

【公開シンポジウム】

オープンサイエンス時代における学術データ・
学術試料の保存・保管、共有問題の現状と将来

「学協会における論文出版と研 究データ保存管理」

大谷 栄治

(日本学術会議連携会員、 東北大学名誉教授・学術研究者)

論文出版と学術の発展のために

1. 論文出版の質の評価とジャーナルインパクトファクター

- ・ JIFの引用数を論文の質とする判断には問題あり。
- ・ 質はJIFに基づいたQ1～Q4の分類に限界がある。
しかし、現状では無視できない
- ・ 研究評価のあり方と評価の基準や評価対象の変容：
(論文からデータの質や公開・保存状況、プレプリント、ソフトウェア、
物理サンプルetc.)
- ・ 学会誌の利点・欠点：利点：商業誌にない**質の高い学会を基盤**，
欠点：国際性，宣伝

2. APC (Article Processing Charge)の高額化への対応

商業出版社プラットフォームは高額：**商業誌と学会誌**

科研費に計上可能，しかしAPCへの国や機関（大学）からの**支援が必要**

3. ジャーナルにとってのData repository

日本発のジャーナルに紐付く信頼性の高いデータを海外に散逸させていいのか？

欧米ファンドのリポジトリに日本の知的財産を集積するのは正しいか？

国内に信頼できるData repositoryが必要

オープンサイエンスと学術成果の発表における 不可欠な媒体

- Preprint server:

(arXiv etc, Jxiv since 2022.3, preprint server by JST)

- Journal:

Open accessの重要性, 査読を通じた質の保証

- Data Repository

(J-Stage Data, since 2021)

大手学術出版社のデータポリシー (主に地球・宇宙科学分野のジャーナルから)

- **Willey/AGU** publication requirement:
” All new **data and software** for submissions **should be placed into community-accepted, trusted repositories** (not in supporting information).”
- **Springer-Nature** /”Research Data Policies”
“to publish the best research, which includes achieving community best practices in **the sharing and archiving of research data.** We also aim to facilitate compliance with **research funder and institution requirements to share data.**”
- **Elsevier** /J. Atmos. Solar Terr. Phys.:
“This journal **requires and enables you to share data**...also encourages you to share your software, code, models, algorithms, protocols, methods and other useful materials...authors are invited to **add a citation to their datasets** by including a data reference **in their Reference List**

出版論文に著者のデータの公開, 使用データを引用すべき. FAIR: findable, accessible, interoperable, reusable

COPDESS

launched on 15 January 2014

AGU理事 K.Lehnert 氏
のスライド (JpGU、
2017) から引用、
村山が修正



Coalition for Publishing Data in the Earth and Space Sciences



COPDESS (地球宇宙科学のデータ出版に関する連合) :
出版者、データ施設、公的機関等とともに、**オープン**
なデータ利活用を目指す。



- 分野内のメタデータ標準化
- データ引用の推進
- 研究者への周知、啓蒙
- データパブリケーション・ワークフロー (査読・エディトリアル作業、エンバゴ期間、etc.)
- 永続的識別子の利用促進 (DOI、ORCID、Fundrefなど)

Time to change the data culture in geochemistry

NATURE REVIEWS | EARTH & ENVIRONMENT
VOLUME 2 | NOVEMBER 2021 | 737

Katy J. Chamberlain¹✉, Kerstin A. Lehnert², Iona M. McIntosh³, Dan J. Morgan⁴ and Gerhard Wörner⁵

Geochemical data are vital for understanding Earth's past, present and future. However, currently only a fraction of geochemical data are findable, accessible, interoperable and reusable, limiting their use in the broadest range of scientific studies. There is an urgent need for international coordination of geochemical data and methods to unlock their full research potential.

Table 1 | Minimum required geochemical metadata

Meta-data descriptor	Required meta-data				
Geographic information					
All samples	Sample number	Sample lat.–long.	Lat.–long. accuracy	Elevation	Elevation accuracy
Sample type					
All samples	Sampling time	Time zone	Collector or scientist	Archiver or curator	Sampling method
Rocks	Chemical characterization	Sediment grain size	-	-	-
Water and gas	Sampling container	Sampling treatment	Fluid type	Fluid classification	-
Fossils	Fossil group	Fossil species	Fossil preservation	-	-
Minerals	Mineral name	Occurrence type	-	-	-
Laboratory information					
Interoperability	Analyst and laboratory	Analytical code	Instrument used	Measured parameter	Technique description
Analysis information					
Reproducibility	Reference sample name	Uncertainty value	Uncertainty unit	-	-
Normalization	Reference sample name	Normalization value	Normalization unit	-	-
Fractionation	Reference sample name	Fractionation parameter	Fractionation value	Fractionation unit	-
Detection Limit	Detection limit	Detection unit	-	-	-
Blank	Blank type	Blank value	Blank unit	-	-

Minimum required metadata¹ that should be submitted alongside all geochemical datasets, irrespective of whether samples are rock, water or gas.

Time to change the data culture in geochemistry

NATURE REVIEWS | EARTH & ENVIRONMENT
VOLUME 2 | NOVEMBER 2021 | 737

Katy J. Chamberlain¹, Kerstin A. Lehnert², Iona M. McIntosh³, Dan J. Morgan⁴ and Gerhard Wörner⁵

Geochemical data are vital for understanding Earth's past, present and future. However, currently only a fraction of geochemical data are findable, accessible, interoperable and reusable, limiting their use in the broadest range of scientific studies. There is an urgent need for international coordination of geochemical data and methods to unlock their full research potential.

RELATED LINKS

AstroMat database: <https://www.astromat.org/>

AuScope Geochemistry Network (AGN): <https://www.auscope.org.au/agn>

DIGIS initiative: <https://www.uni-goettingen.de/de/643369.html>

EarthChem Library: <https://earthchem.org/ecl/>

GEOROC database: <http://georoc.mpch-mainz.gwdg.de/georoc/>

Multi-Scale Laboratories of the European Plate Observing System (EPOS):

<https://www.epos-eu.org/epos-services-list>

North American Volcanic and Intrusive Rock Database (NAVDAT): <https://www.navdat.org/>

OneGeochemistry: <https://www.earthchem.org/communities/onegeochemistry/>

OSIRIS-REx: <https://www.asteroidmission.org>

PetDB database: <https://search.earthchem.org>

FAIR: findable, accessible, interoperable, reusable :
あらゆる地球科学試料のメタデータ（索引）の共通化

Data repositories recommended by Springer–Nature

Mandated data types

Biological sciences

Chemistry and chemical
biology

**Earth, environmental
and space sciences**

Health sciences

Materials science

Physics

Social science

Generalist repositories


Generalist repositories

• Dryad Digital Repository 

• Figshare (2011) 

• Harvard Dataverse 

• Open Science Framework 

• Science Data Bank 
(CAS, 2014)


• Zenodo  (CERN, 2013)

Earth, environmental and space sciences

Broad scope earth and environmental sciences

NERC Data Centres (2010) 

NASA Goddard Earth Sciences Data and
Information Services Center 

PANGAEA (1987, 1992) 

National Tibetan Plateau/Third Pole Environment
Data Center (2019) 

NOAA National Centers for Environmental
Information 

(DOIs only assigned to deposited data on
request)

HydroShare (CUAHSI) 

「オープンサイエンス」を進める上での我が国の問題点

- 研究論文原稿を国際誌に投稿するときに、データの公開・共有が要求される例が増加
- 国内研究者の声の例：
 - 「適切なデータの置き場が国内にない!」
 - 「よいデータを整備、公開したり、オープンサイエンス活動・市民連携を推進しても、学術業績・貢献に認められない!」

⇔ 欧米では「研究・論文のための」科学予算でなく「データやプロジェクト連携を目的とした」科学予算枠も生まれている。

DFGの例：<https://www.uni-goettingen.de/de/digital+geochemistry+infrastructure/643369.html>

DFG project (DIGIS: Digital Geochemistry Infrastructure)

JSPSでは、出版助成はあるが、**科研費のカテゴリーにはない。**

しかし「挑戦的研究（開拓）」に申請・受理の例あり

国立研究開発法人における データポリシー策定のためのガイドライン ～解説資料～（内閣府）

2019年（平成31）年4月2日

1. 本ガイドラインの位置付け

2. データポリシー策定のポイント及び並行して取り組む事項

- (1) ポリシー策定の目的
- (2) ポリシー策定の主体
- (3) 管理対象とするデータが具備すべき要件
- (4) データ利活用のための要件
- (5) ポリシー策定とともに取り組むべき事項
- (6) その他

3. データポリシーで定めるべき項目

- (1) 機関におけるポリシー策定の目的について
- (2) 管理する研究データの定義、制限事項について
- (3) 研究データの保存・管理・運用・セキュリティについて
- (4) 研究データに対するメタデータ、識別子の付与、フォーマットについて
- (5) 研究データの帰属、知的財産の取り扱いについて
- (6) 研究データの公開、非公開及び猶予期間ならびに引用について

国立研究開発法人における データポリシー策定のためのガイドライン ～解説資料～（内閣府）

2019年（平成31）年4月2日

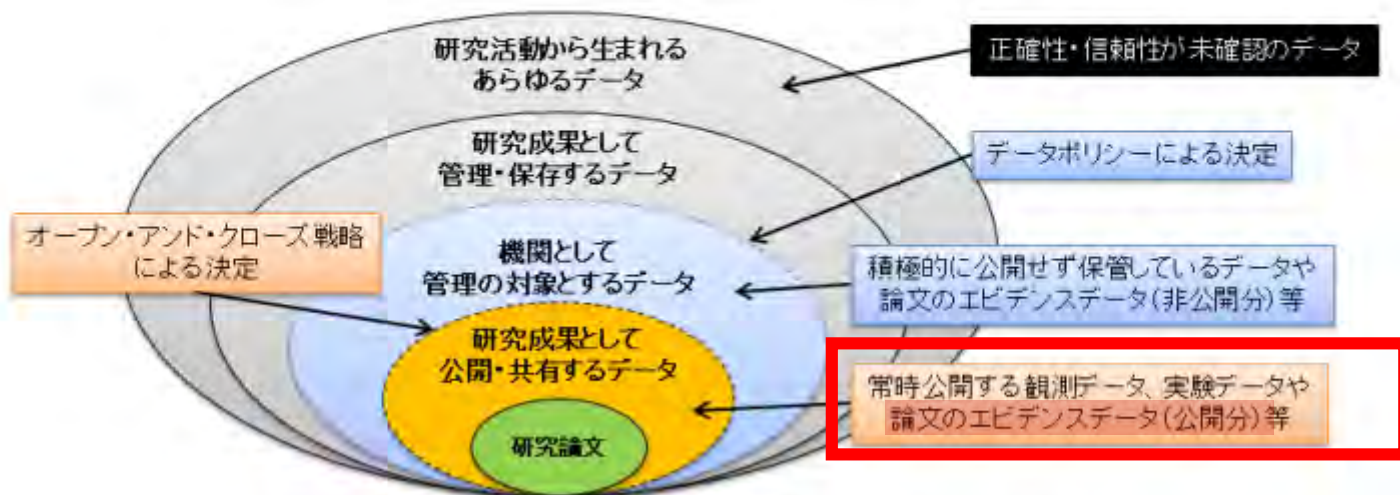
本ガイドラインの位置づけ

研究データの利活用を促進する取組は、そのもたらす利益が国・企業・学界に留まらない活動として、**科学技術に拠って立つ我が国が**、G7等の国際動向を踏まえ、責任をもって、かつ、**率先して取り組むべき課題**と位置付けられる。

「**機関リポジトリ**」: 大学で主に論文の搭載を目的に整備・運用されている

「**分野別リポジトリ**」: 研究分野ごとに整備:国内では生命科学分野のデータリポジトリとして、「**バイオサイエンスデータベースセンター(NBDC, National Bioscience Database Center)**」

国立研究開発法人における データポリシー策定のためのガイドライン ～解説資料～（内閣府）



参考「機関として管理の対象とするデータ」の考え方一例

物材研究機構

JAMSTEC

JAXA

環境研

参考資料：

○国立研究開発法人物質・材料研究機構 研究データポリシー

http://www.nims.go.jp/nims/disclosure/hdfqf10000001742-att/NIMS_research_data_policy_20180801.pdf

○国立研究開発法人海洋研究開発機構

データ・サンプルの取り扱いに関する基本方針（データポリシー）

https://www.jamstec.go.jp/j/database/data_policy.html

○国立研究開発法人宇宙航空研究開発機構 宇宙科学研究所のデータポリシー

<http://www.isas.jaxa.jp/researchers/data-policy/>

○国立研究開発法人国立環境研究所 データの公開に関する基本方針（データポリシー）

https://www.nies.go.jp/kihon/kitei/kt_datapolicy.pdf

東北大学研究データ管理・公開ポリシーの解説 (2021)

(趣旨)

東北大学 (以下「本学」という。) は、建学以来の伝統である「研究第一主義」と「門戸開放」の理念を掲げている。

研究データを適切に管理することは優れた研究を行う上で必要不可欠であり、また、研究データは学術や社会の発展に必須の知の基盤の一つである。

そこで、本学は、研究データを適切に管理し、その公開等を通じて利活用を促進することにより、本学の理念を実践し、もって研究中心大学として人類社会の持続的発展に貢献することを目的として、研究データの管理、公開及び利活用に関するポリシーを以下のように定める。

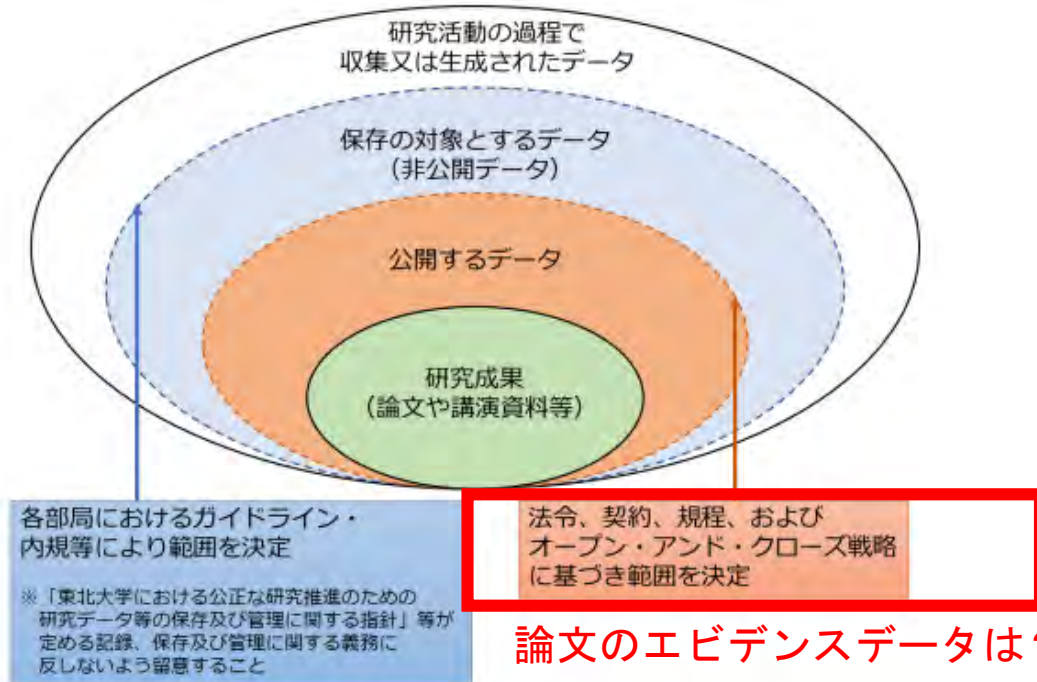


図 1. 研究データとその区分範囲のイメージ

わが国におけるオープンサイエンス： 地球観測・宇宙惑星探査データ

データを**使用する**利用研究者・行政・市民への
サービスを目指す

地球・惑星・環境科学の優れた**大型データの公
開基盤**：さらなる公開の拡大

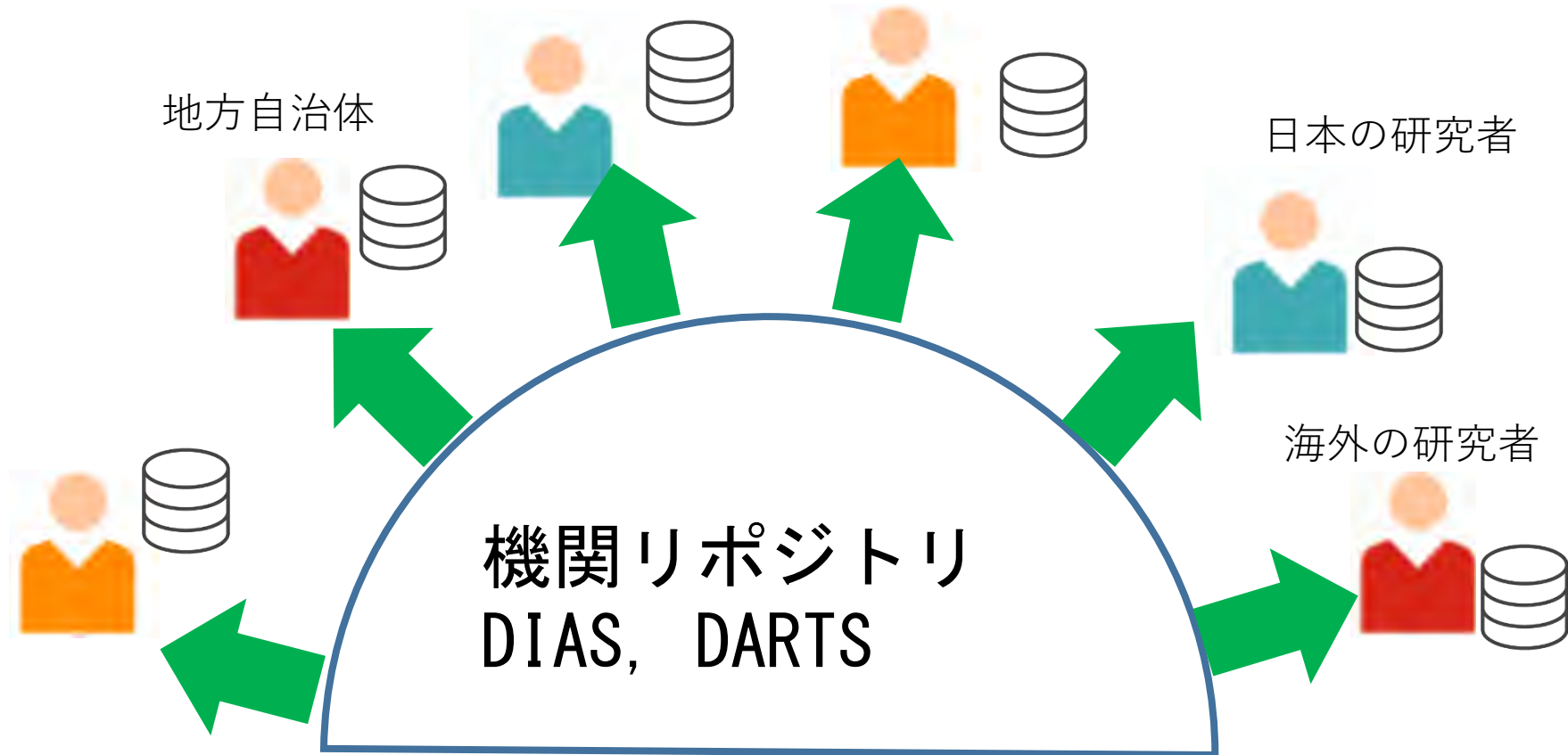
DARTS (JAXA's space science missions)

DIAS (地球規模/各地域の観測で得られたデータ
を収集, 蓄積, 統合, 解析

機関レポシトリー (様々な国研)

・

GED (Global Environmental database) 国立環
境研究所



DIAS: 地球規模/各地域の観測で得られたデータを収集，蓄積，統合，解析するとともに，社会経済情報などとの融合を行い，地球規模の環境問題や大規模自然災害等の脅威に対する危機管理に有益な情報へ変換し，**国内外に提供する。**

全球地球観測システム（GEOSS：Global Earth Observation System of Systems）に参加する世界各国のデータセンターとの接続を実現し，GEOSS への**国際貢献**

<https://www.re3data.org/>

DARTS

ISAS Data Archive and Transmission System

DARTS primarily **archives high-level data products obtained by JAXA's space science missions** in astrophysics (X-rays, radio, infrared), solar physics, solar-terrestrial physics, and lunar and planetary science. **In addition, we archive related space science data products obtained by other domestic or foreign institutes, and provide data services to facilitate use of these data.** 研究者の論文の2次データも保存

DIAS

Data Integration and Analysis System

DIAS aims at collecting and storing earth observation data; analyzing such data in combination with socio-economic data, and converting data into information useful for crisis management with respect to global-scale environmental disasters, and other threats; and to **make this information available within Japan and overseas.**

Life Science Database Archive (LSDB Archive) (NBDC @JST)

The Life Science Database Archive maintains and **stores the datasets generated by life scientists in Japan in a long-term and stable state as national public goods.** The Archive makes it easier for many people to search datasets by metadata (description of datasets) in a unified format, and to access and download the datasets with clear terms of use. (NBDC program, JST)

J-STAGE Data (JST)

J-STAGE Data supports the publication and distribution of data related to J-STAGE articles with the aim of contributing to the promotion of open science in Japan. Journals published on J-STAGE can use J-STAGE Data to publish data related to their own articles. DOI is automatically assigned to the data published in J-STAGE Data

現状の問題点：Data repositoryが 双方向でない

DARTS, DIAS

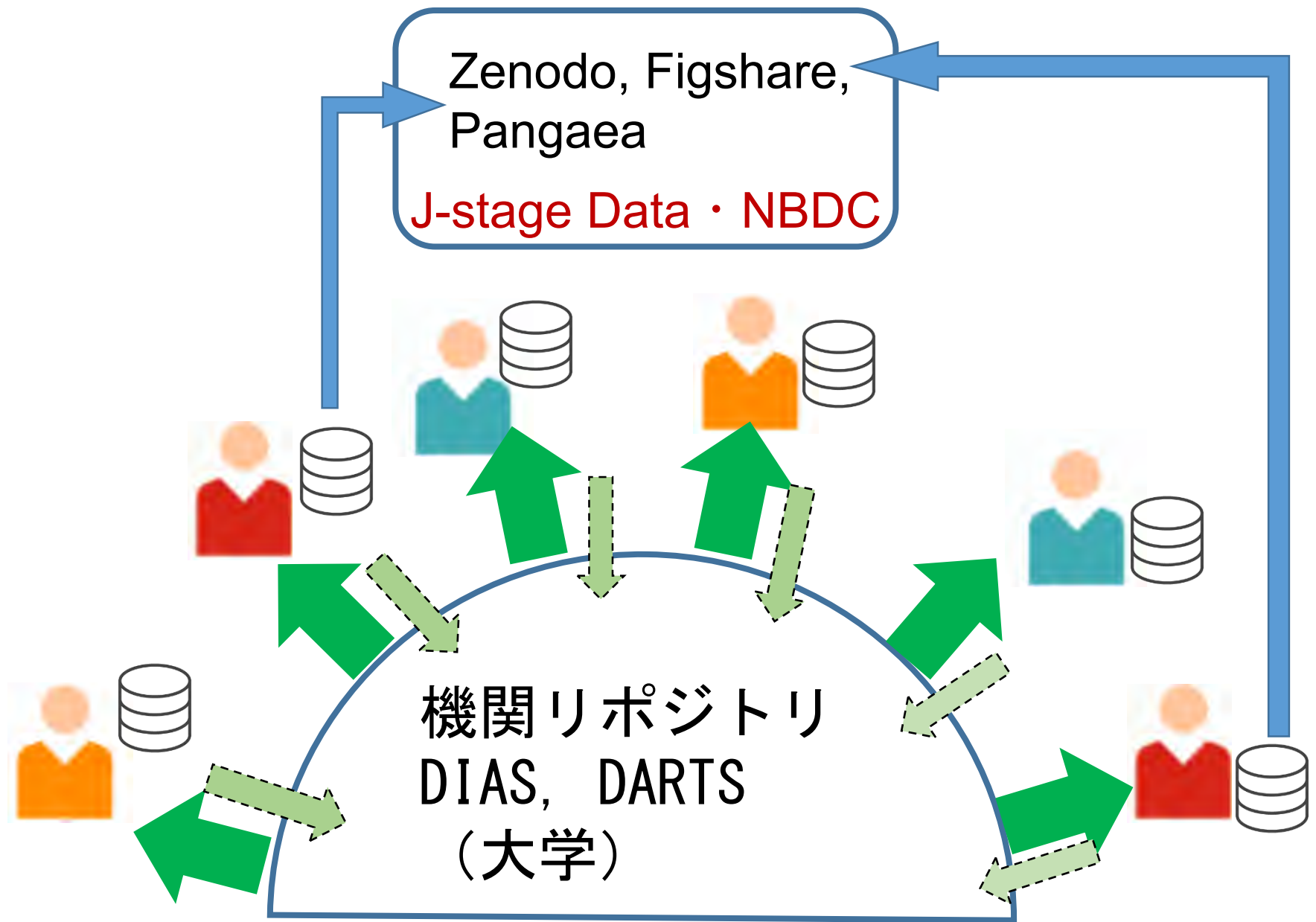
個別の研究者のデータのアーカイブの仕組みはない。

例：

宇宙科学研究所は、論文や学会等で発表された結果を再現するために必要なデータ等、科学的成果のエビデンスとなるデータを公開するという研究データポリシーあり

他の国研の研究データポリシーにも同様？

「プロジェクト」とは独立に、研究者のエビデンスデータの保存・公開の仕組みが必要



わが国における研究者のためのData repositoryの先行例

- ・分野別data repository

NBDC（バイオサイエンスデータベースセンター）の

LSDB Archive：生命科学に限られる。

[NBDC事業推進部 - NBDC \(biosciencedbc.jp\)](http://biosciencedbc.jp) (JST) 2011年に発足

- ・ジャーナル論文のdata repository

J-stage Data (2021年に発足)

J-stage journalsに限られる

- ・全分野

大学の機関リポジトリ

先進的な取り組み：京都大学，名古屋大学

全国の大学；時間がかかりすぎるのではないか？最終目標は？

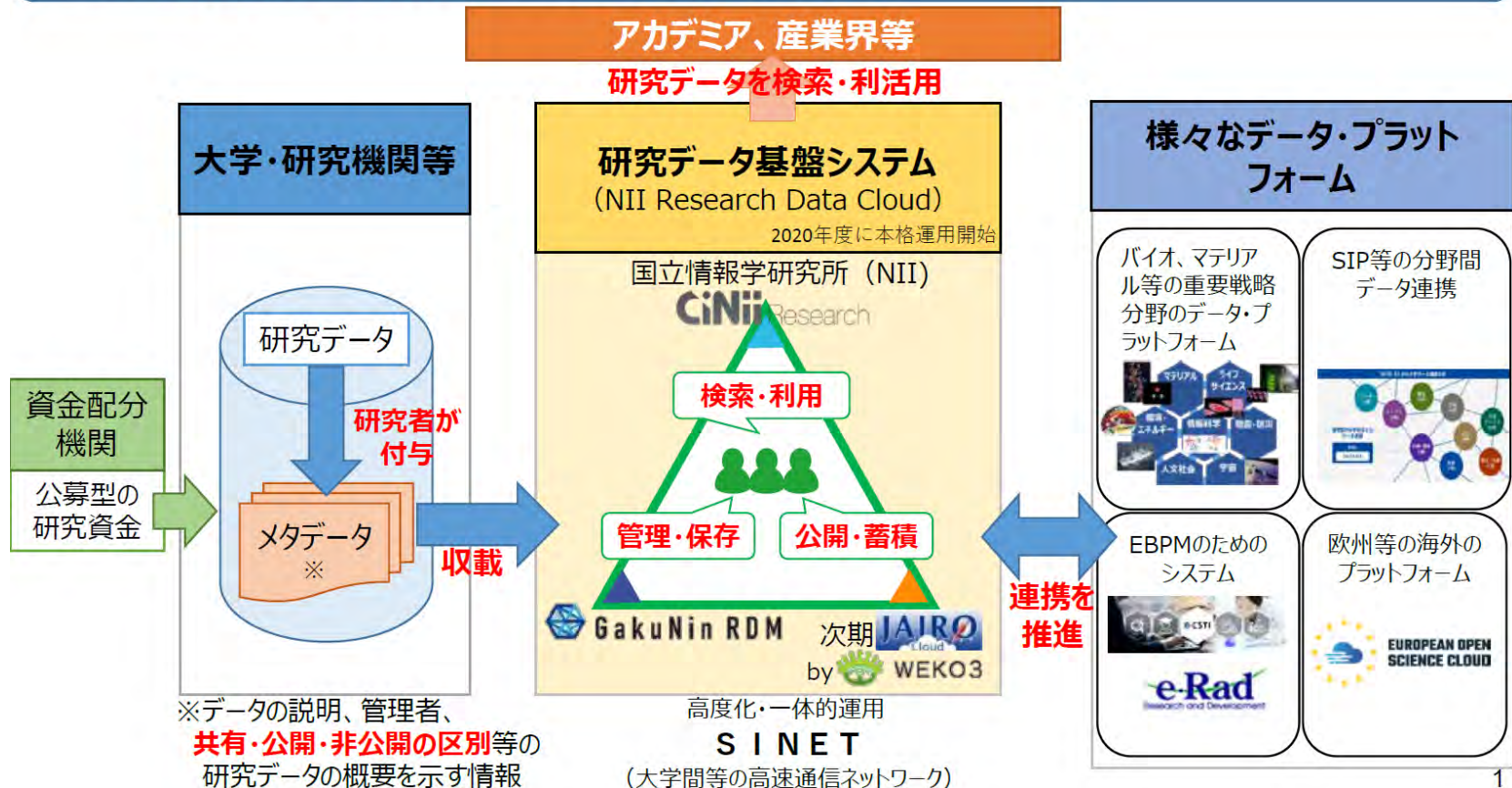
全体像とスケジュール（内閣府：2025年までにRDPを100%に，DMPを2023年から：JSPSの科研費(学術変革領域A, B)に求める（様式A-18）

公的資金による研究データの管理・利活用に関する基本的な考え方について

研究データ基盤システムを中核としたデータ・プラットフォームの構築

- 研究データの公開・共有を推進、産学官のユーザが**データを検索可能**
- ムーンショット型研究開発制度**における試行(2020年度開始)、その後、次期**SIP**に導入

→ **全ての公募型の研究資金**の新規公募分に導入 **DMP (Data Management Plan)** を求める



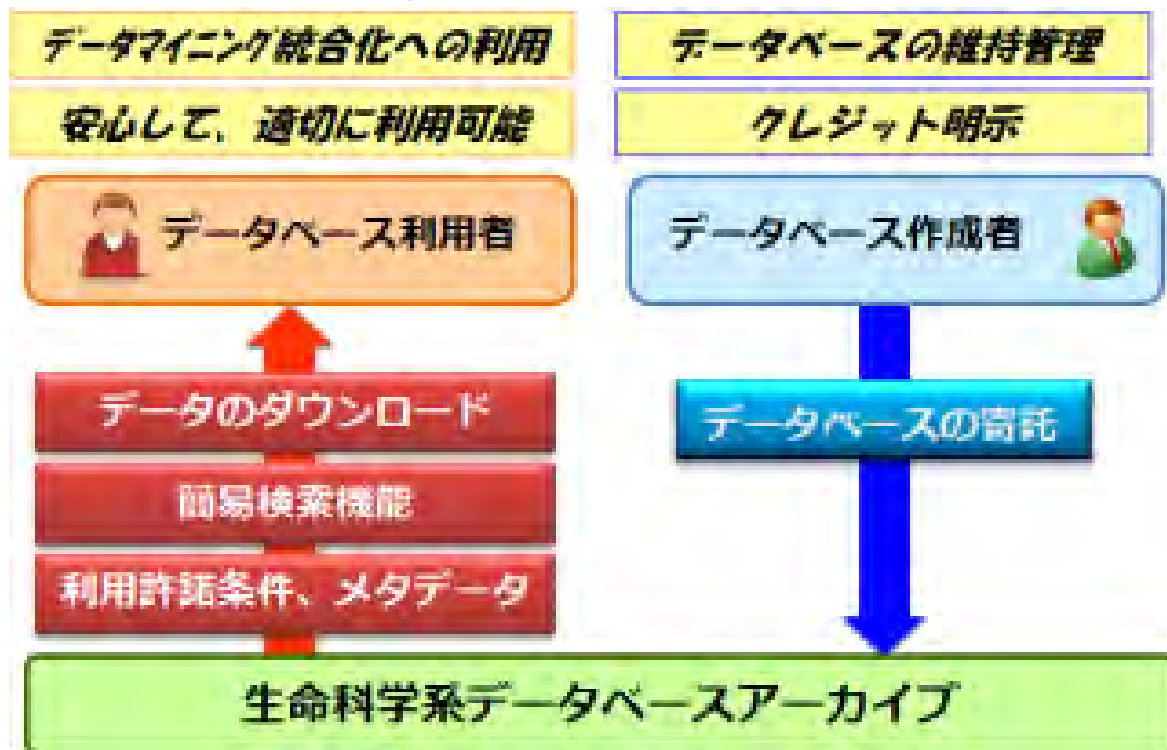
わが国における分野別Data repositoryの例

・ 分野別data repository: **LSDB Archive**

バイオサイエンスデータベースセンター : NBDC (JST) 2011年
に発足 : [NBDC事業推進部 - NBDC \(biosciencedbc.jp\)](http://biosciencedbc.jp)

国内のライフサイエンス研究者が生み出した**データセットをわが国の公共財**
としてまとめて長期間安定に維持保管し、 データ説明 (メタデータ) を統一
して検索を容易にすると共に、 利用許諾条件などの明示を行う

研究者にとって双方向



適用範囲 :

公的資金を用いたプロジェクト等で産生された生命科学分野の研究データ (ゲノム配列、タンパク質、代謝産物、画像等) 及び公的資金を用いたプロジェクト等で産生されたヒトに関するデータ (DDBJと連携、PDBjを一部支援)。

J-STAGE DATA



J-STAGE Data is a research data repository service provided by **JST** (Japan Science and Technology Agency) to advance research data sharing and reuse.

Started from 2021 (registered in re3data.org, Nov.2021)

20 Journals using J-stage platform Journals

J-stage platform (since 1998): 3000 journals from 1500 society/organization in Japan (in 2020)

研究者のための日本のData Repositoryの必要性

データは重要な知的資産

日本発のジャーナルに紐付く信頼性の高いデータを海外に散逸させて良いのか？

欧米ファンドのレポジトリに日本の知的財産を集積するのは正しいか？

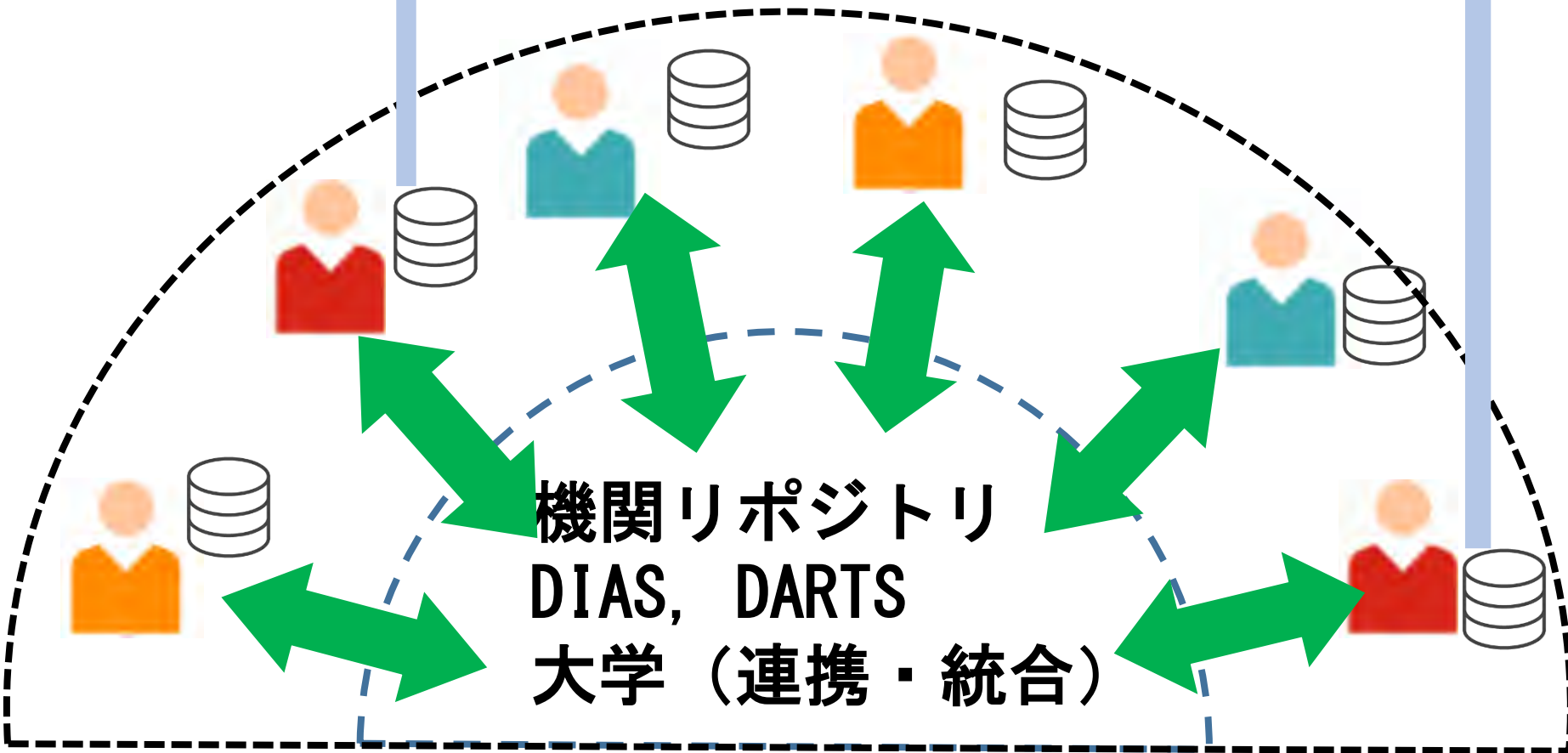
現状では、散逸しつつある→ 早急に対応が必要では。

可能な方法

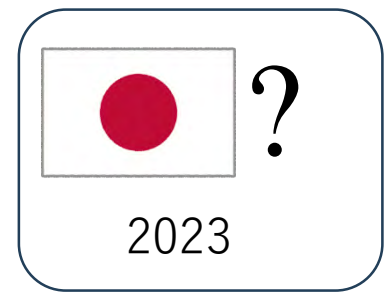
1. LSDB Archive(NBDC)タイプの地球惑星環境分野の National Data repository？
2. 国研のData repositoryの連携・統合？
3. J-stage Dataの対象拡大；J-stage journal以外にも適用？
4. 大学の機関repositoryの早急な確立：大学間の機関repositoryの連携：全体像とスケジュール？

Data Policy策定(100% by 2025)→ Data Management Plan(2023~)
第6期科学技術イノベーション基本計画

Zenodo, Figshare,
Pangaea
J-stage Data · NGDC



Data repository guidance



[Mandated data types](#)

[Biological sciences](#)

[Chemistry and chemical
biology](#)

[Earth, environmental
and space sciences](#)

[Health sciences](#)

[Materials science](#)

[Physics](#)

[Social science](#)

[Generalist repositories](#)

[Earth, environmental and space sciences](#)

Broad scope earth and environmental sciences

[NERC Data Centres](#)



[Generalist repositories](#)

• [Dryad Digital Repository](#)



• [Figshare](#) (2011)



• [Harvard Dataverse](#)



• [Open Science Framework](#)



• [Science Data Bank](#)



(CAS, 2014)

• [Zenodo](#)



(CERN, 2013)

[NASA Goddard Earth Sciences Data and
Information Services Center](#)



[PANGAEA](#)



[National Tibetan Plateau/Third Pole Environment
Data Center](#)



[NOAA National Centers for Environmental
Information](#)



(DOIs only assigned to deposited data on
request)

[HydroShare \(CUAHSI\)](#)



ご清聴ありがとうございました.