



オープンサイエンス政策の 背景と現状 資源としての研究データ

林 和弘

文部科学省 科学技術・学術政策研究所
(NISTEP)

2017年12月27日(水)

内閣府

国際的動向を踏まえた

オープンサイエンスの推進に関する検討会 (第1回)



G7科学技術大臣会合 (つくば 2016.5→トリノ 2017.9)

- オープンサイエンスは、ICT等の活用によって知識がよりオープンになる（開放される）ことによって、研究活動の変容（Transform）を促すものである。
 - 知識生産活動のパラダイムシフトを見越し、標準化に向けた取組と先行者利益を獲得



G7 SCIENCE MINISTERS' COMMUNIQUÉ

Turin, 27 – 28 September

28th September 2017

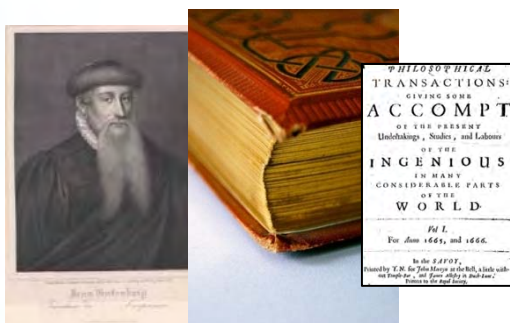
Working Groups under the G7 Science Ministers.

21. In recognizing the valuable contribution provided by the different Working Groups, we encourage cooperation and avoidance of duplication of efforts among the Working Groups themselves.
22. We reaffirm the importance of the actions agreed at the G7 Science and Technology Ministers' Meeting in Tsukuba in 2016 in support of the sustainable use of the seas and oceans and the achievement of United Nations Sustainable Development Goals 13 and 14. We recognise the key role the G7 Future of the Seas and Oceans Working Group can play in delivering these actions through developing stronger scientific knowledge and realizing a more efficient and

歴史に習う

• ポストグーテンベルグ時代の再フレーム

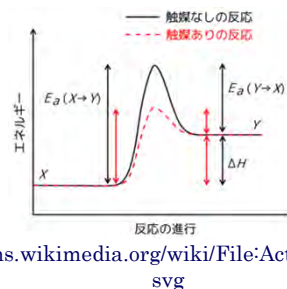
大量印刷ベース



Web ベース

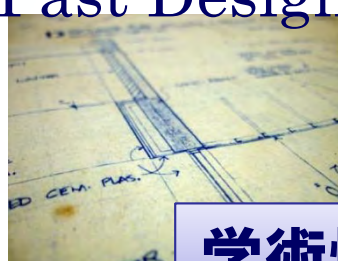
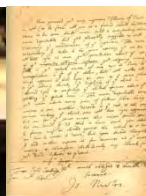


遷移状態

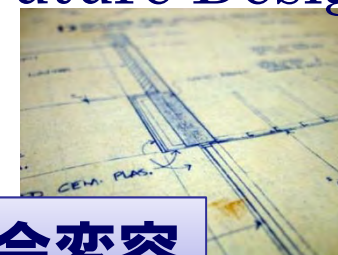


Letter based dissemination

Past Design



Future Design



学術情報流通の再発明を含む社会変容

オープンサイエンスの潜在的便益 (政策的観点)

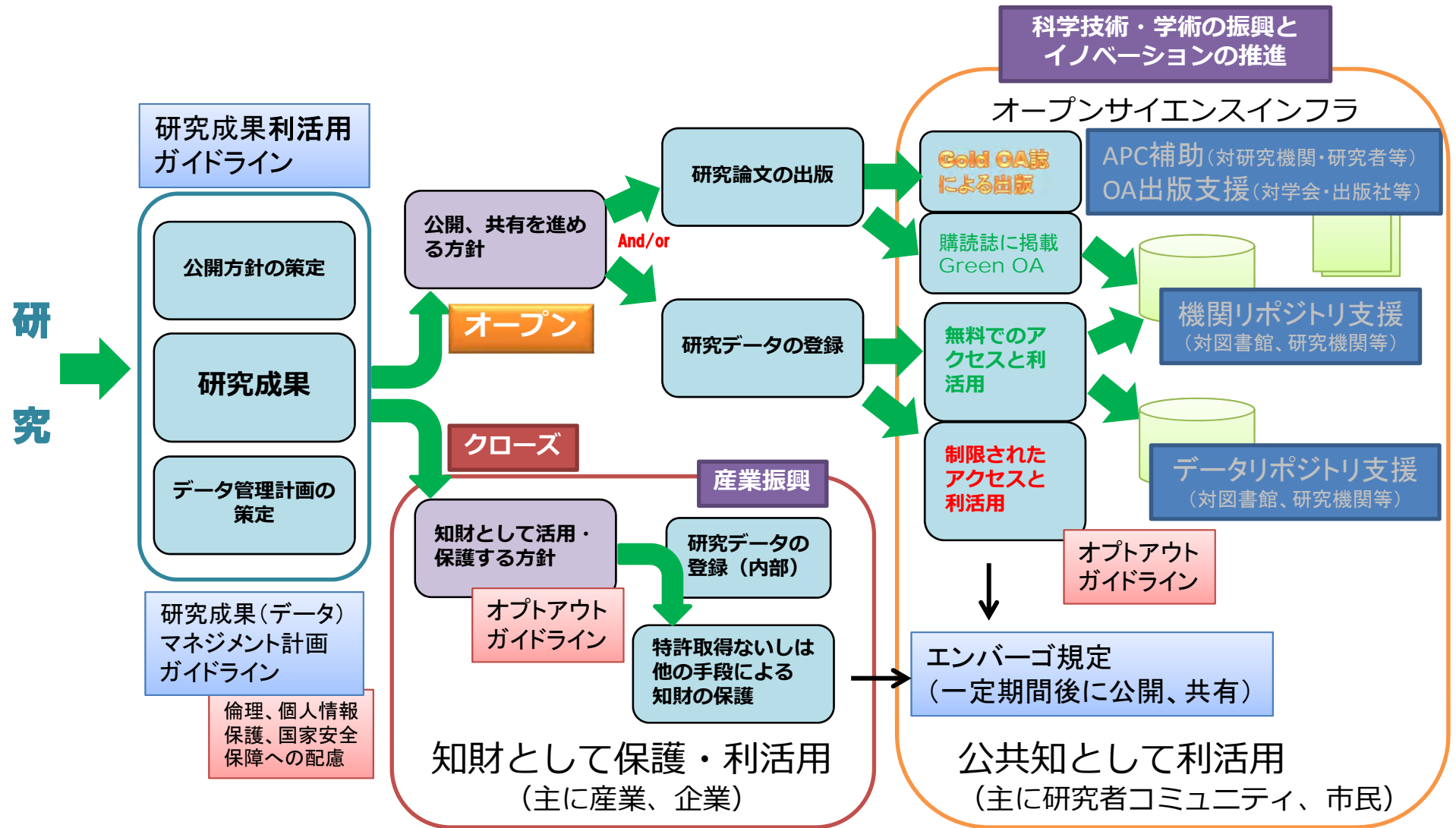
- 研究を加速し成果を見つけやすくすることで 研究開発投資の費用対効果を上げる
- 同じ研究を繰り返すこと避け、研究開発コストを抑える
- 境界領域や多領域にまたがる研究の機会を増やし、多分野の協調を促す
- 研究結果の商業化を早く広い観点から行い、公共研究開発投資の効果を上げ、科学情報を基にした新しい産業を生み出す



Fact sheet: Open Access in Horizon 2020

https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/FactSheet_Open_Access.pdf

研究成果の利活用、オープンサイエンスの推進に係る概念図



下記図表を参考に和訳、改変

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

オープンサイエンス政策、方針の俯瞰

Top
Down

G7科学技術大臣会合
EU (DSM) , OECD (Going Digital)

第5期科学技術基本計画 (Society 5.0)

内閣府オープンサイエンス検討会

文部科学省、他府省庁

研究助成団体

研究機関、大学等研究実施組織

Bottom
Up

草の根活動

欧州単一市場 (Digital Single Market)

- 2015年5月Digital Single Market Strategy (デジタル単一市場戦略 ; DSM)
 - デジタル技術に基づく情報利用・サービス、ネットワークや経済の向上を実現 (5億人、50兆円)
 - データ・情報通信の標準化および相互運用性 (interoperability) の確保が優先事項
 - EUの試算によればEuropean Open Science Cloud(EOSC)構築に67億ユーロ、うち20億ユーロはホライゾン2020予算、残り47億ユーロは他の公的・民間資金を併用して投資するとしている

村山、林、欧州オープンサイエンスクラウドに見る
オープンサイエンス及び研究データ基盤政策の展望

STI Horizon, Vol. 2, No. 3, p. 49-54.

<http://doi.org/10.15108/stih.00044>



欧州サイエンスクラウド (EOSC)

- 既存の研究データ基盤構築の施策との連携調整：EUDAT(HPCI)、GÉANT[NW]、LIBER[Library]、OpenAIRE[Repository]、EGI[Data Center]等
 - 分野的・地理的・施策上別箇に整備されたシステムを結合
 - 欧州全体の研究データ利活用基盤→世界的な共通基盤 (“Global Open Science Cloud”) へ？
 - NIIのOpenAIRE連携

村山、林、欧州オープンサイエンスクラウドに見る
オープンサイエンス及び研究データ基盤政策の展望

STI Horizon, Vol. 2, No. 3, p. 49-54.

<http://doi.org/10.15108/stih.00044>



先導するオーストラリア



Data is Tran

- Government make life ea
- Investments problems to
- This requires wide variety



NCRIS: Australian Infrastructure Approach

- **Stable** for 10 years – bilateral support
- \$AU150M/annum
- Invests in **collaborative** infrastructure
- Both physical and data
- **Data is infrastructure**
- Separate from research funding
- Substantial national data assets created
- \$20M/annum on data and collaboration services (inc. ANDS 10M, RDS, 5M, Nectar 5M)

9

The Value and Opportunities for Sharing Research Data – an AU perspective, Ross Wilkinson, Australian National Data Service (NISTEP, Tokyo, 2017)

第5期科学技術基本計画(2016-2020)

- “「国際的動向を踏まえたオープンサイエンスに関する

第4章「科学技術イノベーションの基盤的な力の強化」

・今後起こり得る様々な変化に対して柔軟かつ的確に対応するため、若手人材の育成・活躍促進と大学の改革・機能強化を中心に、基盤的な力の抜本的強化に向けた取組を進める。

(2) 「知の基盤の強化」③「オープンサイエンスの推進」

・オープンサイエンスの推進体制を構築し、公的資金による研究成果については、その利活用を可能な限り拡大することを、我が国のオープンサイエンス推進の基本姿勢とする。

第4章2節3項「オープンサイエンスの推進」

http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu22/siryu/___icsFiles/afiedfile/2016/12/08/1380241_04.pdf

資料3
科学技術・学術審議会
総合政策特別委員会
(第15回) H28.11.24

オープンサイエンスの推進について

文部科学省

研究振興局参事官(情報担当)付

科学技術・学術政策局科学技術・学術戦略官
(制度改革・調査担当)付



文部科学省

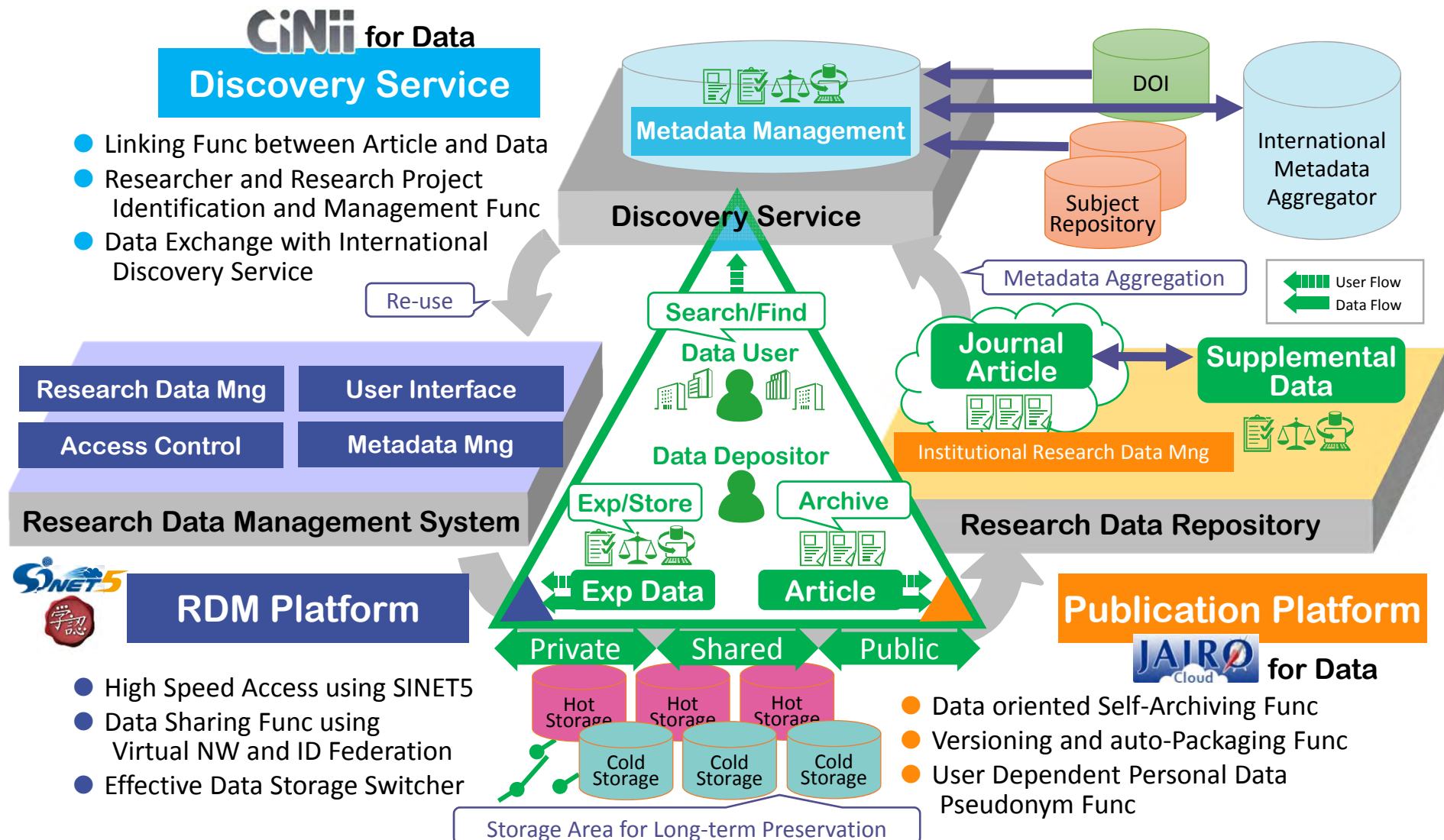
MEXT

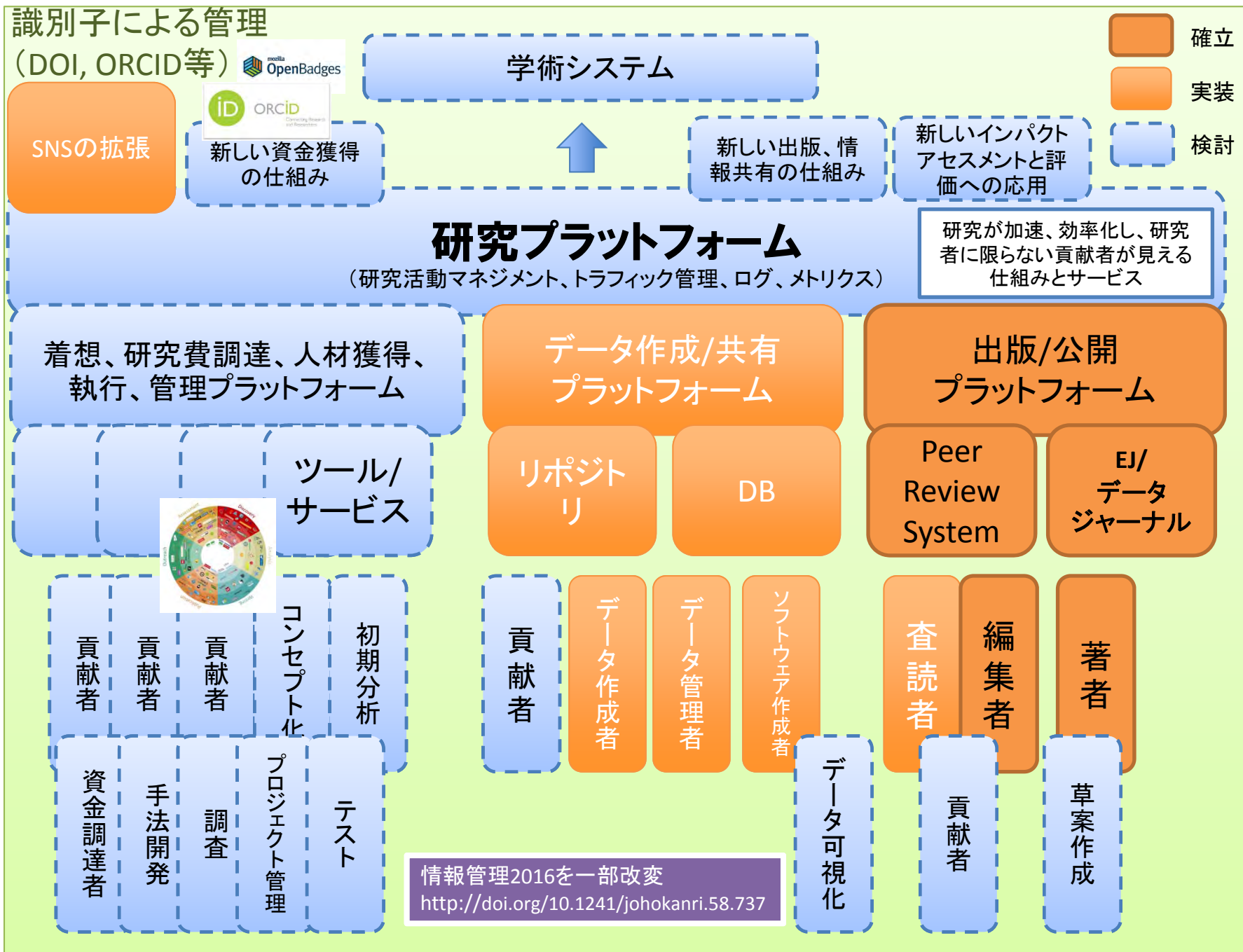
MINISTRY OF EDUCATION,
CULTURE, SPORTS,
SCIENCE AND TECHNOLOGY-JAPAN

文部科学省としての基本方針は示されている

- 内閣府の方針の再確認（論文とエビデンスデータから）
- 研究成果の利活用促進を前提としたオープン化
 - すでにオープンサイエンスの文脈を含んだ施策、活動の整理
 - 既存の施策とのひも付けとマッピング
- オープンサイエンスの文脈で拡充、改善
 - 研究データ基盤の整備など新しい試みもある
- 人材等不透明な課題も残ってはいる

Research Data Infrastructure for Open Science

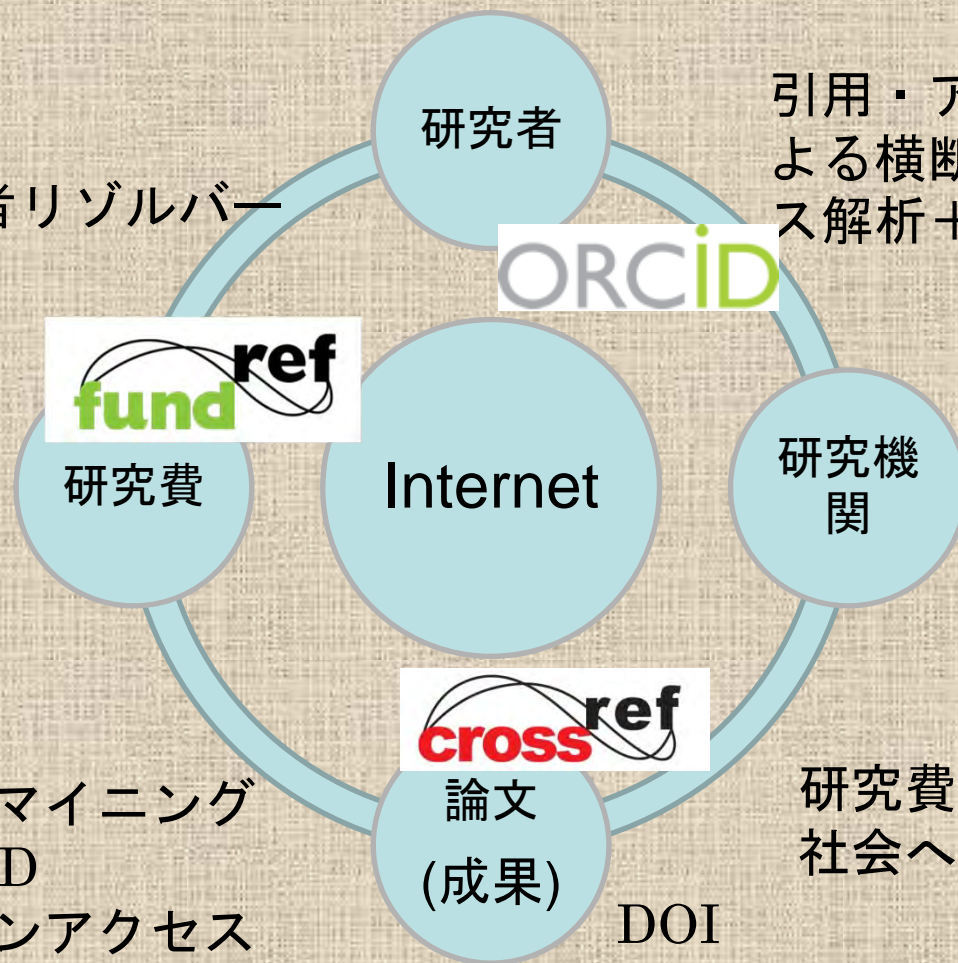




ID (識別子) の浸透により、 どの研究機関の誰がどの研究費を使ってどんな研究をし、 その成果とインパクトはどうだったかがわかる時代へ

ORCID
E-Rad
KAKEN-研究者リゾルバー
READ

引用・アクセス数などによる横断的パフォーマンス解析 + Altmetrics



データマイニング
Open ID
オープンアクセス

研究費の透明性の確保
社会への説明



論文誌の電子ジャーナルをめぐる最近の動き, 科学技術動向, 2009/7, 100, 10-18.
(改変)



安心・安全なデータ共有文化作り

- 簡便な蓄積
 - 研究データ基盤整備（FAIR原則に準拠）
- 簡便な利用とユースケース作り
 - サービスデザイン（これがないと研究ができない）
 - 利用に関するライセンス(例えオープンでも)
 - マシンリーダブル
- 健全な評価と報酬（Reward）
 - 当面はデータ引用文化の醸成と評価への展開
- 文化醸成のためのステークホルダーの対話の繰り返し
- 主体的に取り組むことが重要（ルールメイク）

FAIRデータ文化と信頼の確保



EUROPEAN COMMISSION
DIRECTORATE-GENERAL FOR RESEARCH & INNOVATION

The Director-General

Brussels, 26 October 2017

EOSC Declaration

RECOGNISING the challenges of data driven research in pursuing excellent science;

GRANTING that the vision of European Open Science is that of a research data commons, widely inclusive of all disciplines and Member States, sustainable in the long-term,

CONFIRMING that the implementation of the EOSC is a process, not a project, by its nature iterative and based on constant learning and mutual alignment;

UPHOLDING that the EOSC Summit marked the beginning and not the end of this process, one based on continuous engagement with scientific stakeholders, the European Commission,

PROPOSES that all EOSC stakeholders consider sharing the following intents and will actively support their implementation in the respective capacities:

Data culture and FAIR data

- **[Data culture]** European science must be grounded in a common culture of data stewardship, so that research data is recognised as a significant output of research and is appropriately curated throughout and after the period conducting the research. Only a considerable cultural change will enable long-term reuse for science and for innovation of data created by research activities: no disciplines, institutions or countries must be left behind.
- **[Open access by-default]** All researchers in Europe must enjoy access to an open-by-default, efficient and cross-disciplinary research data environment supported by FAIR data principles. Open access must be the default setting for all results of publicly funded research in Europe, allowing for proportionate limitations only in duly justified cases of personal data protection, confidentiality, IPR concerns, national security or similar (e.g. 'as open as possible and as closed as necessary').



The screenshot shows a news article on the Jisc website. The article title is "Concordat on open research data launched" and the date is "28 July 2016". The text of the article states: "Four of the UK's leading research organisations - Higher Education Funding Council for England (HEFCE), Research Councils UK (RCUK), Universities UK (UUK) and Wellcome Trust - have today launched a concordat that proposes a series of clear and practical principles for working with research data." Below the text, there is a list of participating organizations: "The Concordat on Open Research Data has been developed by a UK multi-stakeholder group - Arts and Humanities Research Council (AHRC), Natural Environment Research Council (NERC), RCUK, Newcastle University, UUK, HEFCE, University of Warwick, Research Information Network, Springer Nature, British Library, Wellcome Trust, University of Essex, The Russell Group and Jisc - and is a set of expectations of best practice".

EOSC Declaration

RECOGNISING the challenges of data driven research in pursuing excellent science;

GRANTING that the vision of European Open Science is that of a research data commons, widely inclusive of all disciplines and Member States, sustainable in the long-term,

CONFIRMING that the implementation of the EOSC is a process, not a project, by its nature iterative and based on constant learning and mutual alignment;

UPHOLDING that the EOSC Summit marked the beginning and not the end of this process, one based on continuous engagement with scientific stakeholders, the European Commission,

PROPOSES that all EOSC stakeholders consider sharing the following intents and will actively support their implementation in the respective capacities:

Data culture and FAIR data

- [Data culture] that research data throughout and enable long-term disciplines, insti
- [Open access b efficient and cr Open access m allowing for pr confidentiality,



人文学オープンデータ共同利用センター
Center for Open Data in the Humanities

日本語

<http://codh.rois.ac.jp/seminar/coretrustseal-20171204>

★ / CODHセミナー / 第5回CODHセミナー

第5回CODHセミナー

信頼できるデータリポジトリ

～CoreTrustSeal認証に関する実践的情報共有の場～



CoreTrustSeal



ICSU
WORLD DATA SYSTEM
ICSU World Data System (WDS)

研究データ利活用協議会 Research Data Utilization Forum (RDUF)

Organizational Members



- 多様なステークホルダー
RDUF
Research Data Utilization Forum

Toppage

Topics

Japanese



Landscape of Open Science/Research Data Sharing



2008-
ICSU-World Data System



International Council
for Science



Future Earth
(ICSU, UNESCO, UNEP,
UNU, Belmont Forum,...)



1966-
CODATA (Committee on Data
for Science & Technology, ICSU)



European Open Science Cloud
ESFRI, EUDAT, GEANT. . .



2012-
Research Data Alliance



OECD:
Projects of Intl. e-Infra./
interoperable policy



Group on Earth Observation
Global Earth Observation
System of Systems



The Internet
of Things



Y. Murayama

Landscape of Open Science/Research Data Sharing



2008-
ICSU-World Data System



International Council
for Science

Future Earth

(ICSU, UNESCO, UNEP,
,...)

国際ネットワークと、国際的に不利にならない立ち回り

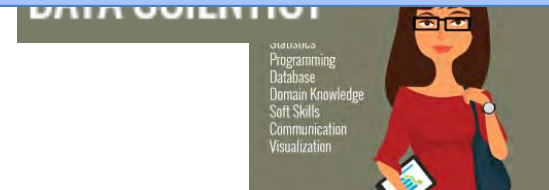
(理想的には)

- 国際的な取り組みの中での日本の明確な貢献とプレゼンスの向上を図る。

(より身近なところでは)

- 主要学術ジャーナルにおけるデータポリシーに対する対応 (根拠データの提出と公開)
- 今後の国際共同研究における研究データの取り扱いへの対応

The Internet
of Things



Y. Murayama

Data sharing、Open dataの意義の例

- ① 共同解析・・・大量の実験データをグループでは解析しきれない
 - ✓ LHC-ATLAS実験：約1000回／秒の陽子・反陽子衝突の即時解析のためには100万台のCPUが必要 ⇒ 170研究機関を通信ソフトで結び共同解析 = ヒッグス粒子を数十事象／110億衝突事象 を発見
- ② 共同観測・・・観測する対象が大きすぎて1機関では扱えない
 - ✓ GEOSS(全球地球観測システム)：世界全域を対象とした人工衛星や地上観測など多様な観測システムが連携した、包括的なシステムを構築
- ③ ユニークデータの共有・・・何度も同じ解析をしない
 - ✓ ヒトゲノム計画：ヒトのゲノムの全塩基配列(約30億塩基対)を解析するプロジェクトの実施
- ④ 失敗データの再利用・・・ダメも多く集まれば、成功が見える
 - ✓ マテリアル・インフォマティクス：調べられた化合物データの解析により新たな物質を探索 物性を予測して合成が可能な時代へ
- ⑤ ボランティアの活躍・・・趣味が役に立つ
 - ✓ Galaxy Zoo：ハッブル望遠鏡がとらえた何百万の銀河系の画像を20万人以上の市民ボランティアが分類・整理 新たな銀河発見にも繋がる

研究者の意識

■ ある ■ ない ■ わからない ■ データは用いない ■ OAの論文がある ■ OAの論文はない ■ わからない



図 4 公開データとOA論文の有無 (いずれもn=1,398)

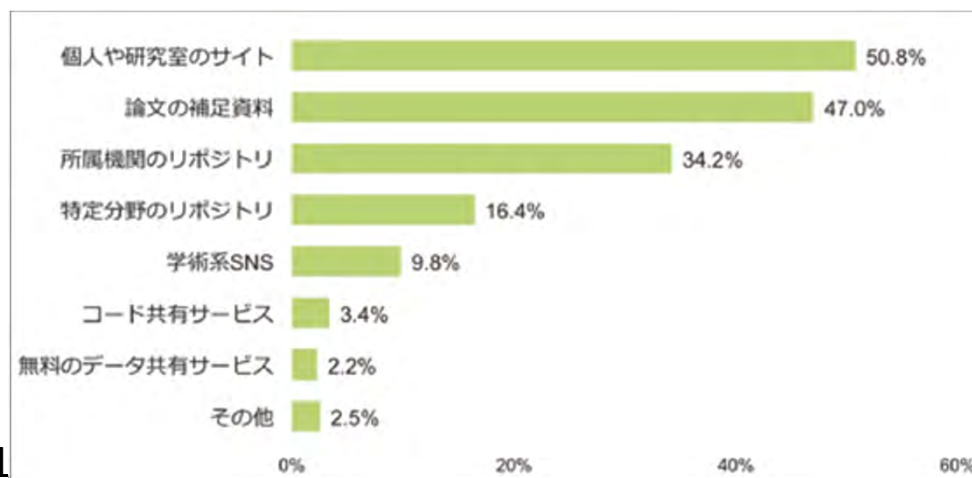
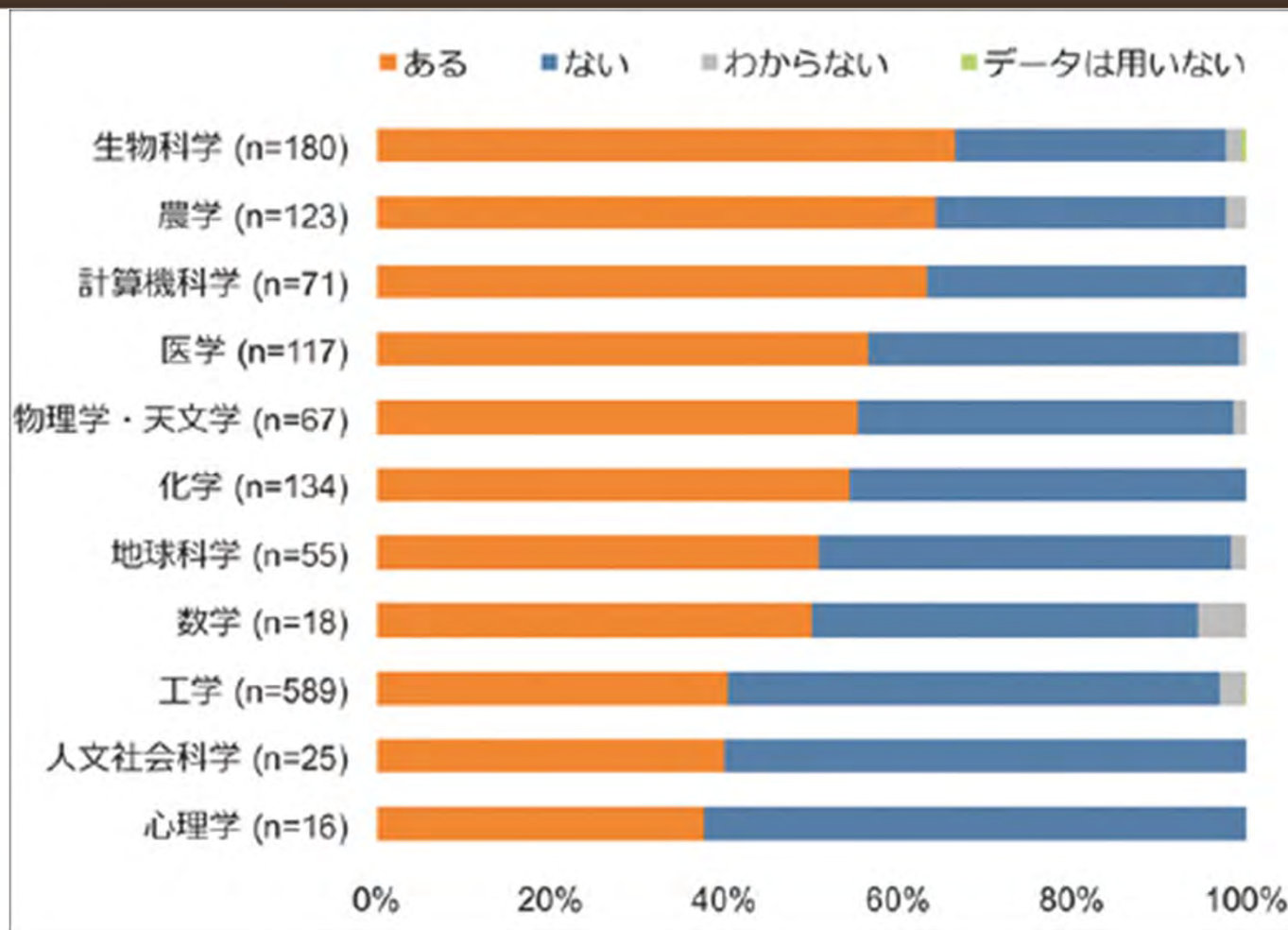


図 11 の公開意思 (データn=595, 論文n=379)

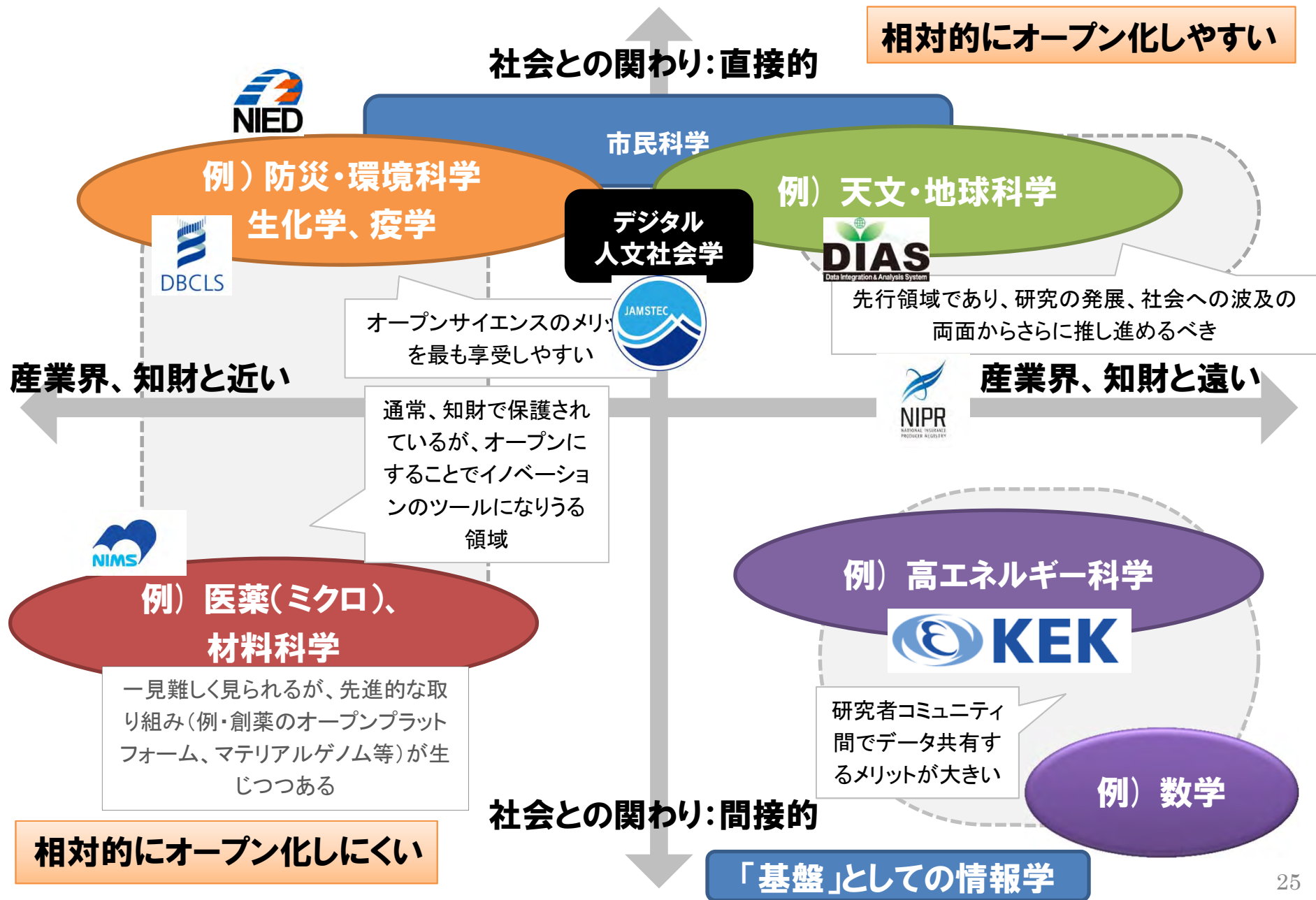
池内有為, 林和弘, 赤池伸一 (2017). 研究データ公開と論文のオープンアクセスに関する実態調査. 文部科学省科学技術・学術政策研究所科学技術予測センター. <http://doi.org/10.15108/stih.00106>

分野特性



池内有為, 林和弘, 赤池伸一 (2017). 研究データ公開と論文のオープンアクセスに関する実態調査. 文部科学省科学技術・学術政策研究所科学技術予測センター. <http://doi.org/10.15108/stih.00106>

オープンサイエンスに係る分野別相対マッピング例





日本としてのオープンサイエンス

- 国外の動向としては、国の政策レベルでは、研究データの利活用についての戦略が（研究分野別対応を含めて）相当進んでいる国（EU含む）が見られる。
（個々の研究者レベルの理解が進んでいるかについては注意が必要）
- 我が国においても、取組の進展が見られる中、国の一部の機関で始められている取組を横展開・全体化し、我が国の研究者及び研究コミュニティが国際的に不利な状況にならないようにする。
- 日本発の研究データを“資源”として捉え、その利活用戦略（データ（マネジメント）ポリシー）をもって、日本の研究力向上と産業育成を目指し、デジタルネイティブなイノベーション基盤を構築する。