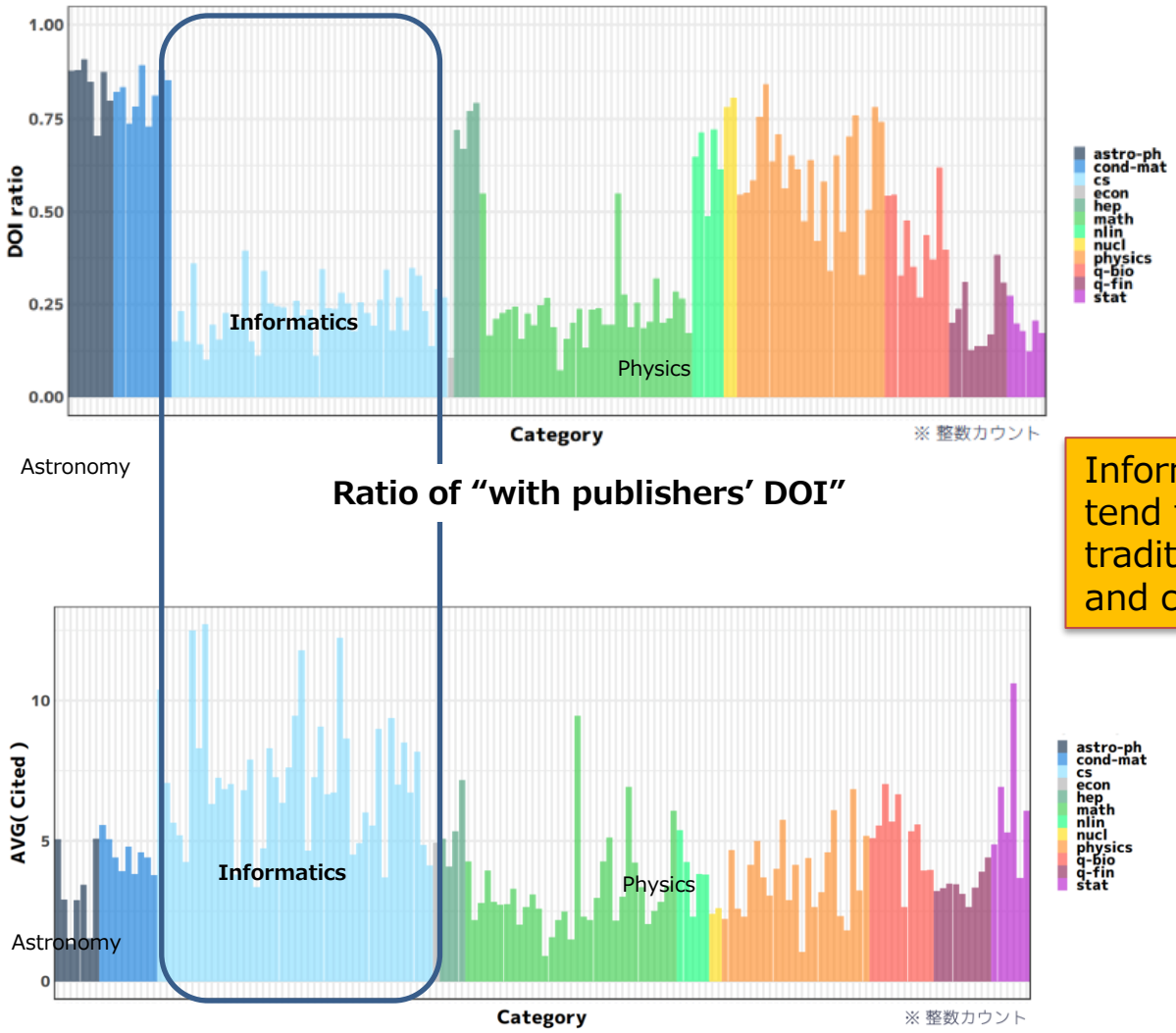
▼

- **Analysis using AI-related technologies such as machine learning and natural language processing**
  - Contribute to the promotion of DX and EBPM as a ministry or agency by systematizing and providing as much as possible.
    - Realize a new style of research activities that can be carried out together with administrative officials and other users.

Our Research Scope of Open Data



Ecosystem with preprints only (arXiv)



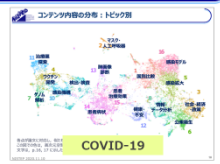
Informatics does not tend to publish in traditional journals and cites preprints

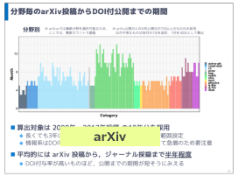
* NISTEP, Discussion Paper doi/10.15108/dp187

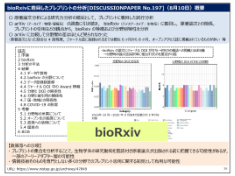
Other Examples

Analysis of preprint

- New information source preprints in addition to journal articles, patents, etc.
  - Since there is no commercial DB, it is necessary to work on collection and shaping individually.
  - There is a possibility that the power structure will change, and it will take a few more years to reach regular operation.

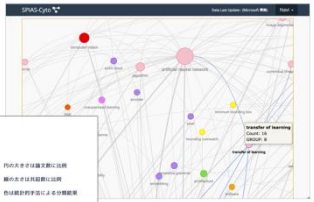


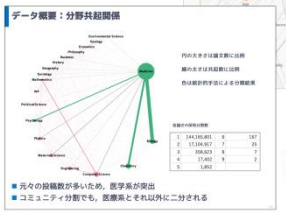




Example of open article DB analysis

Example of Microsoft Academic-based Topic Co-occurrence Analysis

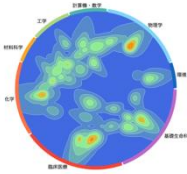




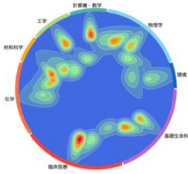
An example of Semantic Scholar-based disciplinary co-occurrence analysis

Example of field integration analysis

- Example of visualization of the relationship between the integration of disciplines and impact
  - This is an important point of view when we consider the relationship between humanity and society in the future.



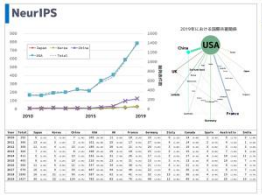
Japan 2012



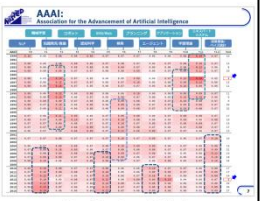
Japan 2016

Example of analysis of an international conference

- Collection and analysis of data from top information-related conferences, including NeuIPS, AAAI, and other AI-related conferences
  - RM253(2016) : http://doi.org/10.15108/rm253
    - Reported by the Committee on Information Science and Technology (94th, 95th meetings), etc.



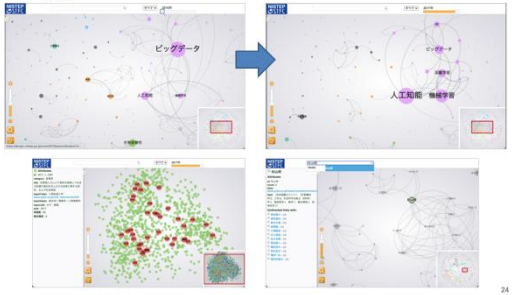
Number of posts by country



Trend of topics on LDA basis

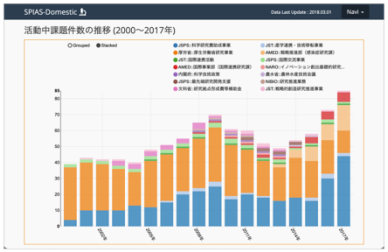
Example of survey-based trend analysis

- Utilized from 2015 to 2017 for descision making
  - Analyze and visualize changes in keyword co-occurrence relationships, etc.



Examples of analysis of various public competitive funds

- Understanding domestic trends of various funds from open data
  - Separate U.S. and European versions exist.



Changes in the number of issues related to "infection control" in general

These are “triggers” to communicate potential users (hopefully policy makers) to improve them AND to enhance data literacy of potential users

Summary

1. **Open Science in Japan**
   - ✓ Policy: 2021 Research Data Sharing, 2024 Open Access mandata from 2025
   - ✓ Data Platform: NII Research Data Cloud and its application program
   - ✓ Monitoring Surveys: to understand the actual status of open science

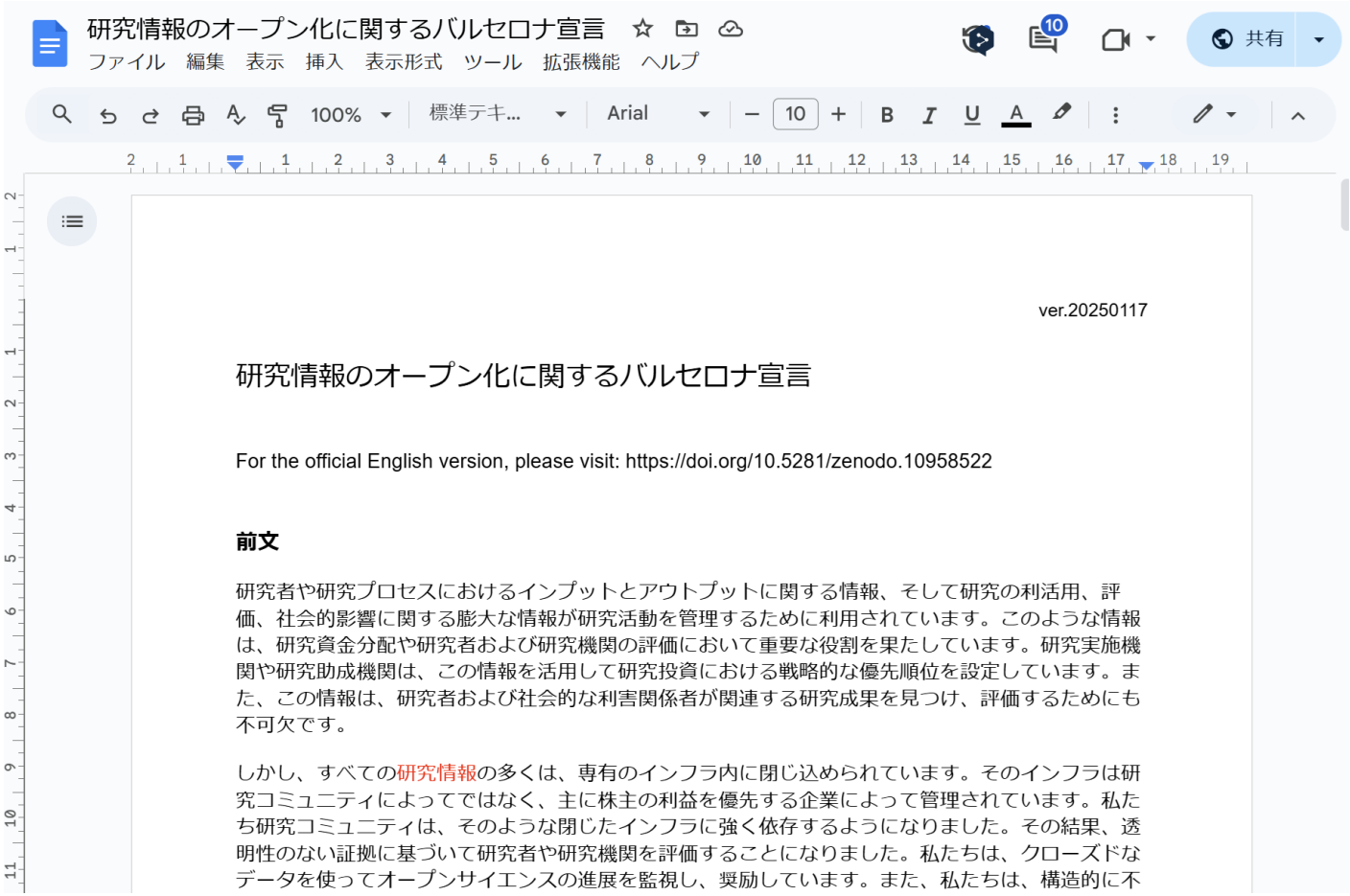
2. **Open Research Information System at NISTEP**
   - ✓ 10 years trials for utilizing Open Data for EBPM
   - ✓ Development of new survey research and data analysis methods using data science and AI-related technologies, etc.
   - ✓ Seeking an alternative methodology to the one with propritary database or even traditional one itself.

■ **Contribute to the promotion of EBPM through the development and deployment of analysis systems and the utilization of open data**

For promoting DX and Open Science

One more thing…

1. Japanese Translation of Barcelona Declaration Project
2. Volunteer Recruiting to check it



Contact us by Jan 31th!