

# 第1回

# DIASオープンサイエンスセミナー DOIの利活用を中心とした世 界の動向とDIASの方針

国立情報学研究所

北本 朝展 (KITAMOTO Asanobu)

<https://dias.ex.nii.ac.jp/>

本日の発表資料は後日公開します。

# DIASとオープンサイエ ンス

# オープンサイエンス

- サイエンスをより「オープン」にすること。ただし「**オープン**」は**多義的**である。
  1. **共有**：誰でもアクセスして利用できるように公開すること。
  2. **透明性**：誰でもデータを検証できるように公開すること。
  3. **参加・協働**：誰でも研究活動に参加できるように境界を下げること。

# オープンサイエンスへの収束

透明性

オープンアクセス

共有

オープンピアレビュー

オープンデータ

研究の再現性・  
透明性・研究  
データ保存

研究データ管理  
データ出版  
データリポジトリ

市民科学・クラウド  
ファンディング

コラボレーション・オープンイノベーション

超学際研究

参加

協働

メタ研究 = 研究（システム）に関する研究

# 第5期科学技術基本計画

## ③ オープンサイエンスの推進

オープンサイエンスとは、オープンアクセスと研究データのオープン化（オープンデータ）を含む概念である。オープンアクセスが進むことにより、学界、産業界、市民等あらゆるユーザーが研究成果を広く利用可能となり、その結果、研究者の所属機関、専門分野、国境を越えた新たな協働による知の創出を加速し、新たな価値を生み出していくことが可能となる。また、オープンデータが進むことで、社会に対する研究プロセスの透明化や研究成果の幅広い活用が図られ、また、こうした協働に市民の参画や国際交流を促す効果も見込まれる。さらに、研究の基礎データを市民が提供する、観察者として研究プロジェクトに参画するなどの新たな研究方策としても関心が高まりつつあり、市民参画型のサイエンス（シチズンサイエンス）が拡大する兆しにある。近年、こうしたオープンサイエンスの概念が世界的に急速な広がりを見せており、オープンイノベーションの重要な基盤としても注目されている。

こうした潮流を踏まえ、国は、資金配分機関、大学等の研究機関、研究者等の関係者と連携し、オープンサイエンスの推進体制を構築する。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。その他の研究成果としての研究二次データについても、分野により研究データの保存と共有方法が異なることを念頭に置いた上で可能な範囲で公開する。

ただし、研究成果のうち、国家安全保障等に係るデータ、商業目的で収集されたデータなどは公開適用対象外とする。また、データへのアクセスやデータの利用には、個人のプライバシー保護、財産的価値のある成果物の保護の観点から制限事項を設ける。な

32

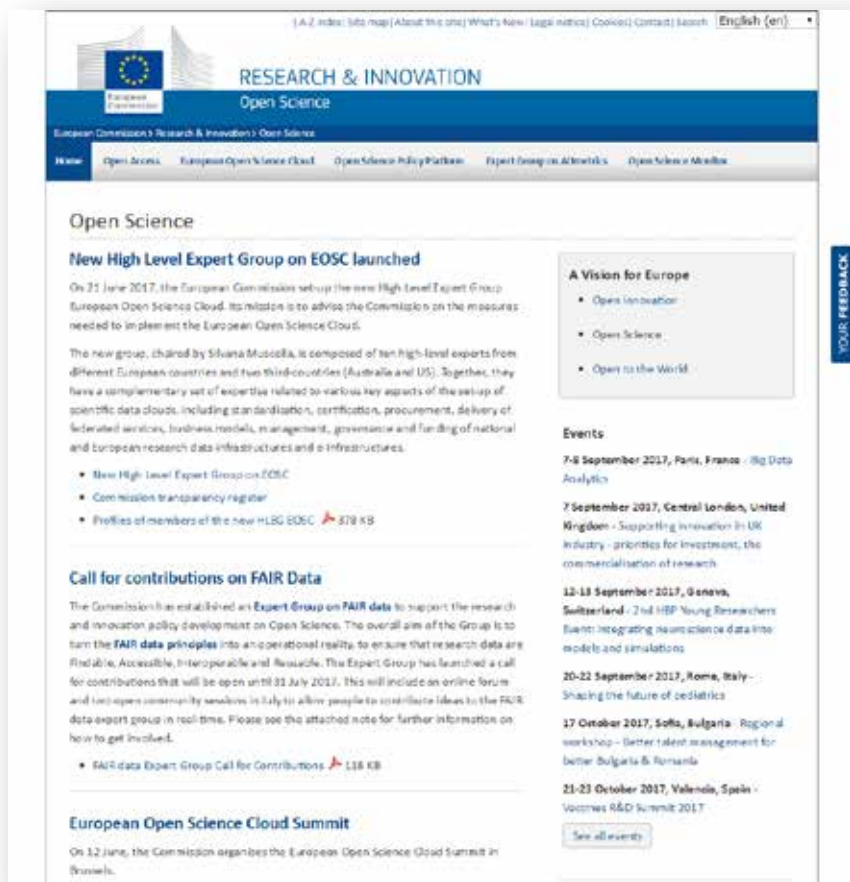
お、研究分野によって研究データの保存と共有の方法に違いがあることを認識するとともに、国益等を意識したオープン・アンド・クローズ戦略及び知的財産の実施等に留意することが重要である。

また、国は、科学研究活動の効率化と生産性の向上を目指し、オープンサイエンスの推進のルールに基づき、適切な国際連携により、研究成果・データを共有するプラットフォームを構築する。

- 平成28年度～平成32年度の基本計画。
- 研究データの公開や活用、透明化、プラットフォームなどを列挙。
- オープンアクセスや市民科学、オープンイノベーションなどにも言及。

<http://www8.cao.go.jp/cstp/kihonkeikaku/5honbun.pdf>

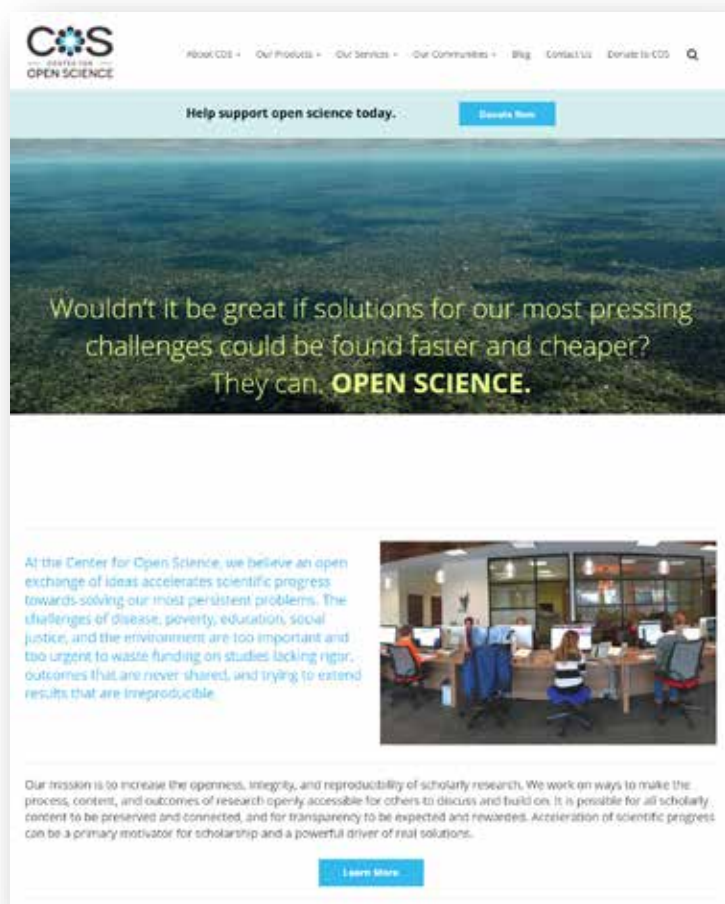
# European Open Science Cloud



<https://ec.europa.eu/research/openscience/index.cfm>

- EUもオープンサイエンスを傘として、その下に多くの取り組みを集約している。
- European Open Science Cloudなどでe-Scienceを発展させる計画。
- 学術出版の未来や研究データの問題も取り上げている。

# Open Science Framework



- The Center for Open Scienceが取り組む、新しい研究とデータの基盤。
- コミュニティのための、より早く便利な情報共有。
- 研究とはどうあるべきか？の研究も。

<https://cos.io/>

# 研究データ利活用協議会

<http://japanlinkcenter.org/rduf/>



- 国際的にはRDA (Research Data Alliance)等の有力なコミュニティが存在する。
- ジャパンリンクセンター (JaLC) が、研究データのDOI登録実験を実施。
- この活動が母体となって協議会を設立。分野横断型の研究データコミュニティで、研究会なども開催予定。

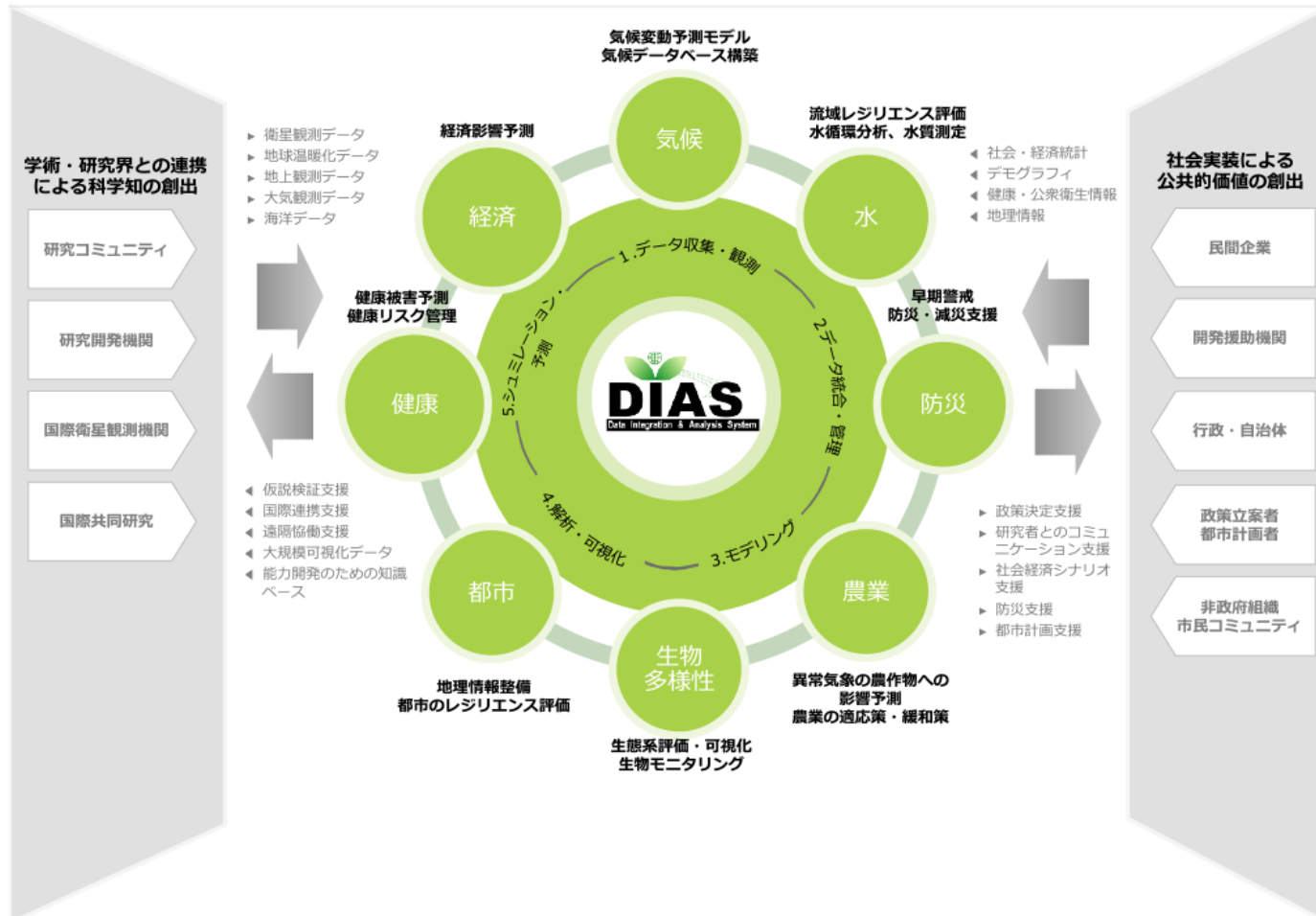


# DIASとオープンサイエンス

1. DIASでは「**オープンサイエンス分科会**」を設置し、オープンサイエンスへの対応を進めている。
2. データ共有の促進という観点から、**DOIの付与やデータ引用・データジャーナルへの対応**から開始している。
3. **認証の多様化**などにも着手し、**学認**などからのログインも可能にする計画。

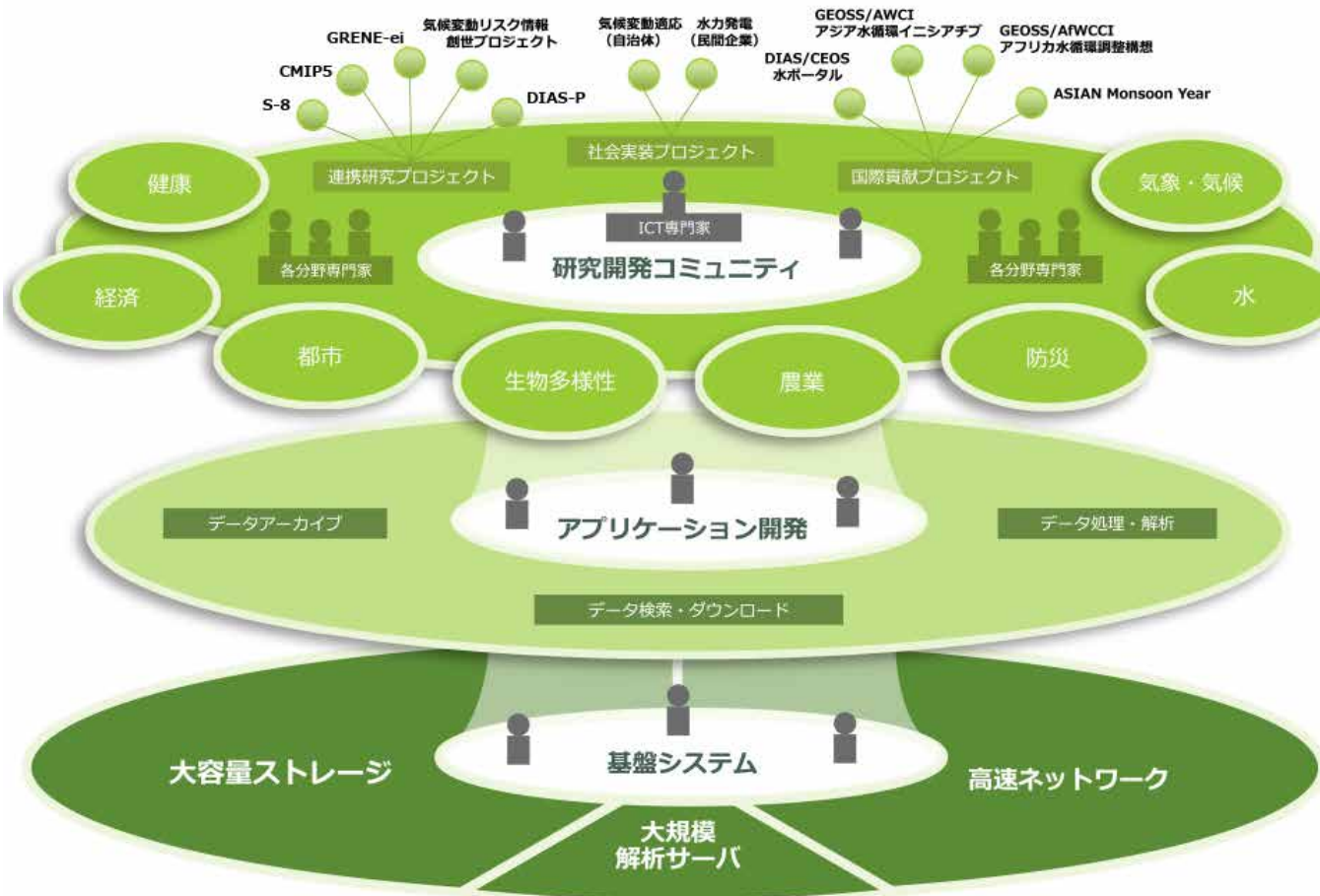
# DIAS = データ統合解析システム

<http://www.diasjp.net/>



# DIAS = データ + コミュニティ

<http://www.diasjp.net/>



コミュニティ基盤

データ基盤

# 3つのDIASコミュニティ

1. **DIASリサーチコミュニティ**：ドメイン専門家と情報学専門家が協働して価値を生み出すコミュニティ。
2. **DIASプロバイダコミュニティ**：DIASにデータを提供して安定的な利用を促進するコミュニティ。
3. **DIASユーザコミュニティ**：DIASが公開するデータやサービスを自分の目的に活用するコミュニティ。

# DOIとは？

# FAIR Data Principles

<https://www.force11.org/group/fairgroup/fairprinciples>

- **FAIR原則**：データ駆動型サイエンスにおいては、研究データを **Findable, Accessible, Interoperable, Re-usable** な状態にしておくことが重要である。
  1. **Findable**: 識別子、メタデータ、リポジトリ。
  2. **Accessible**: オープンなプロトコル、認証。
  3. **Interoperable**: 共通フォーマット、語彙。
  4. **Re-usable**: ライセンス、来歴、標準。

# DIASにおけるDOI第1号

DIAS データ統合・解析システム

お知らせ

2017年4月5日

データ統合・解析システム (DIAS) におけるDOI付与の開始

ニュース

- トピックス
- イベント
- データセット

10.20783/DIAS.496

～地球環境データへのアクセス向上によりデータの公開と利活用を推進～

DIASは公開データに対するDOI (デジタルオブジェクト識別子) の付与を開始しました。DOIは、論文や研究データなど様々な学術情報への持続的なアクセスを支援する識別子システムとして、世界的に利用が拡大しています。DIASでもDOIを活用したデータ利活用を促進するために、DOIを付与するための体制を整え、こ

<http://www.diasjp.net/infomation/press-release-dias-first-doi-registration/>

DIAS データ検索・検索システム Dataset Search and Discovery

Home How to use About

GAME Tibet

Data File Download with DIAS data download system

IDENTIFICATION INFORMATION

Name	GAME Tibet
DOI	doi:10.20783/DIAS.496
Metadata Identifier	GAME_Tibet:20170725133015-DIAS20170725102541-en

CONTACT

CONTACT on DATASET

Name	Tamakawa Katsunori
Organization	International Centre of Excellence for Water Hazard and Risk Management
Address	6, Minamihara, Tsukuba, Ibaraki, 305-8516, Japan
TEL	028-879-6779
E-mail	tamakawa@icerm.org

CONTACT on PROJECT

Data Integration and Analysis System

Name	DIAS Office
Organization	Remote Sensing Technology Center of Japan
Address	TOKYU REIT Toranomon Building 2F 3-17-1 Toranomon, Minato-ku, Tokyo, 105-0001, Japan
E-mail	dias-office@diasjp.net

[http://search.diasjp.net/en/dataset/GAME\\_Tibet](http://search.diasjp.net/en/dataset/GAME_Tibet)

# 識別子としてのDOI

10.20783/DIAS.496

オブジェクトと文字列の紐付けをグローバルに管理し、オブジェクトへの永続的アクセスを保証する仕組み。

- 単純化すれば、たったこれだけ！
- 単純に見えて実は奥深く、しかも核心的な機能、それが識別子である。



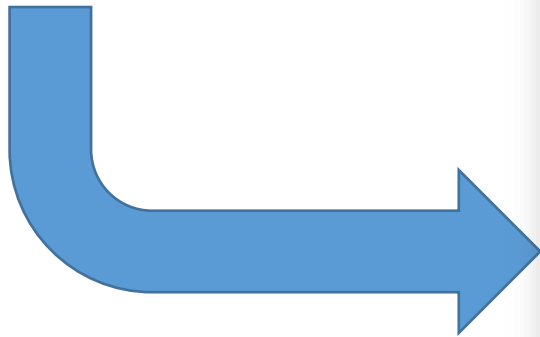
# DOIの仕組み

## DOI = Prefix/Suffix

1. Prefixは国際DOI財団（IDF）が一元管理  
→ グローバルに通用する識別子となる。
2. Suffixは独自に管理 → 意味を持たせる流儀と、持たせない流儀が混在する。
3. レゾルバの運用 → <https://doi.org/<DOI>>  
はランディングページのURLに自動変換（Handleシステムの機能も利用）。

# DOIレゾルバ

<https://doi.org/10.20783/DIAS.496>



**DIAS** データ検索・検索システム  
Dataset Search and Discovery

Home How to use About

### GAME Tibet

Data File Download  
with DIAS data download system

#### IDENTIFICATION INFORMATION

Name	GAME Tibet
DOI	<a href="https://doi.org/10.20783/DIAS.496">doi:10.20783/DIAS.496</a>
Metadata Identifier	GAME_Tibet:20170725133015-DIAS20170725102541-en

#### CONTACT

##### CONTACT on DATASET

Name	Tanaka Katsunori
Organization	International Centre of Excellence for Water Hazard and Risk Management
Address	1-6, Minamihara, Tsukuba, Ibaraki, 305-8516, Japan
TEL	029-879-6779
E-mail	<a href="mailto:tanakawa@icerm.org">tanakawa@icerm.org</a>

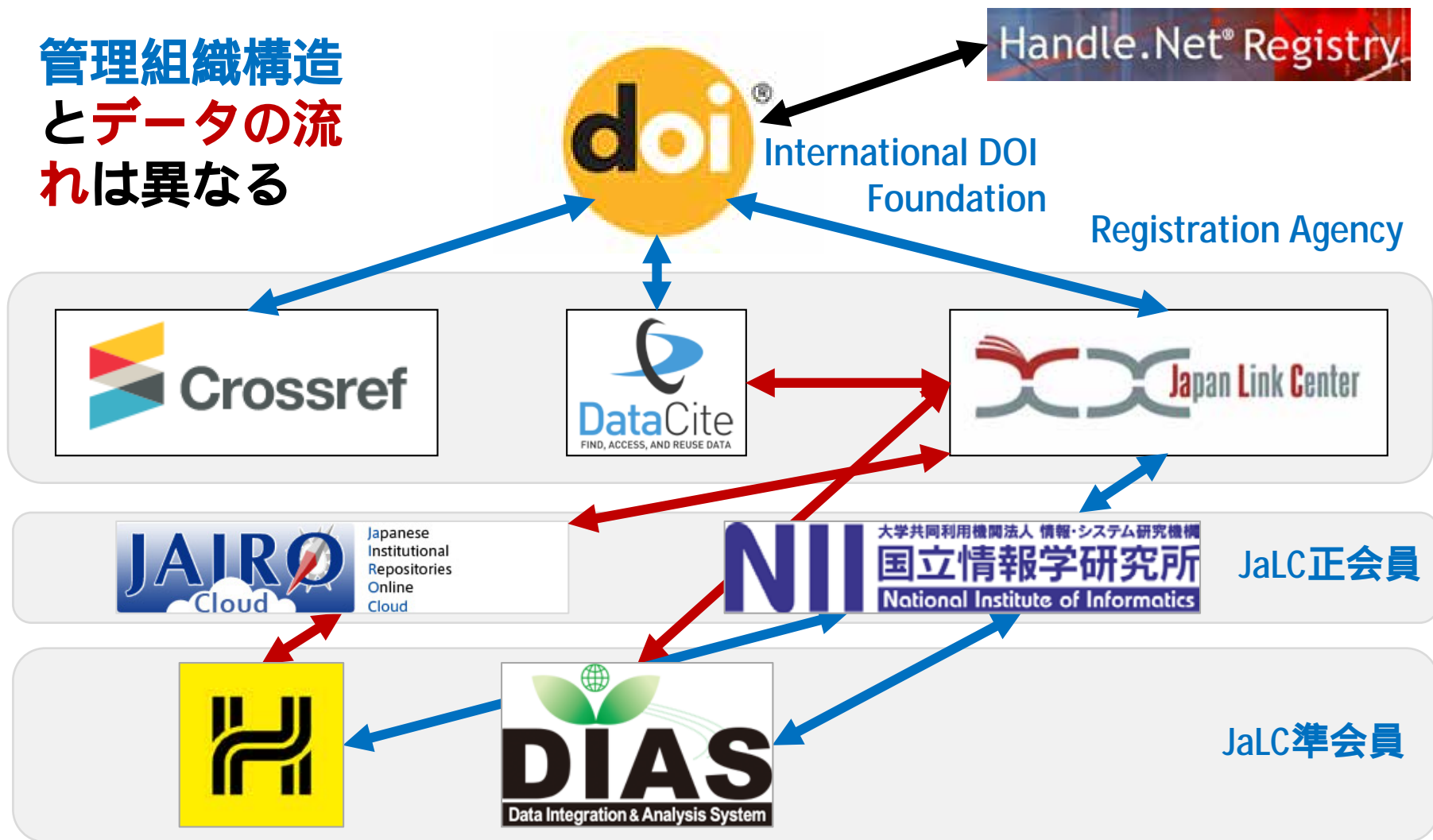
##### CONTACT on PROJECT

Data Integration and Analysis System

Name	DIAS Office
Organization	Remote Sensing Technology Center of Japan
Address	TOKYU REIT Toranomon Building 2F 3-17-1 Toranomon, Minato-ku, Tokyo, 105-0001, Japan
E-mail	<a href="mailto:diass-office@diass.jp">diass-office@diass.jp</a>

# DOIシステムの全体像

管理組織構造  
とデータの流  
れは異なる

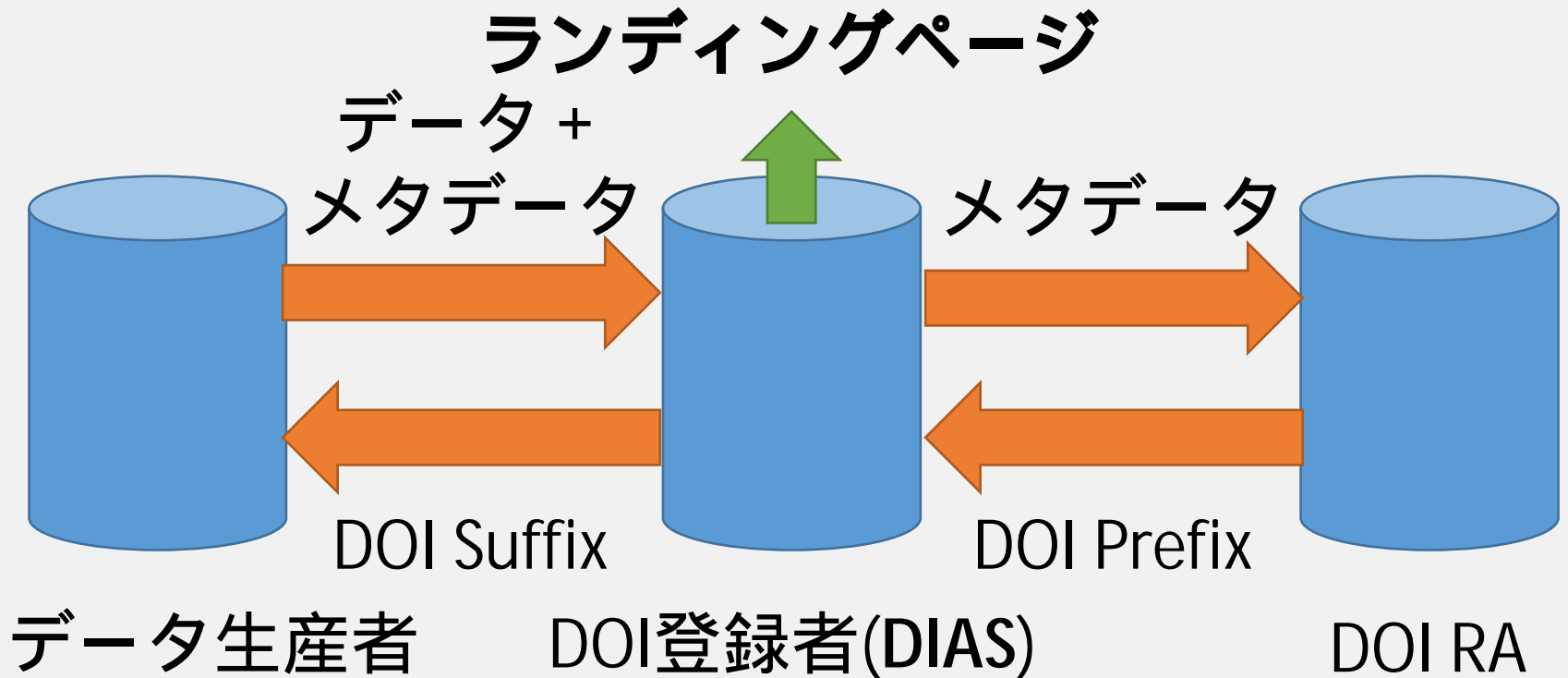


# DataCiteへの登録

<https://search.datacite.org/works/10.20783/DIAS.496#>

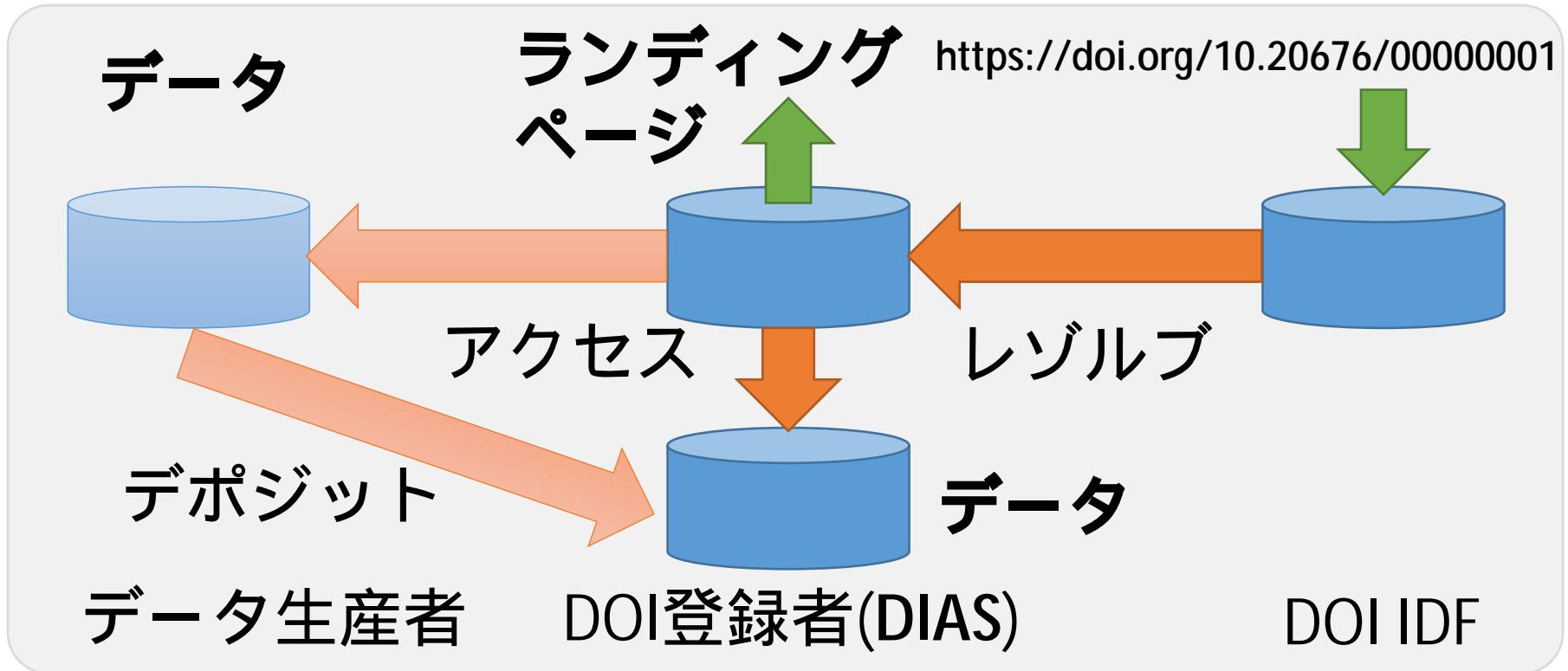
The screenshot shows the DataCite search interface. At the top, there is a navigation bar with links for 'Works', 'People', 'Data Centers', 'Members', 'Support', and a 'Sign in' button. The main content area features a search result for 'GAME Tibet' by Koike Toshio. The title 'GAME Tibet' is highlighted in blue. Below the title, it states 'Work published 2017 via Data Integration and Analysis System (DIAS)'. The description of the project is provided, and the DOI link <https://doi.org/10.20783/dias.496> is shown with a 'Cite' button. To the right of the main content, there are sections for 'Data Center' (Japan Link Center), 'Member' (Japan Link Center), and 'Download' options (DataCite XML, RDF-XML, Schema.org JSON-LD, Citeproc JSON). Below these are social media share icons for Twitter and Facebook. The footer contains four columns of links: 'About DataCite' (What we do, Board, Steering groups, Staff, Job opportunities), 'Services' (Assign DOIs, Metadata search, Event data, Profiles, re3data, Citation formatter, Statistics, Content negotiation, OAI-PMH, Test environment), 'Resources' (Metadata schema, Support), 'Community' (Members, Partners, Steering groups, Events, Roadmap, User Stories), and 'Contact us' (social media icons, Terms and conditions, Privacy policy, Acknowledgements, and a status indicator 'All Systems Operational').

# メタデータとDOIの交換



DOIをGiveする代わりにメタデータをTakeすると、  
**DOIチェーンの上流にメタデータが集まる。**

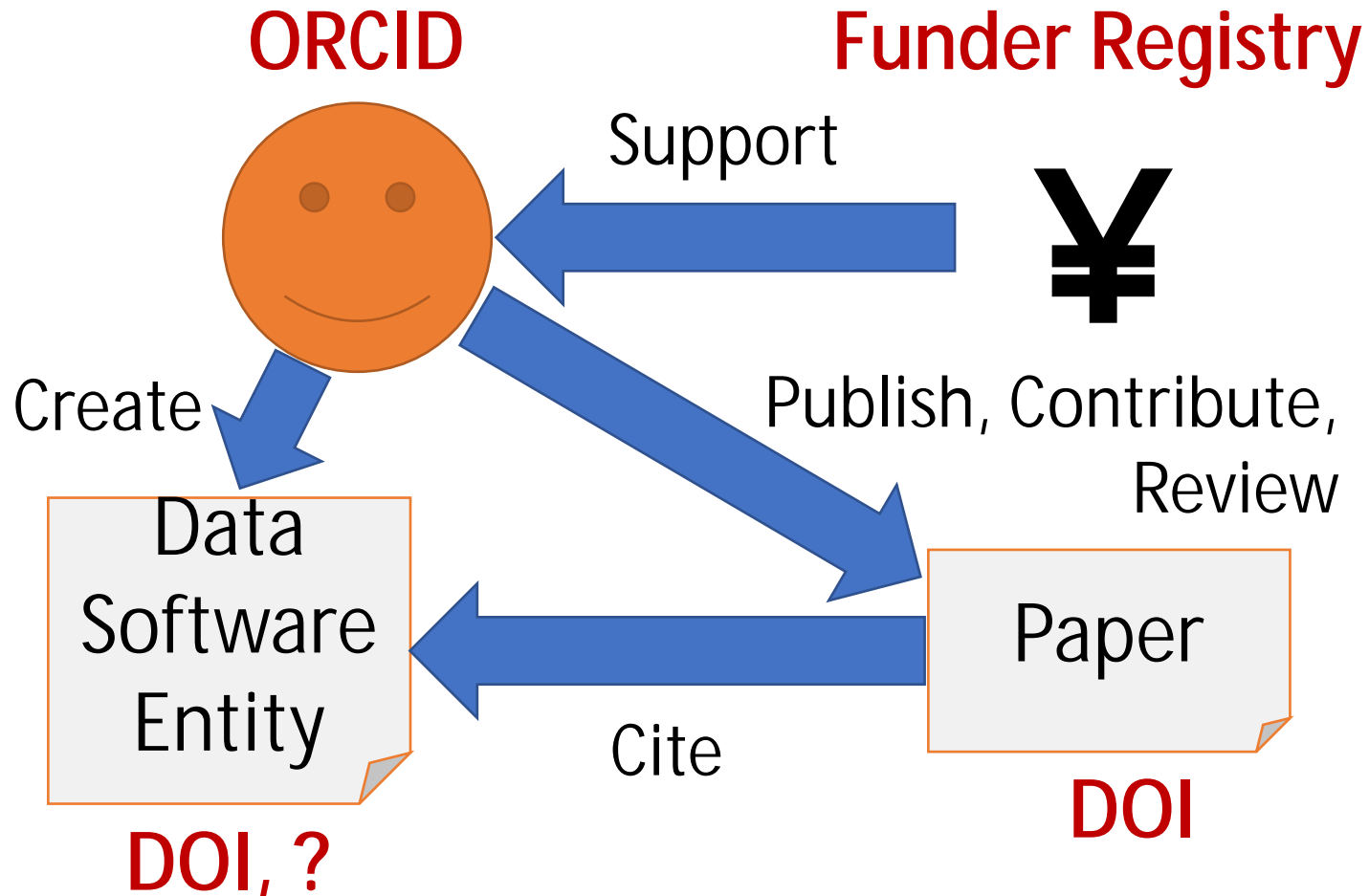
# ランディングページと責務



ランディングページには、**オブジェクトのメタデータとアクセス手段**を明示。

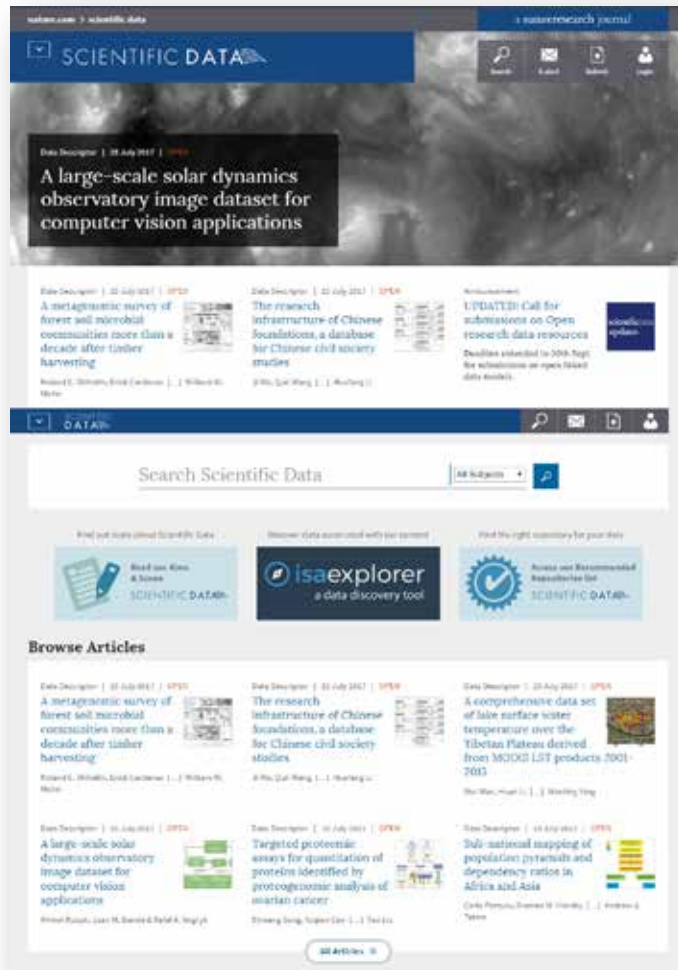
# DOIの利活用

# グローバルな識別子の普及





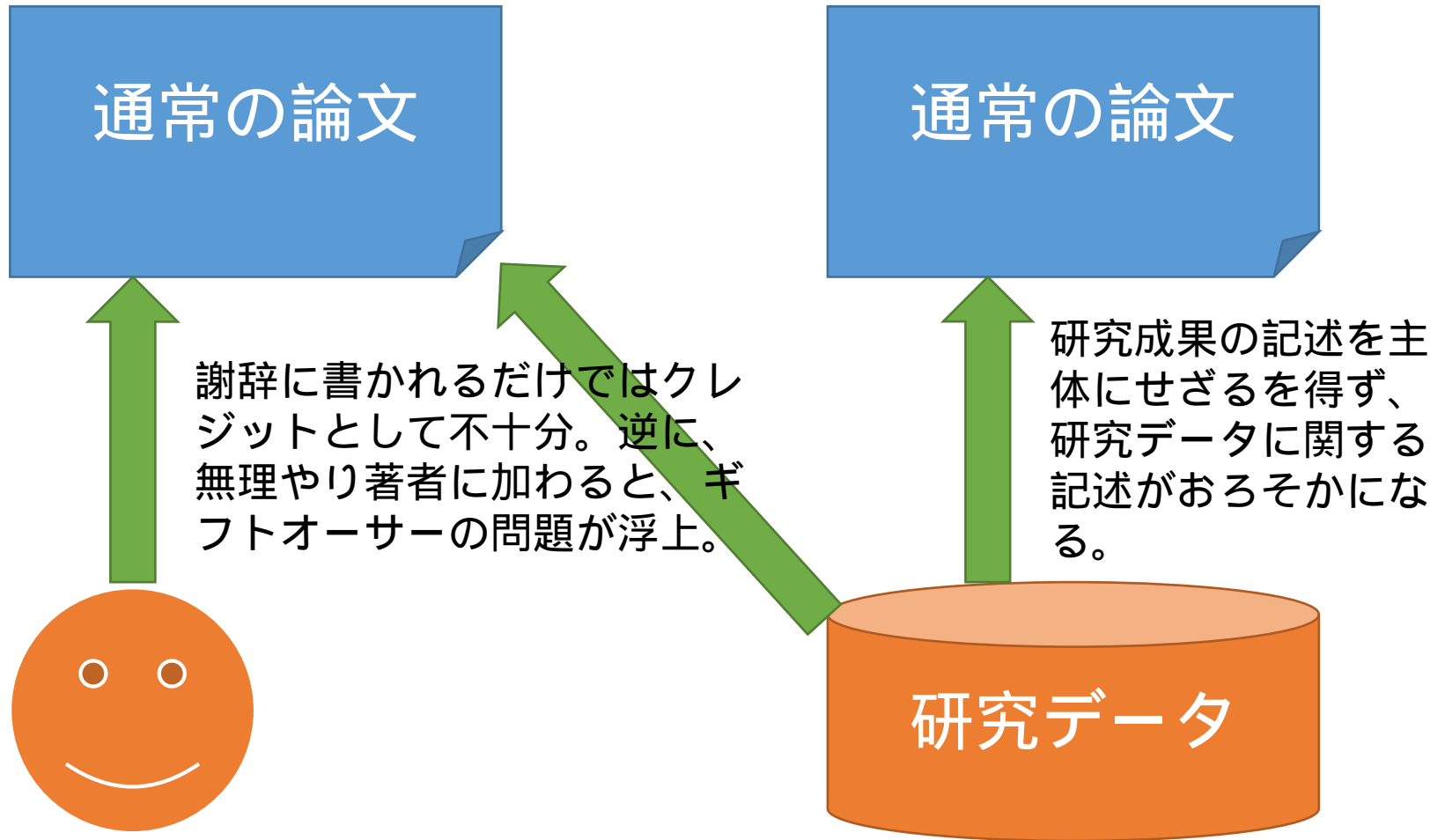
# 研究データの重視



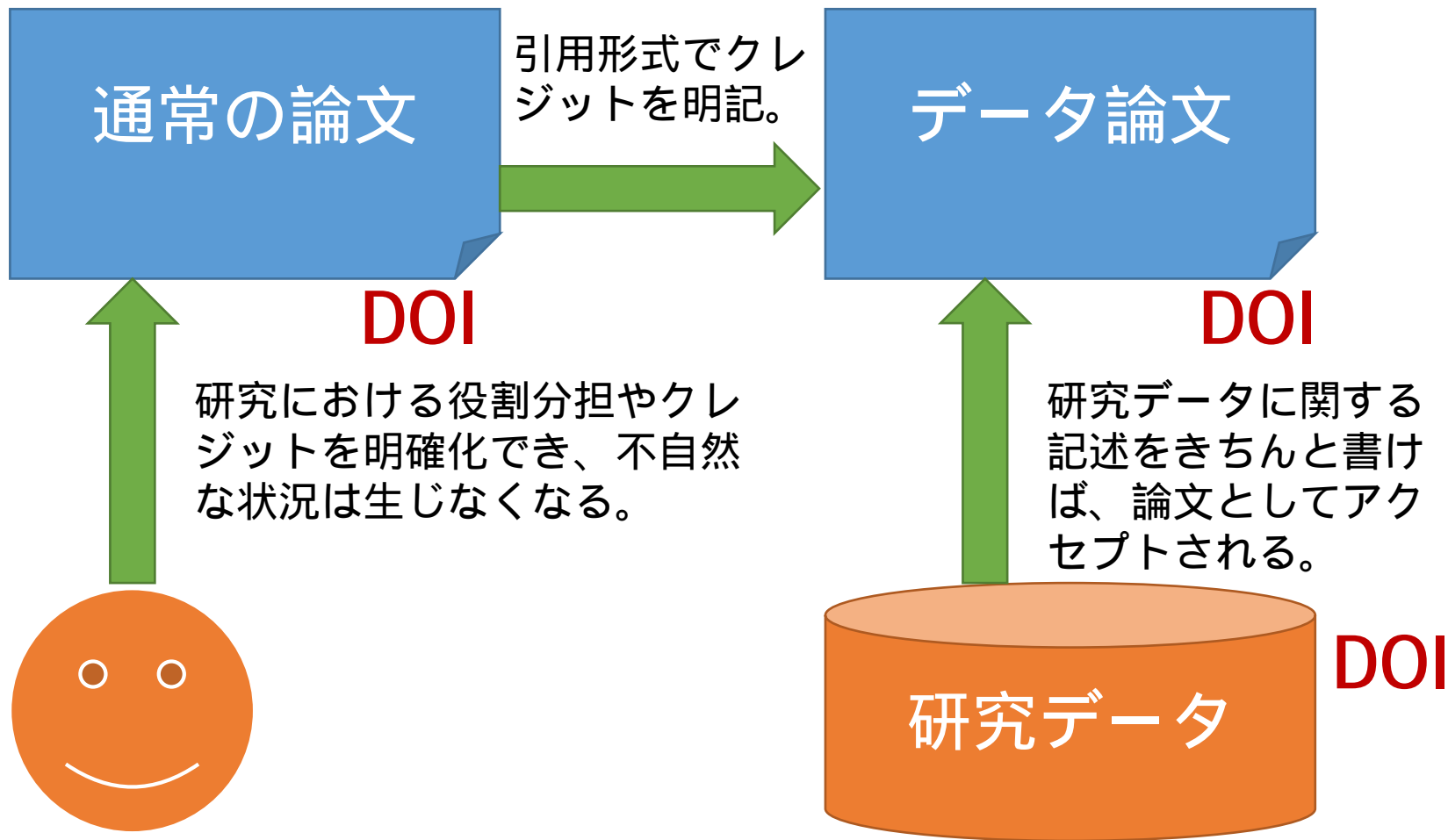
<https://www.nature.com/sdata/>

- 研究論文の根拠となるデータも、紐付けてオープン化すべき。
- 研究資源となるデータは、FAIRデータとして管理すべき。
- データ論文により、きちんと引用して評価すべき。

# 従来の学術出版



# データ論文を含む学術出版



# データジャーナル投稿手順

1. データジャーナルごとに投稿規程が存在し、**データをデポジットできるリポジトリの条件**が示されている。
2. 有名なりポジトリはすでに条件を満たしているが、**DIASが条件を満たしているか**は確認が必要。
3. ジャーナル編集者などと相談しながら、DIASがその条件を満たすように解決策を探る。

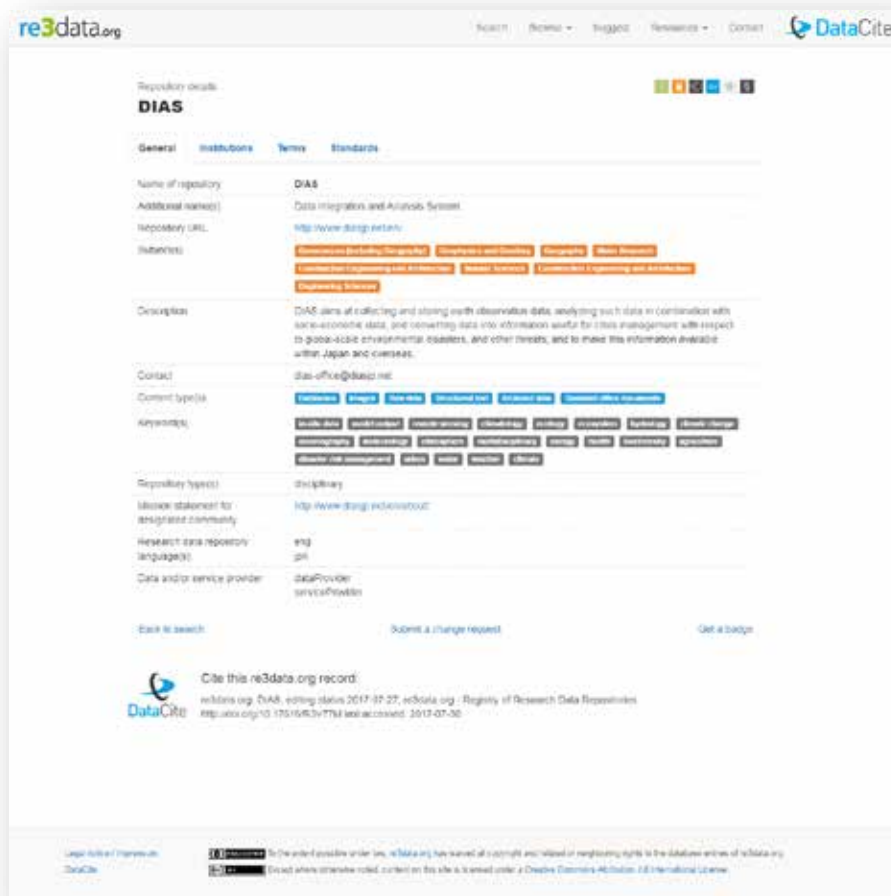
# Earth System Science Data Journal

<https://www.earth-system-science-data.net/>

1. **Persistent identifier:** The data sets have to have a digital object identifier (DOI).
2. **Open access:** The data sets have to be available free of charge and without any barriers except a usual registration to get a login free-of-charge.
3. **Liberal copyright:** Anyone must be free to copy, distribute, transmit, and adapt the data sets as long as he/she gives credit to the original authors (equivalent to the Creative Commons Attribution License).
4. **Long-term availability:** The repository has to meet the highest standards to guarantee long-term availability of the data sets and permanent access.

# 研究データリポジトリの登録

<https://doi.org/10.17616/R3V77M>



- 研究データリポジトリの概要をメタデータ化して登録。
- リポジトリの信頼性の判断を助ける情報を提供（認証は別）。
- データジャーナル等への投稿には、所定の条件を満たすリポジトリの利用が必須。

Re3data: Registry of Research Data Repositories

# リポジトリの種類

## 汎用リポジトリ

データの種類や由来を問わず、広くデポジットを受け付ける。



## 機関リポジトリ

機関が生み出した成果物を、一般的な標準を適用し、広く発信する。



## 分野リポジトリ

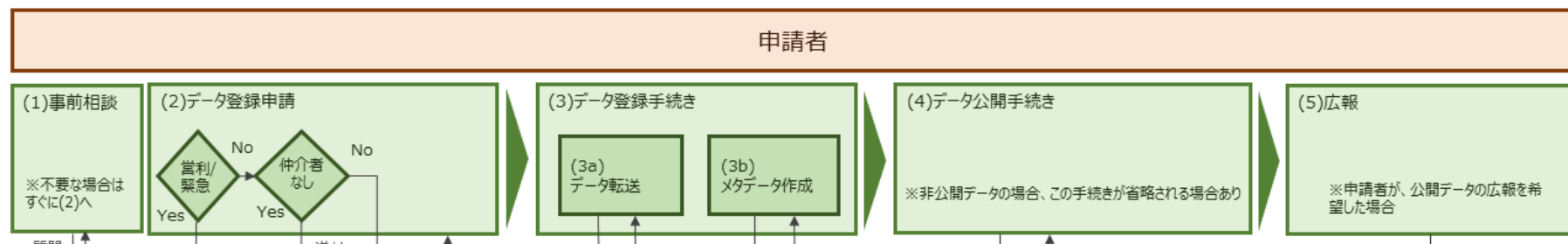
分野のコミュニティが、自らの標準を適用し、成果を共有する。



# DIASにおけるDOI付与 の拡大



# DIASへのデータ登録



- DIASは「分野リポジトリ」であり、どんなデータでも受け入れるわけではない。
- DIASのアーカイブ対象となるデータかを判断するために、登録申請が必要である。
- データ出版だけを目的としたデポジット先としては使えない。

# 既存データセットへのDOI付与

1. まずDIASアーカイブズ済み**既存データセット**を対象に、DOI活用事例を探る。
2. DOIの利用は今後確実に増えるため、**できるだけ早期にDOIを付与**しておきたい。
3. DIASプロバイダコミュニティへ方針を通知し、DOI付与を具体化したい。
4. DOI付与に伴う**デメリット**は**基本的にな**いが、一つだけ考慮すべき点がある。

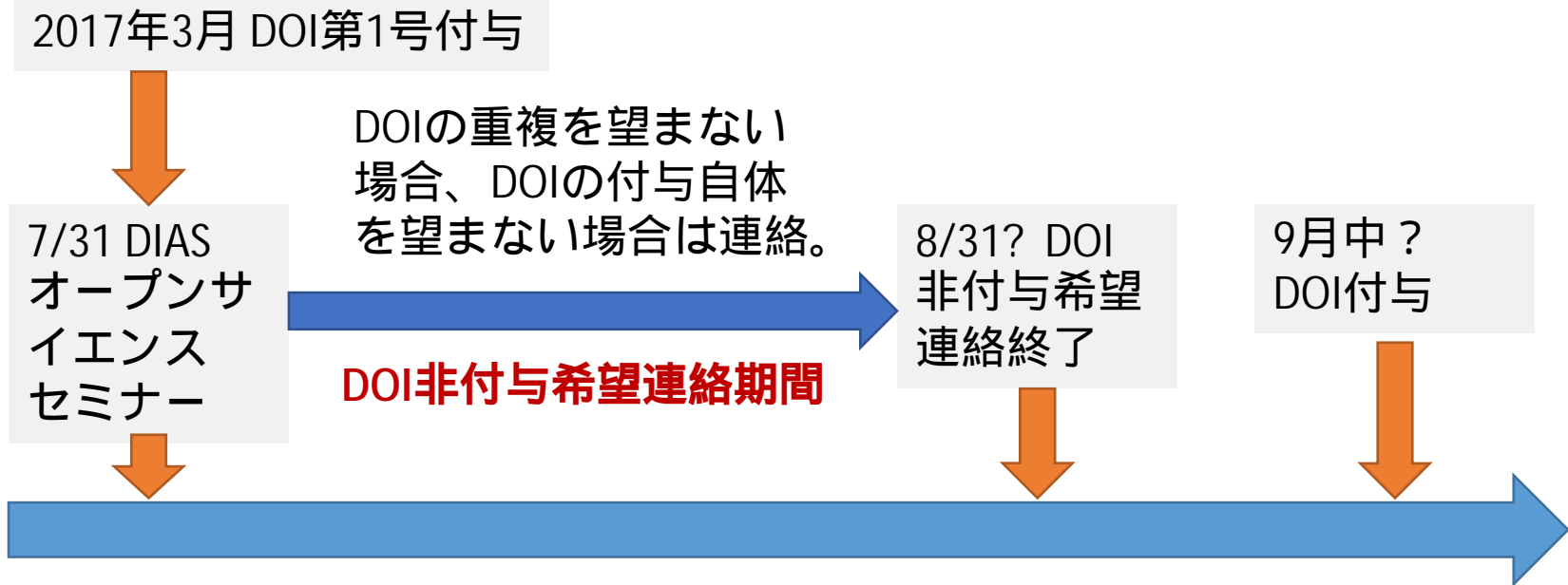
# DOIの重複付与問題

1. DOIはデータセットを特定する**グローバルに唯一な識別子**であって欲しい。
2. **努力目標**：DOIの**重複付与を防ぐ**。ただしそれを防止する技術的方法はない。
3. **複数DOI付与**：DOIの**分散**が発生し、**同一性に関する混乱**の心配がある。
4. DOIの**分散**：引用数などの点でも**不利** = 将来的には別のサービスで解決？

# DOI付与対象

1. DIASにメタデータとデータを登録済みのデータセットをDOIの付与対象とする。
2. 逆にメタデータしか存在しない場合はDOIを付与しない。
3. 一部の大規模データ（温暖化予測データ等）は、進め方を関係者と協議する。
4. 現在の登録データの粒度に合わせて、DOIを付与する。

# DOI付与スケジュール（仮）



DOIの重複付与を、DIAS側で調査することは困難 **オプトアウト式**（連絡がなければ問題なしとみなす方式）で付与。

# ご意見

1. DOI非付与希望連絡期間が短すぎるのではないか？夏休みで連絡が取りづらい。
2. 登録アドレスにメールが届かない場合はどうするか？
3. いったん認めた後に、後から取り下げたくなった場合はどうするか？
4. もし登録データの粒度を変えたいとなったらどうするか？

# DIASプロバイダコミュニティ

- 関係者のコミュニケーションを増やすことが、オープンサイエンスの精神に沿う。
  1. **ワークフローの改善**：より使いやすいシステムに向けたロードマップ。
  2. **ライセンスの設定**：オープンデータ時代に合わせたライセンス種類と表記。
  3. **利用促進方策**：サイテーション機能、データジャーナル対応、検索高度化等。

# オープンサイエンスの展望

1. DIASは、データ駆動型科学の発展に不可欠な**研究データ基盤**を構築する。
2. DIASは、地球環境データへのアクセスを適切な条件の下で**オープン化**する。
3. DIASは、データ提供者が高く**評価**されるような文化を作り出す。
4. DIASは、研究データ基盤を支える**人材**の育成にも取り組む。



# DIASウェブサイト

日本語 | English



データ統合・解析システム

ホーム

DIASとは

データ・アプリケーション

適用分野

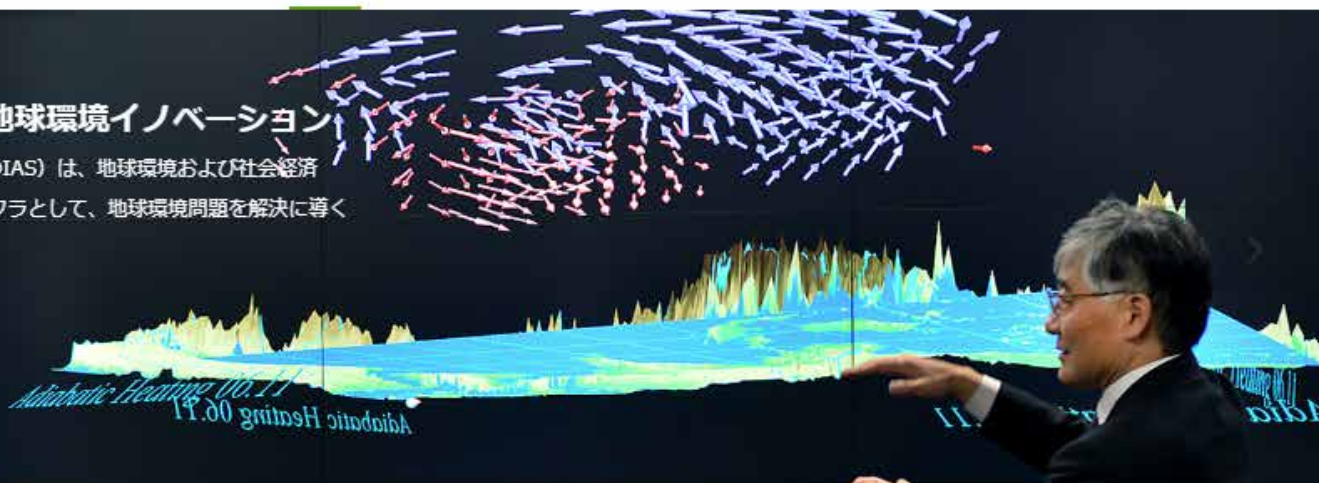
活動紹介

お知らせ

## 最先端技術による地球環境イノベーション

データ統合・解析システム (DIAS) は、地球環境および社会経済データを収集・解析するインフラとして、地球環境問題を解決に導く知と公共利益を創出します。

DIASとは



<http://www.diasjp.net/>

## 適用分野



気象・気候



水



都市



防災



農業



生物多様性



健康



エネルギー