

我が国のオープンサイエンス動向と データ共有の今後について

村山泰啓

NICT

目次

- オープン・データ
 - 国際的な動向、SGEPSS将来構想
 - 国内動向へ
- 内閣府・CSTI
 - データ共有を国としてすすめるには
- 国立国会図書館
 - 図書館は「無料の貸本屋」か？ 予算の減額等。
 - 電子データは図書館の永続コンテンツになるか

G8 (2013) における 研究データオープン化の 合意

(↔ Open Government Data)

G8 Science Ministers Statement London UK, 12

Introduction

We, the G8 Science Ministers met in London on Wednesday of our respective national science academies, as part of this unique meeting we discussed how our nations could lead in transparency, coherence and coordination of the global scientific in order to address global challenges and maximise the social 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



SGEPSS将来構想セッションでの講演から

(村山ほか、SGEPSS、2012年10月)

SGEPSS将来構想セッション： データシステムの新たな取り組み

村山泰啓、堀 智昭、篠原 育
SGEPSS将来構想検討ワーキンググループ
山本 衛

過去のSGEPSS将来構想における「データ」

- 平成3年「地球電磁気学の発展的将来」、全67ページ。
(全67ページ;日本学術会議地球電磁気研究連絡委員会)
- データネットワークのセクションを設けて、提言。
- 平成17年「21世紀の地球電磁気学」全14ページ。(日本学術会議地球電磁気研究連絡委員会)
- データのセクションはないが、DB重要性の記述が各分野に埋め込まれていた。

第5章 研究推進のために必要な施策 ……………

1 宇宙空間科学および中層・超高层大気科学 ……………

1. 1 惑星研究の推進 ……………

1. 2 国際協力の推進 ……………

1. 3 海外拠点観測所 ……………

1. 4 航空機観測体制の確立 ……………

1. 5 理論・シミュレーション ……………

1. 6 データネットワーク・総合解析 ……………

1. 7 研究・教育体制 ……………

(地球電磁気学の発展的将来、1991)

SGEPSS将来計画に書き込みました。

データパブリケーション

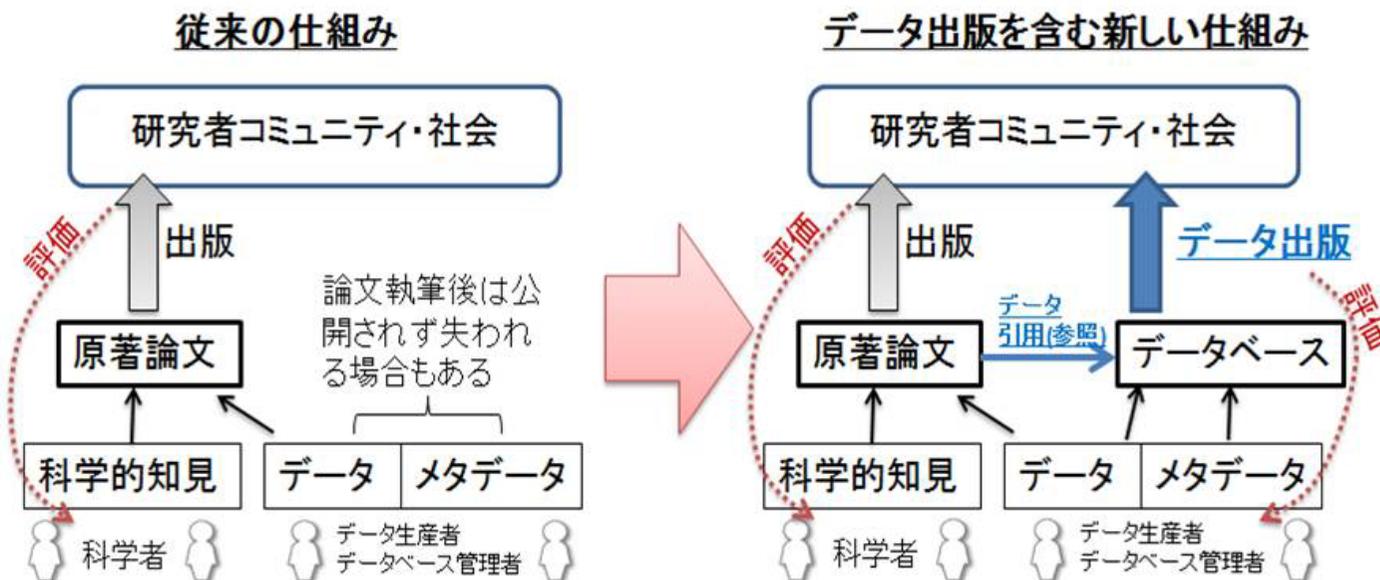
- データを「出版」する仕組み:
- 課題: データの「査読」「固定」「公表」等をどうするか。
- 課題: ID標準化、引用ルール確立、評価手法など国際団体等で模索中

データサイテーション

- データを文献のように「引用」「参照」する仕組み
- 課題: ID標準化、引用ルール確立、評価手法など国際団体等で模索中

データを出版・引用・参照すると

- 論文・書籍と同様、知的生産力の基準に。→ 研究職・教育職の業績評価。
- 信頼できるデータ生成・提供は現代では科学者の仕事。← 評価



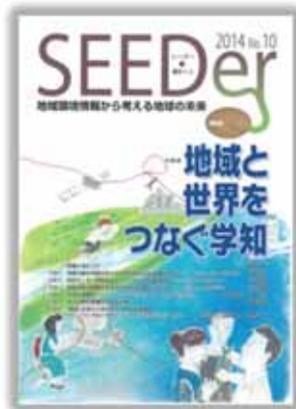
[地球電磁気・地球惑星圏学会, 2013]

オープンサイエンス系のコンセプトの国内周知

「地球電磁気学・地球惑星圏科学の現状と将来」(2013)とその図面、URLをリファーしています。
(doiをつけてほしい?)

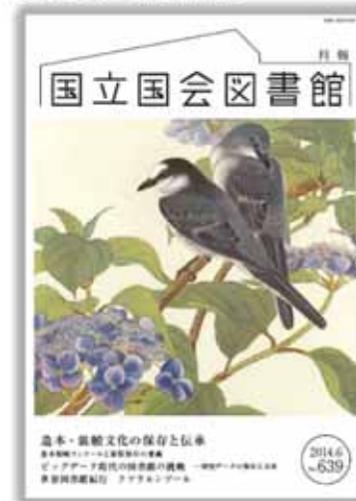


学術の動向、2013年9月
(日本学術会議)



SEEDer (シーダー) 10号(2014年4月)
特集：地域と世界をつなぐ学知
(総合地球学研究所・昭和堂)

国立国会図書館月報
639号 2014.6



文科省「科学技術動向」誌、
連載(2014年9・10月号、
11・12月号...)



科学技術オープンデータの背景

- 科学(技術)と社会

- 社会と科学者の関わりが問われている

- データ

- 科学技術活動の重要な成果物
 - 公的資金研究によるデータの公開原則
 - さらなる研究の加速 (e-science, data intensive science...)
-

- オープンサイエンスデータの検討

- World Data System事業 (International Council of Science)等
- G8+O6 data infrastructure WG
国際コンソーシアム「RDA (Research Data Alliance)」
- 内閣府・CSTI「国際動向を踏まえたオープンサイエンスに関する検討会」
- 国立国会図書館「科学技術情報整備審議会」/ 第四期科学技術情報整備基本計画策定に向けた基本方針検討部会

内閣府・CSTIの「オープンデータ議論」

オープン化に関する諸外国及び我が国の動向

平成26年12月9日

内閣府

第4期科学技術基本計画における推進方策

4. 国際水準の研究環境及び基盤の形成

(3) 研究情報基盤の整備

【推進方策】

- 国は、大学や公的研究機関における機関リポジトリ*の構築を推進し、論文、観測、実験データ等の教育研究成果の電子化による体系的収集、保存やオープンアクセスを促進する。また、学協会が刊行する論文誌の電子化、国立国会図書館や大学図書館が保有する人文社会科学も含めた文献、資料の電子化及びオープンアクセスを推進する。

*論文等のデータを機関毎に保存・公開する電子アーカイブシステム

- 国は、デジタル情報資源のネットワーク化、データの標準化、コンテンツの所在を示す基本的な情報整備、更に情報を関連付ける機能の強化を進め、領域横断的な統合検索、構造化、知識抽出の自動化を推進する。また、研究情報全体を統合して検索、抽出することが可能な「知識インフラ」としてのシステムを構築し、展開する。

内閣府/CSTIにおける研究データ共有のガイドライン

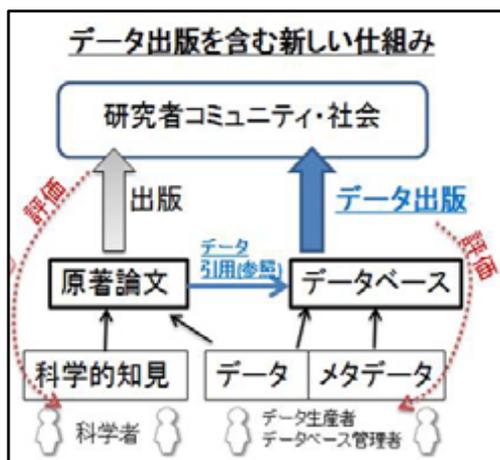
国際的動向を踏まえたオープンサイエンスに関する検討会の開催について

平成 26 年 11 月 13 日

内閣府 政策統括官（科学技術・イノベーション担当）決定

1. 趣旨

オープンサイエンスに係る世界的議論の動向を的確に把握した上で、我が国としての基本姿勢を明らかにするとともに、早急に講ずべき施策及び中長期的観点から講ずべき施策等を検討するため、「国際的動向を踏まえたオープンサイエンスに関する検討会」（以下「検討会」という。）を開催する。



内閣府の会議資料にも入れました。

[内閣府、2014、
<http://www8.cao.go.jp/cstp/sonota/openscience/1kai/1kai.html>、2015年2月20日取得]

I. 国際的動向からみたオープン化に関する現状認識

1. 認識すべき重要性
2. 基本概念
 - (1) オープンサイエンス
 - ① オープンアクセス
 - ② オープンデータ
 - ③ オープン化すべき科学研究データ
3. 国際的動向にみるオープン化の必要性

II. 国際的動向にみるオープン化に関する課題と検討すべき方向性

1. 国際的動向にみる我が国の現状
2. 日本におけるステークホルダーに求められる役割と課題

内閣府・CSTIの「オープンデータ議論」

国際的動向からみたオープンサイエンスに関する検討資料

平成 26 年 12 月 9 日

③ オープン化すべき科学研究データ

科学研究データは原著論文と同等又はそれ以上の研究開発成果として位置づけられる。つまり、オープンサイエンスの概念のもと、オープンすべきデータは数値データ、テキストレコード、イメージ、ビジュアルデータ、メタデータなど多様なデータが含まれる。ただし、諸外国におけるデータ共有・管理ポリシーにおいては、機密、法律、知的財産権等で制限事項が設けられているほか、国益を損なうことについてのオープン化は認められていない。

1. 国際的動向にみる我が国の現状

国際的なオープン化の潮流は、留まることを知らず論文の次は、研究データのオープン化が加速的に議論・推進されている。日本はこうした潮流に乗り遅れる可能性がある。

(1) 潮流に乗り遅れるデメリット

① 地球規模研究におけるデメリット

共有すべき研究データの共有が進まないことで、地球規模の研究に参加できず、国際的貢献に寄与できないとともに、競争力低下を招く。

② 研究活動における効率化、国際発信におけるデメリット

オープン化のデファクト・スタンダードが海外主導によって定められた場合、日本の研究活動及び国際発信に不利益が生じる。

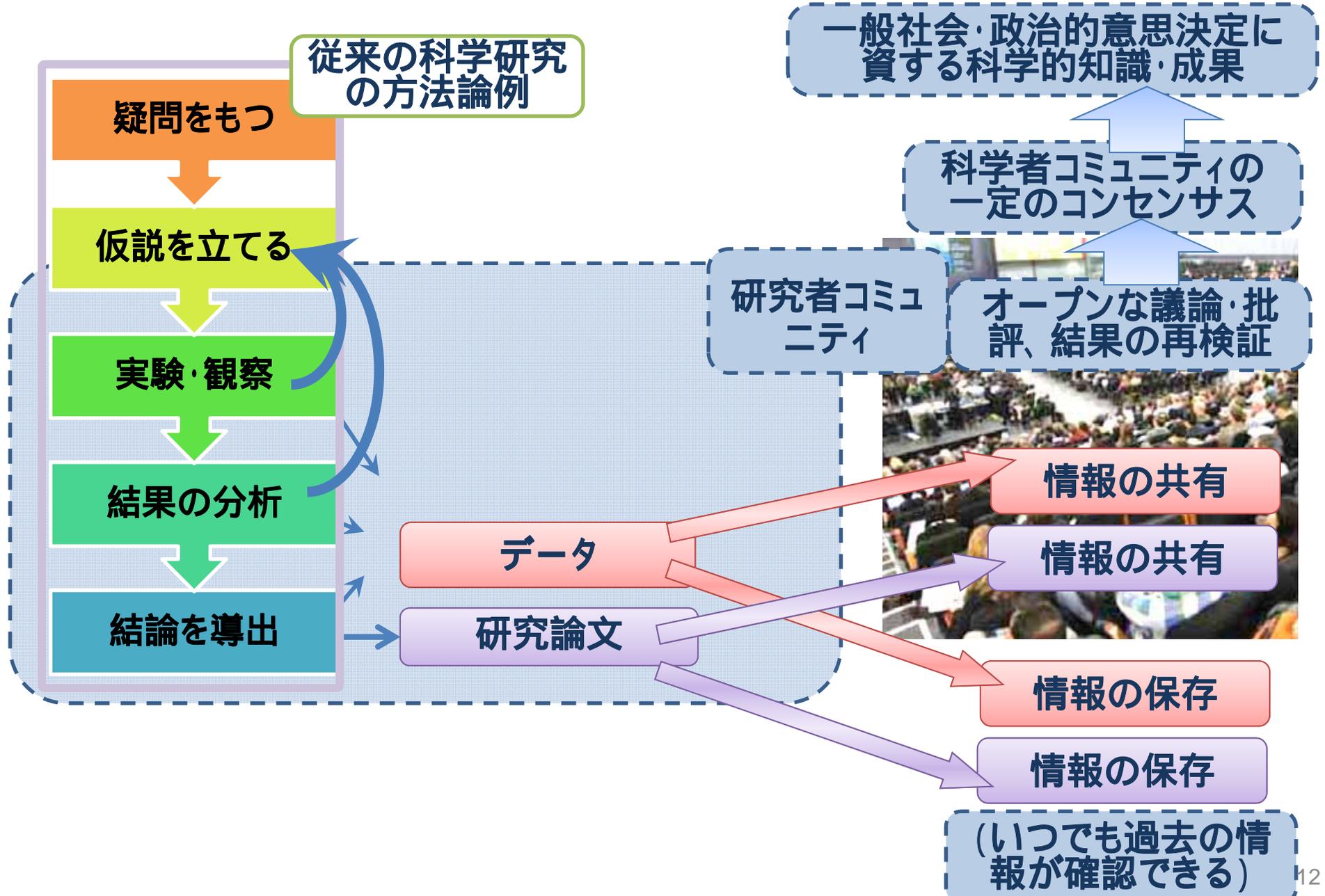
③ 日本の「見えない化」

世界でオープン化が進み、研究論文、データの相互交流が活発になると、日本で行われた閉じた研究が相対的に見えなく恐れがある。その結果、日本の研究者の業績が正当化されず、海外での研究や共同研究の機会を失う恐れがある。

オープンサイエンス、データ共有の 体制にむけたアイデア

データからみた研究過程	実施主体の検討
データ生成	研究現場。研究機関、大学
データ処理	同上。研究者が必要。
データ利用、研究、論文	同上
データ保存(永続保存)	<p>永続保存が可能なのは？</p> <ul style="list-style-type: none"> × 研究機関は時限活動 図書館 Libraryは永続保存! <p>→ 課題: 文献・書籍業務から踏み出せるか？</p>
メタデータ(書誌情報のようなもの)整備	<p>研究者: 苦手(な人が多いのでは?)</p> <p>図書館司書: 書誌管理のプロ</p> <p>→ 両者の相互協力は有用ではないか。</p>
...	

科学研究とデータ・情報



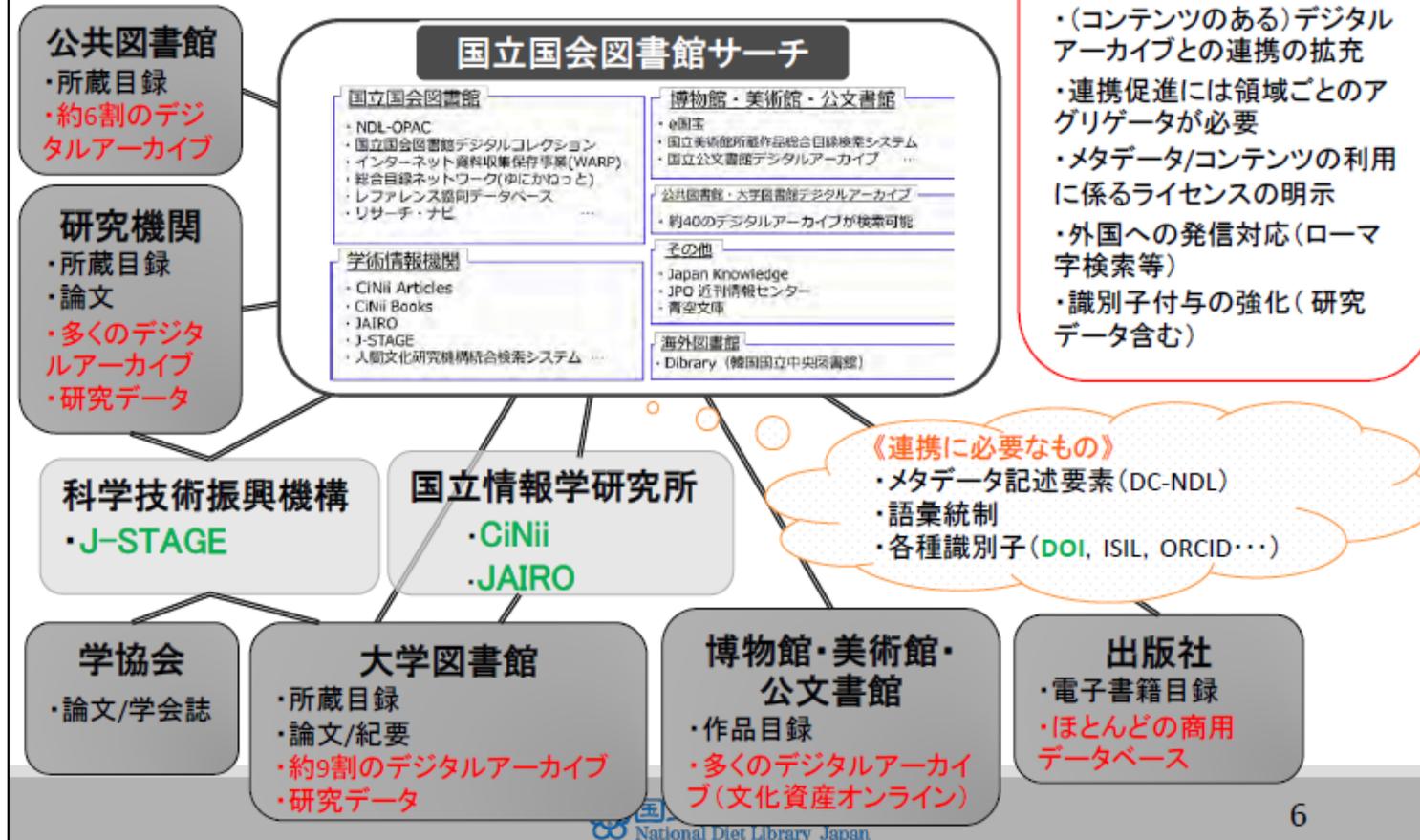
国会図書館と科学技術情報

③ 電子情報資源の組織化と連携

【現状】 国立国会図書館サーチによる連携先データベースは約100
赤字が検索対象になっていないもの。

【課題】

- ・(コンテンツのある)デジタルアーカイブとの連携の拡充
- ・連携促進には領域ごとのアグリゲータが必要
- ・メタデータ/コンテンツの利用に係るライセンスの明示
- ・外国への発信対応(ローマ字検索等)
- ・識別子付与の強化(研究データ含む)



第四期科学技術情報整備基本計画策定に向けた基本方針検討部会の設置について

1 基本方針検討部会の役割

国立国会図書館の第三期科学技術情報整備基本計画は、平成 27 年度までの計画となっており、国立国会図書館では、平成 28 年度から新たに 5 か年計画の第四期科学技術情報整備基本計画を策定することを予定している。計画の策定に当たっては、科学技術情報整備審議会から提言を頂戴し、提言を踏まえた計画にしたいと考えている。

基本方針検討部会は、科学技術情報整備審議会での提言をまとめるに当たり、以下の役割を担うことを想定している。

- (1) 提言において重要となる論点に関する議論
- (2) 提言素案の作成
- (3) 提言素案に係る科学技術情報整備審議会での議論を踏まえた提言案の作成

2 基本方針検討部会での検討テーマ（案）

(全体テーマ)

イノベーションを支える科学技術情報基盤—知識インフラの深化に向けて

(個別テーマ)

- ① 我が国で生成される電子情報資源への永続的アクセスの保障について
- ② 我が国における学術情報流通の現状と課題、今後の方向性について
- ③ オープンデータ促進と利活用における課題、社会への情報還元¹の在り方
- ④ 国としての外国雑誌等の科学技術資料・情報整備の在り方

議論はこれから

② インターネット資料・オンライン資料等

【現状】 枠内が収集・提供できているもの。枠外はできていないもの

【コンテンツ構築】

- ・インターネット資料(ウェブサイト)の収集:国・地方公共団体、大学等
- ・無償かつDRMなしのオンライン資料(PDF・EPUB・DAISY)の収集
- ・国のウェブサイトからのオンライン資料の切り出し
- ・電子版博士論文(機関リポジトリ以外)の収集

【コンテンツ提供】

- ・「国立国会図書館デジタルコレクション」&「WARP」による提供

国立国会図書館

【課題】

- ・学協会がもつ会議録等の収集(DRM付きor別フォーマット)
- ・文献と結びついた研究データの収集・蓄積・保存
- ・公的機関のオープンデータへの対応(アーカイブの検討)
- ・長期利用保証のためのデータ変換



科学の体制(これまでは主に論文文献の世界)

研究事業

研究主体・科学研究者



研究情報の出版・保存・利用事業

出版社・出版組織



「ライブラリ」、 学術研究情報組織



NII 機関リポジトリ

学術機関リポジトリ構築連携支援事業
NII Institutional Repositories Program

研究情報マネジメント、 インフラ構築



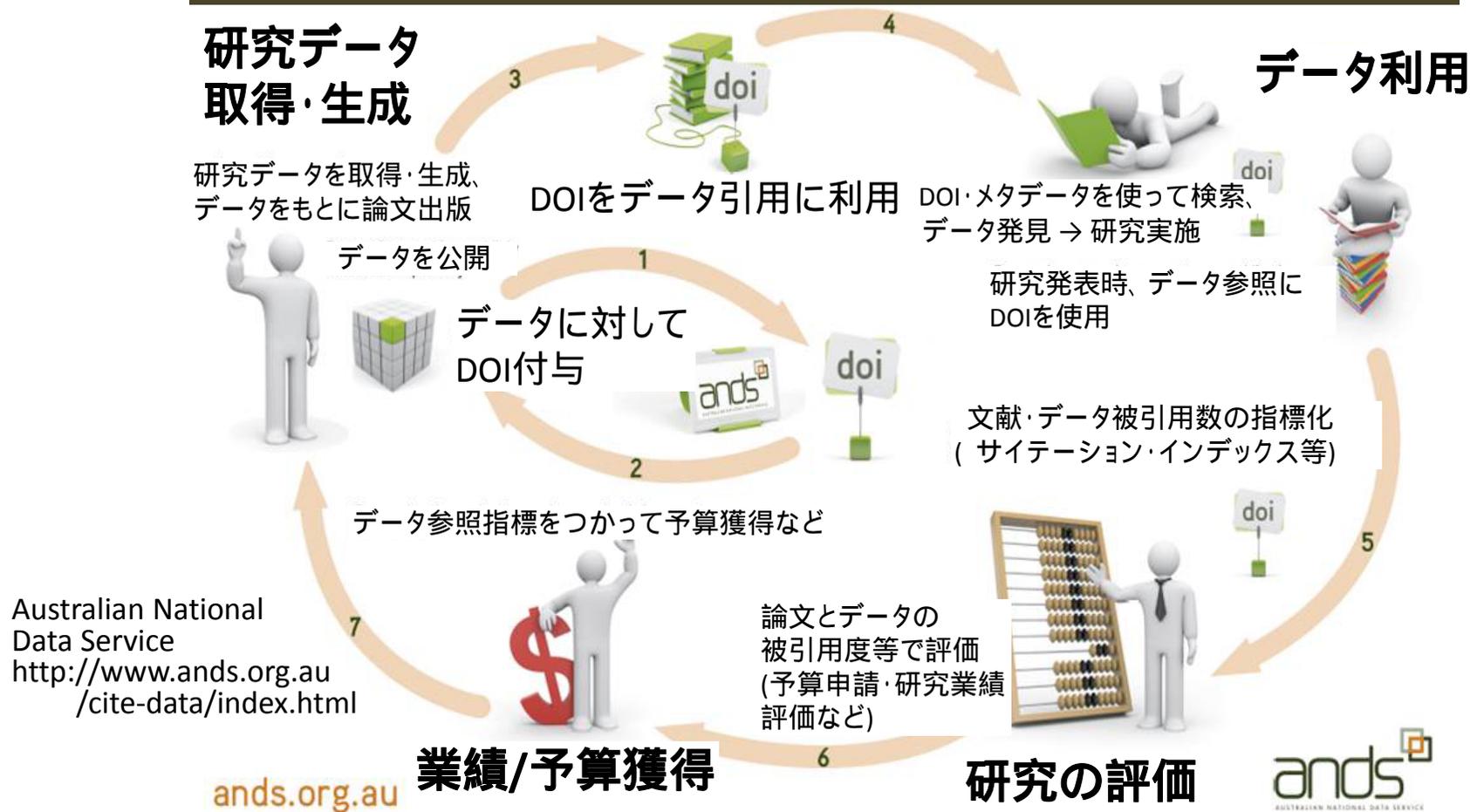
研究結果、データ、メタデータ、のやりとり



(参考:豪政府機関 Australian National Data Service による取組み)

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

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



Example of DOI-minting to Earth Science database in NOAA/NGDC

- EMAG2: Earth Magnetic Anomaly Grid (2-arc-minute resolution)

doi:10.7289/V5MW2F2P



http://www.ngdc.noaa.gov/nmmrview/metadata.jsp?id=gov.noaa.ngdc.mgg.geophysical_models:EMAG2&view=iso2html

Digital data

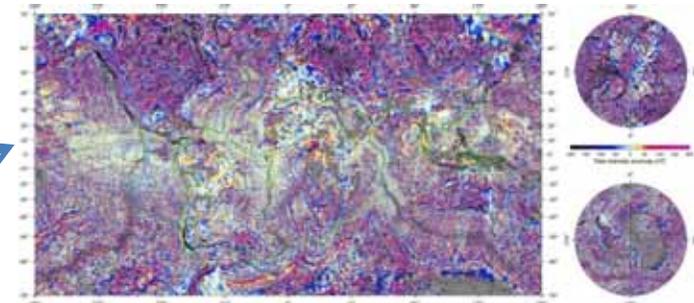
```

-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 -89.900000 -56.089413
-54.800000 -89.900000 -56.081806
-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
    
```

Landing Page

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

Data plot



Instruction of data citation

Maus (2009): EMAG2: Earth Magnetic Anomaly Grid (2-arc-minute resolution). National Geophysical Data Center, NOAA. Model, doi:10.7289/V5MW2F2P [access date]

SGEPSS将来構想から

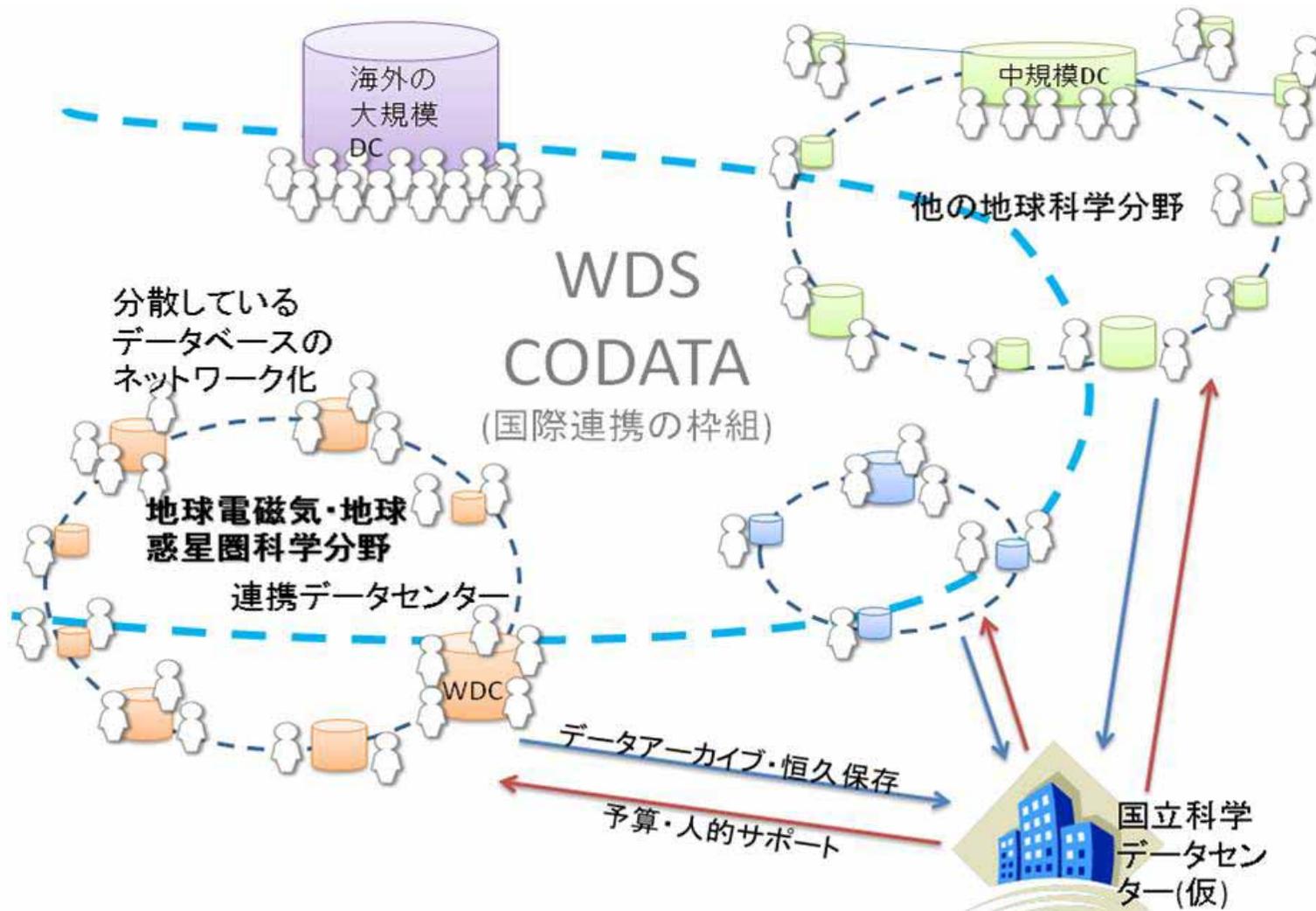


図4.3.1 連携データセンターと国立科学データセンターによる科学データサポート体制

「地球電磁気学・地球惑星圏科学の現状と将来」
(地球電磁気・地球惑星圏学会、2013年)

まとめにかえて

- オープン・データ
 - 科学と社会の関係に踏み込む話題
 - EU、米NSF、豪政府機関は動き出した。 RDA
- G8のプレッシャーと内閣府・CSTI
 - 日本政府が動き出した
 - 今後国内各省庁・機関がどの程度うごくか？
 - 「オープン」の意味に注意！
- どの機関が何をするか
 - データ保存・メタデータ整備と図書館
 - 図書館は危機感をもっている
 - しかしまだ、データに踏み出せてない
 - 研究組織と図書館組織の連携が重要ではないか。
- 科学のシステムを科学者(≠研究者)が守れるか。

日本学術会議主催：日本学術会議主催学術フォーラム 「科学を変えるデータジャーナル

- 科学技術データの共有・再利用の新たなプラットフォーム構築へむけて -」

平成27年3月4日(水)、日本学術会議講堂にて

– 登壇予定者(敬称略)：

- 原山優子(総合科学技術イノベーション会議議員)
 - 大竹 暁(科学技術振興機構理事)
 - 喜連川優(国立情報学研究所所長)
 - 岩田修一(DSJ編集委員長, 国際サイエンスデータ分科会)
 - Krishan Lal(AASSA(アジア学術アカデミー)会長)
 - Paul F. Uhler(全米科学アカデミー データ政策顧問)、
他
- 科学技術データ共有と活用に関する研究集会(仮)
 - 平成27年3月5日(日本学術会議)
 - RDAなど国際データ活動の対応体制構築へ向けて
 - JST、AIST、NII、 . . .