ご挨拶

- 研究データに関する世界的な国際会議である研究データ同盟（Research Data Alliance：RDA）東京大会がこのたび開催されることとなりました。これを機会に、研究データ共有に関する国内外の最新動向を広く共有するためのフォーラムを開催します。
- 日本でも研究データに関する動きが最近は加速しつつあります。例えば、ジャパンリンクセンターでは研究データへのDOI登録実験プロジェクトが進み、研究データの共有を目指す新たなコミュニティが広がりつつあります。また、国もオープンサイエンスの推進を前面に打ち出しており、研究活動や図書館活動等における仕事の進め方にインパクトを与えることが予想されます。
- こうした潮流をどのように捉えればよいでしょうか？RDA東京大会における議論や国内外の動向を踏まえながら、研究者や図書館員をはじめ様々な立場の人々が共に考えていける場にしたいと考えています。基礎的知識に関する講演もありますので、研究データに興味はあるものの敷居が高いと感じている方も、どうぞお気軽にご参加ください。

フォーラムの構成

1. **講演**
   - 3名が、基礎的知識やコミュニティ形成などについて、コンパクトにまとめる。
2. **RDA東京大会参加報告**
   - 7名が、RDAで感じた、学んだことをもとに、それぞれの視点から意見を述べる。
3. **フロアも交えたディスカッション**
   - みんなで、研究データとオープンサイエンスについて考える。

研究データとオープンサイエンスに関する基礎的知識

国立情報学研究所 / 総合研究大学院大学
北本 朝展（KITAMOTO Asanobu）
http://agora.ex.nii.ac.jp/~kitamoto/

自己紹介

- 研究分野は情報学。特に画像処理や画像データベース、パターン認識など。
- その後、データ駆動科学へ。特に気象情報、地球環境情報、災害情報、人文科学情報等の分野。
- **高品質データ基盤の構築や、異分野融合研究への評価を、もっと高めたいという立場。**

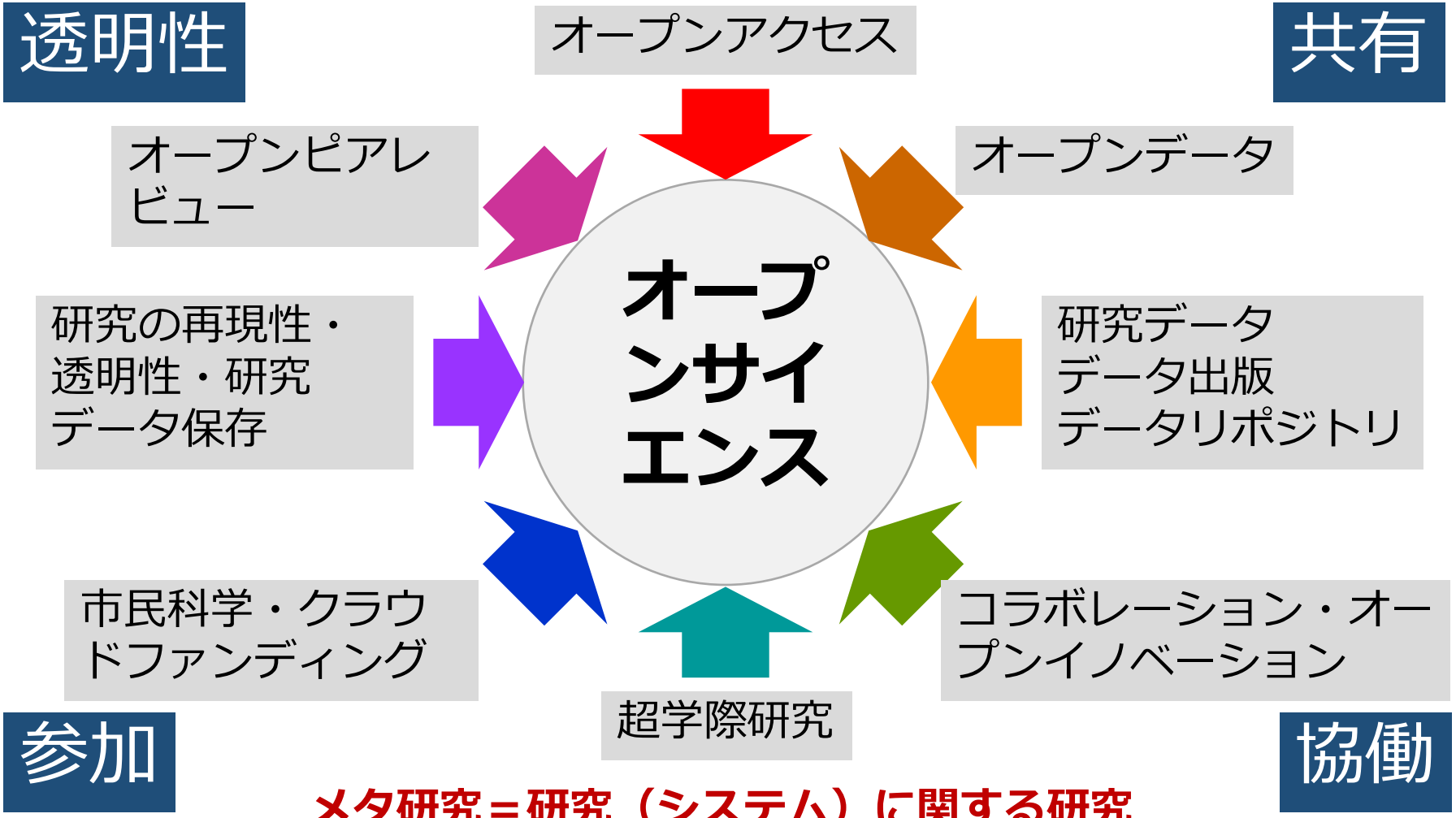
デジタル台風

311メモリーズ

ディジタル・シルクロード

オープンサイエンス

オープンサイエンスへの収束



メタ研究＝研究（システム）に関する研究

オープンサイエンスとは？

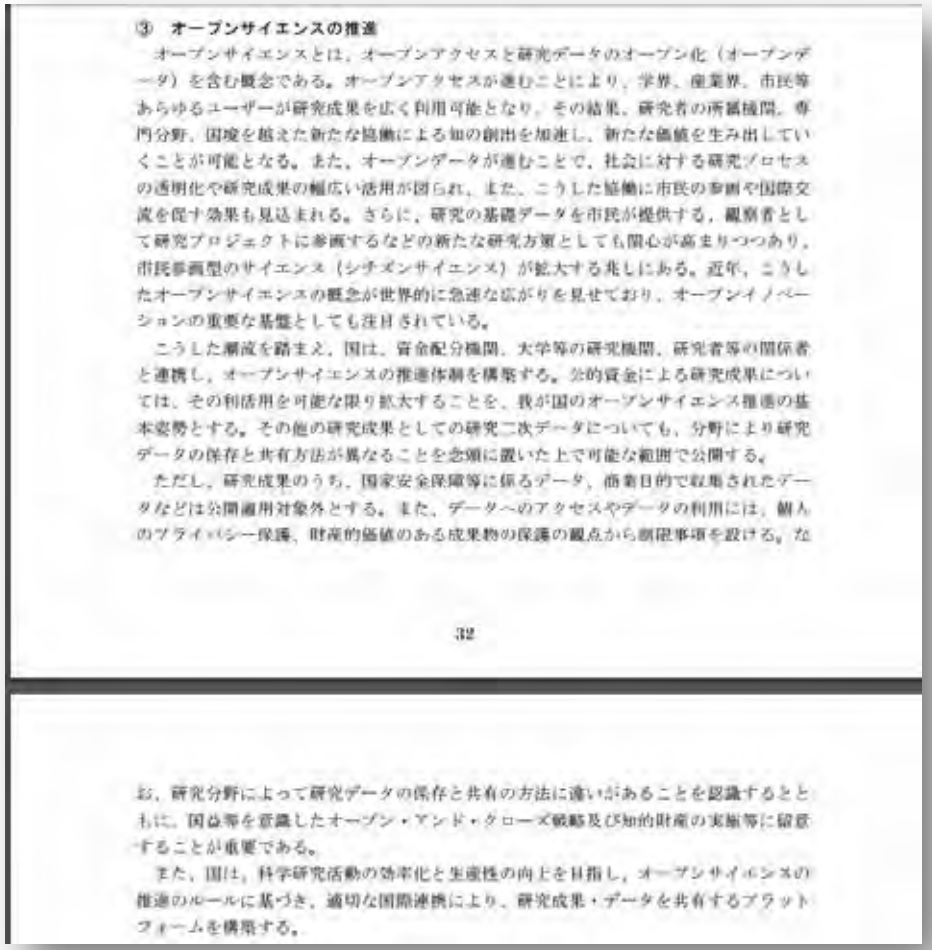
- **「オープン」という言葉に夢／野望を託し、現状の研究システムを変革すること。**
- 共通目標は、社会との関わりの中で、サイエンス（研究）をもっと良くすること。
- 大同団結？同床異夢？コミュニケーションしながら、それぞれの立場を理解する。
- 分野の壁、専門性の壁を越えて、現在の問題点を共有し、全体最適化を図る。

- 科学はオープン化を妨げる**独自の問題**に直面している。
- （CCなどのツールは）**科学者の報酬は論文の出版によって得られる**もの（中略）という問題を直接解決してくれるわけではない。
- オープンサイエンスは（中略）一般的な運動から多くを学べるが、それと同時に**新たな考え方を必要**としている。

オープンサイエンス革命
マイケル・ニールセン (著)
紀伊國屋書店, 2013
右の引用はp.340.

ニールセンの「夢」を実現する方向に進んでいるわけでは必ずしもない。

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。

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

- 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が記述の中心。
- 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触れる。

EU Open Science

https://ec.europa.eu/research/consultations/science-2.0/science_2_0_final_report.pdf

Science 2.0からOpen Scienceへの変遷

- 「Science 2.0」の代替として、43%が「Open Science」と答えた。
- ‘participatory science’, ‘science highway’, ‘better science’, ‘open research’ ‘open scholarship’等も候補に挙がる。
- 「サイエンス」は、人文学を含むものと解釈。

オープンサイエンスに対する意見

- 研究者の関心は研究成果の普及や共同研究にあり、市民科学への関心は薄め。
- 懸念には、品質の保証、クレジット付与の欠如、（若手への）インセンティブの欠如などがある。

「オープン」が意味する範囲

オープンデータ	**狭すぎる**
オープンソース	**ぴったり**
オープンサイエンス	**広すぎる**

オープンサイエンスは範囲が広く、境界を定義しきれない。むしろ、動きの方向「よりオープンに」を共通項とした方法論の束ではないか。

研究データ

研究データの多彩なテーマ

1. データシェアリング（オープンデータ）
2. 永続的識別子、メタデータ、キュレーション
3. データ出版、データ論文、データジャーナル
4. データリポジトリ、相互運用性、ライセンス
5. データサイテーション、評価メトリクス
6. 研究データ／リポジトリの評価・認証
7. 研究データのワークフロー、ライフサイクル
8. 人材育成、ビジネスモデル
9. データ駆動科学（＝研究データを活用した個別分野の研究）

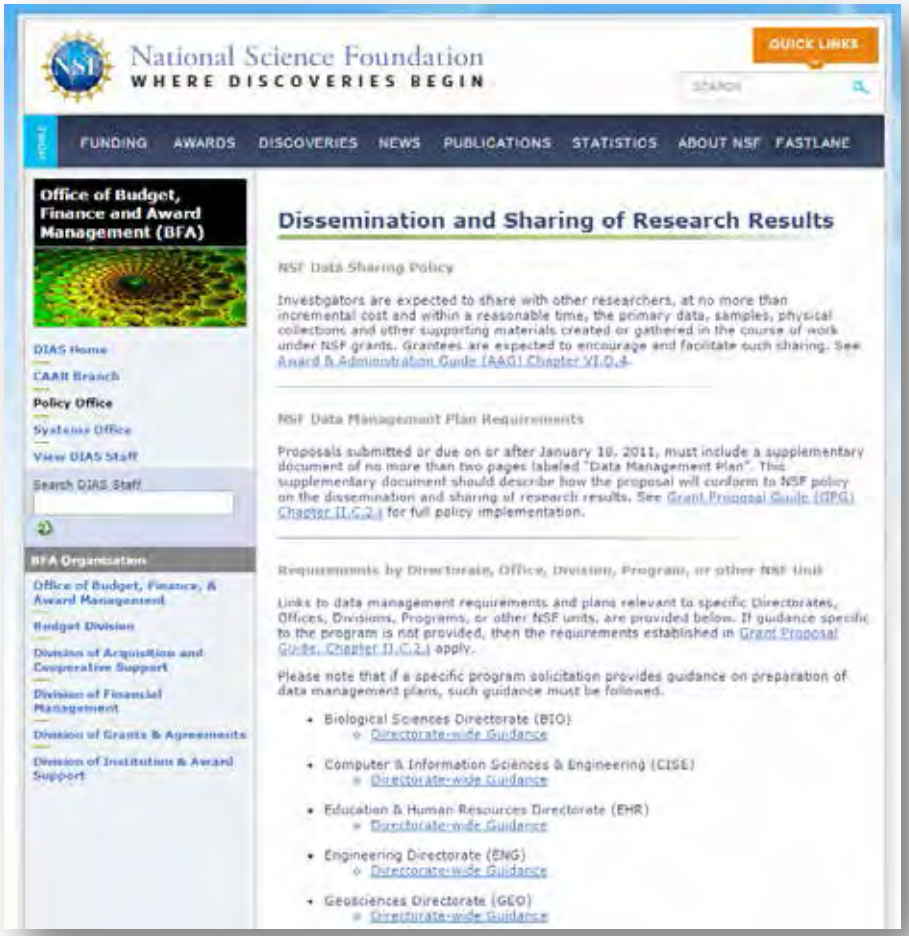
オープン・クローズ戦略

政府	研究	ビジネス
公共の利益に資すること（**open by default**）	新しい方法を生み出すこと（**open by strategy**）	組織の利益を上げること（**closed by default**）

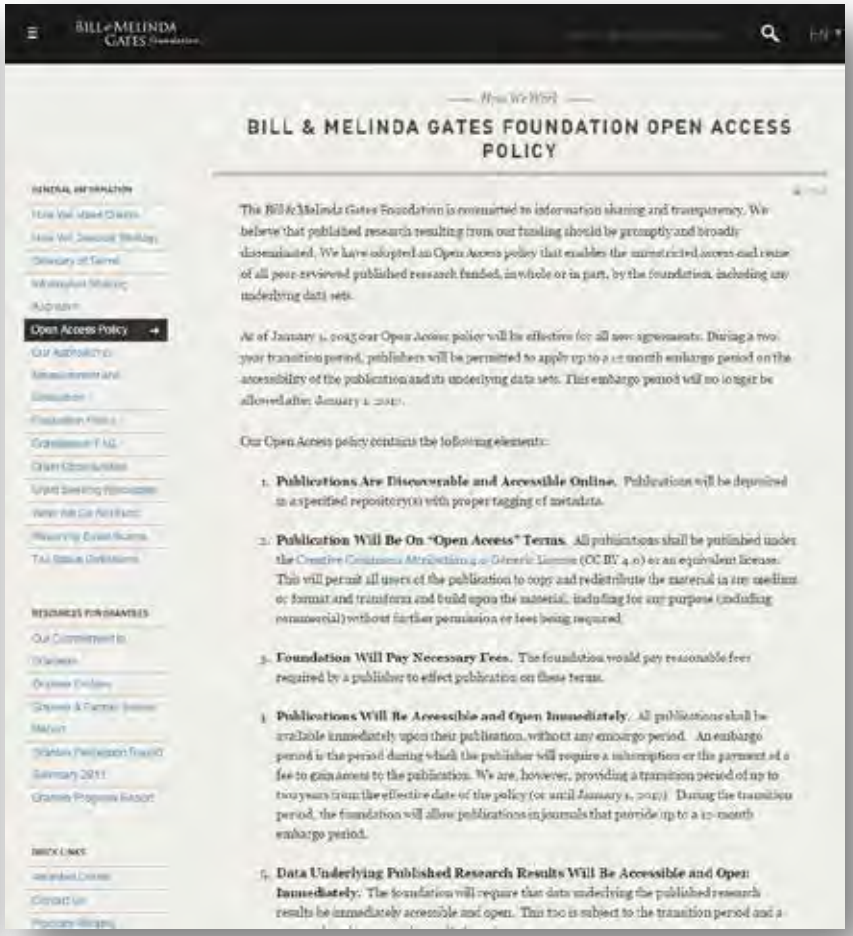
様々なオープン化

- データ・シェアリング：グループ内や研究コミュニティ内でデータを共有。
- （広義の）オープン化：市民も含めた社会からのアクセスと活用を促進する。
- （狭義の）オープンデータ：誰でも制限なく、機械処理が容易な形式で公開する。
- オープン化の一つの根拠は研究資金のポリシー。公的助成・基金の使命に基づく。

研究助成機関の動き



NSF Data Management Plan



Open Access Policy

京都大学オープンアクセス方針

- 研究者間でコラボレーションを促進し、研究分野を超えた新たな知を創出。
- 学術研究に従事する者が社会に対して果たすべき説明責任を明確にする。
- 京都大学は、全学方針の採択により、大学としてオープンアクセスを推進する姿勢を明らかにする。

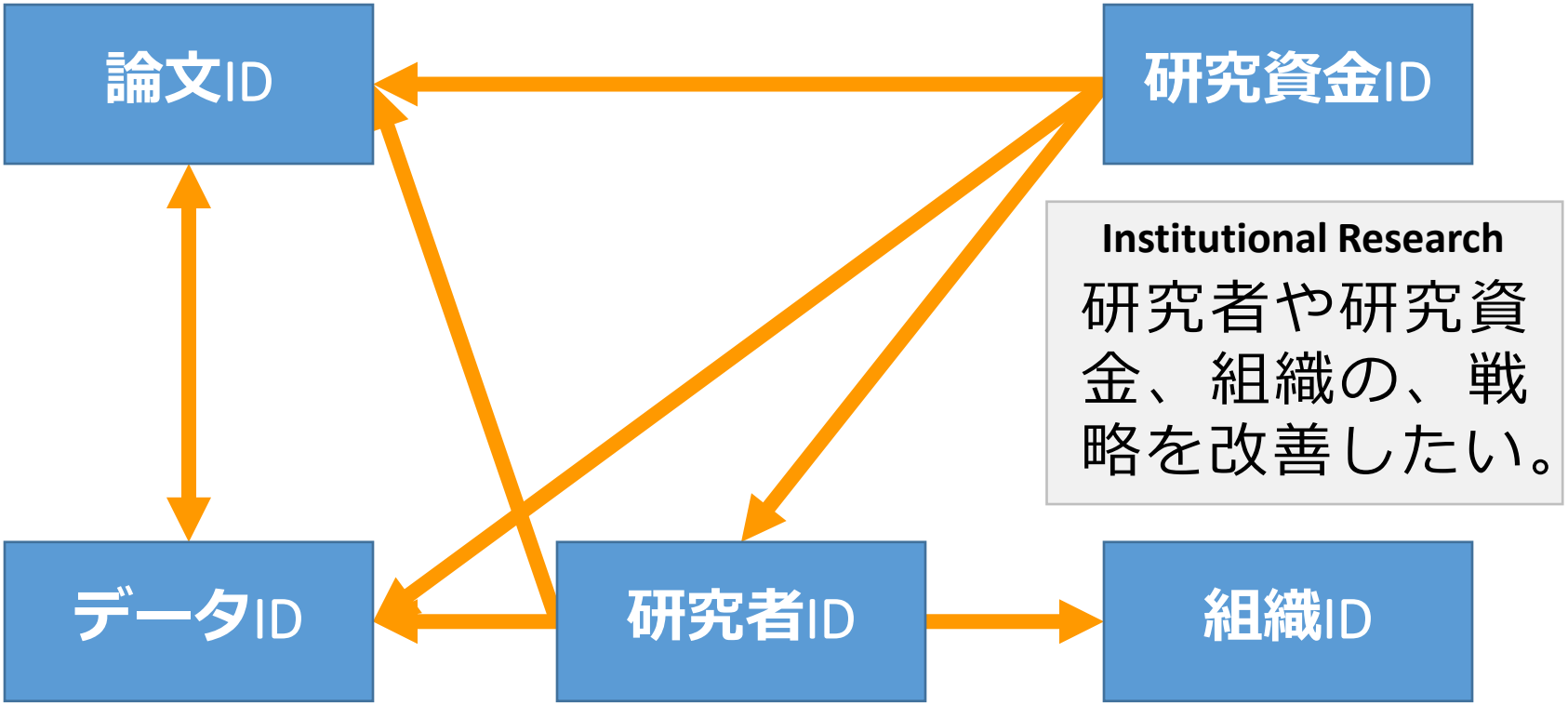
http://www.kulib.kyoto-u.ac.jp/modules/content0/index.php?content_id=92

研究データリポジトリ

https://www.datacite.org/node

- DataCite（2009～）が世界の研究データの事実上のハブ。
- WDS等の学術的、Figshare等の商業的なリポジトリが普及。
- 日本は存在感が薄い。JaLCや機関リポジトリの活用で挽回？

研究データと「何でもID」



研究成果は、永続的識別子（PID）をリンクしたグラフとして、表現かつ計測できる。

インセンティブ

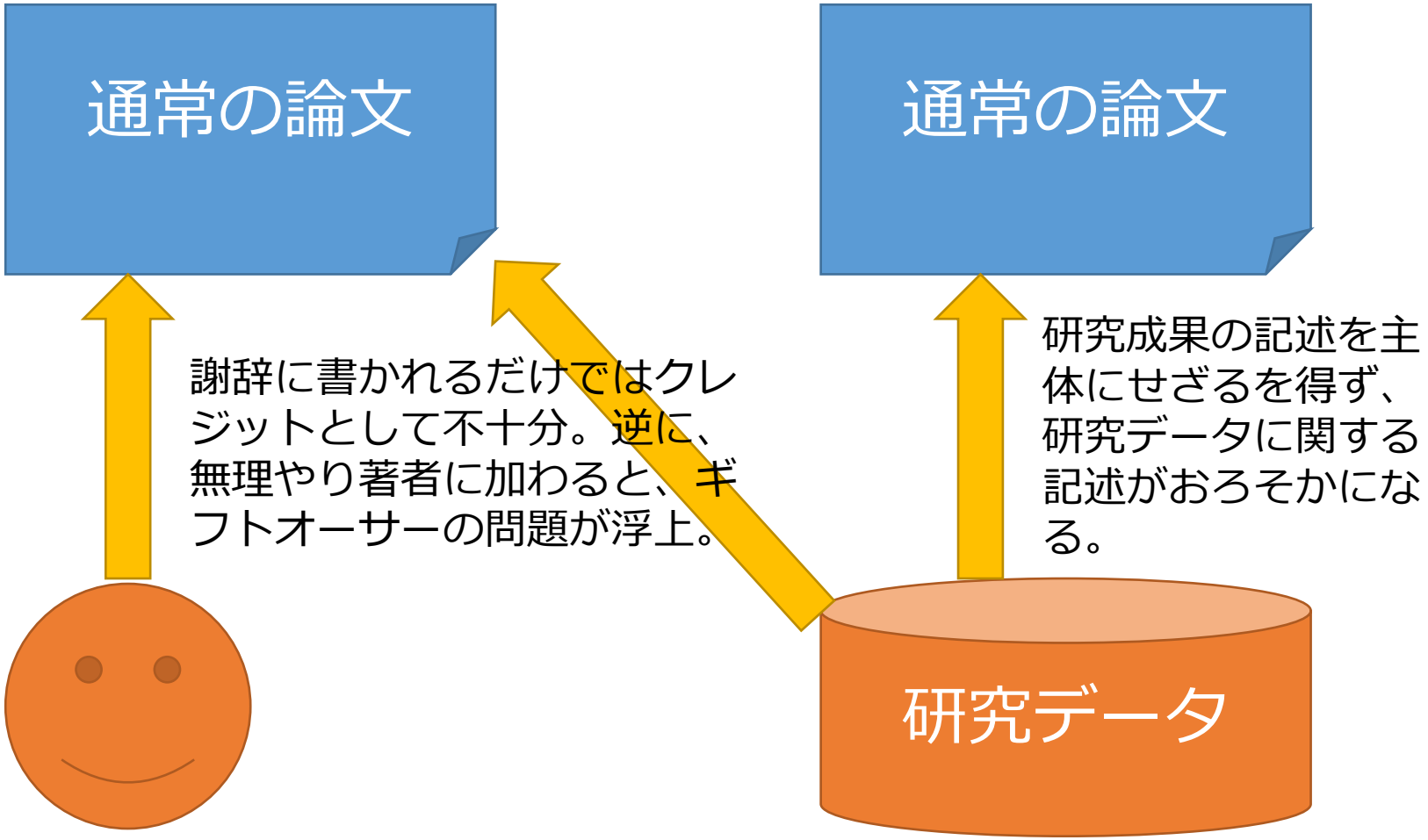
現場からの声

- 外からあれこれ言うのは勝手だけど、現場のことを考えてるの？
- 研究者：オープンサイエンスは、キャリアパスとして有利なの？
- 図書館員：オープンサイエンスは、図書館が今後やるべき仕事なの？
- 外からの押しつけは失敗する。インセンティブをきちんと設計する必要あり。

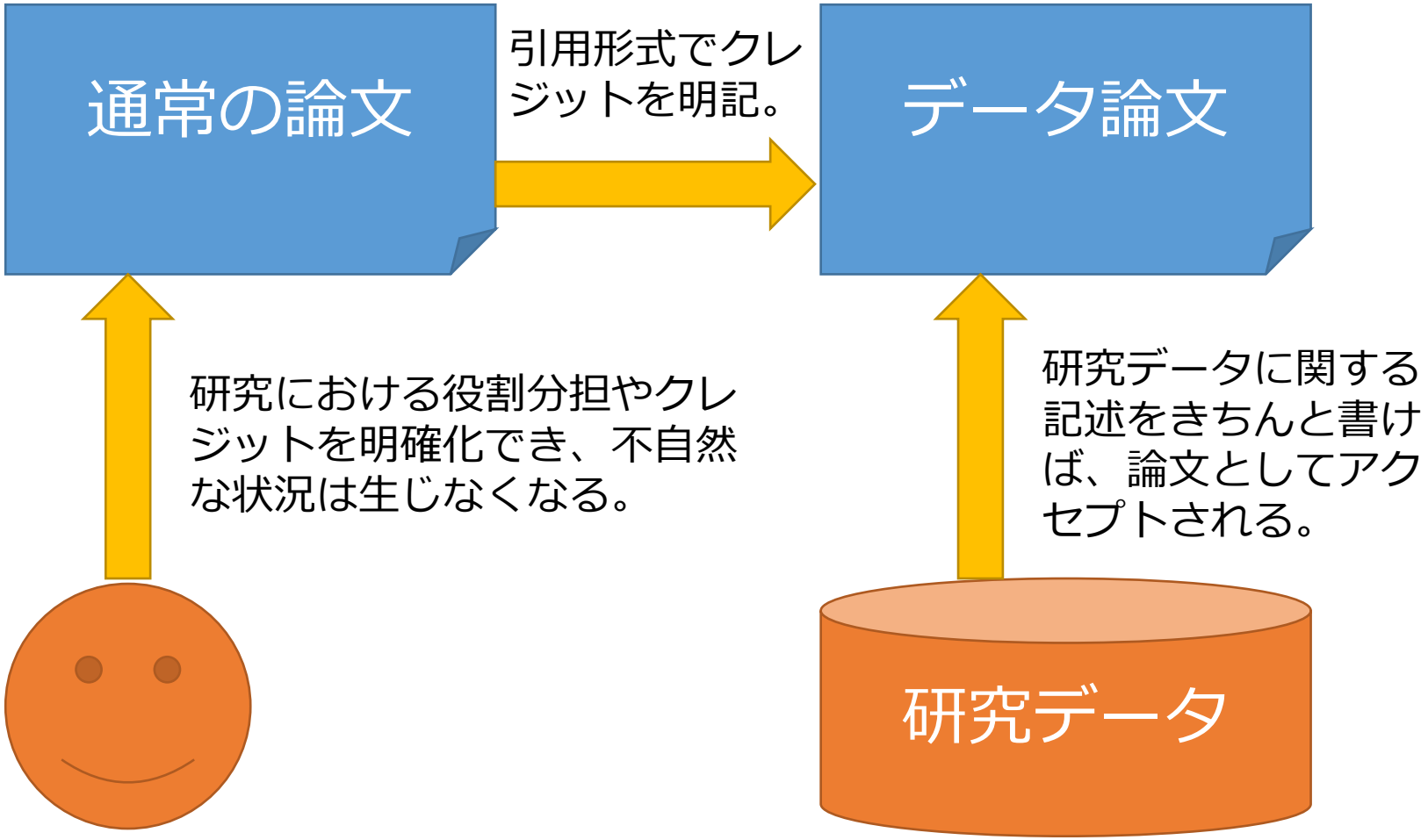
データジャーナルの査読

- Scientific Dataの例では基準は以下の通り。
  1. Experimental Rigour and Technical Data Quality
  2. Completeness of the Description
  3. Integrity of the Data Files and Repository Record
- データから生まれる発見の重要性や新規性は問わない。著者によるデータの解釈も不要。
- データ論文は、データの生成方法、品質管理、データ生成プロセスの厳密さなどを査読し、きちんとしたデータの再利用を促進するもの。

従来の学術出版


通常の論文
通常の論文
謝辞に書かれるだけではクレジットとして不十分。逆に、無理やり著者に加わると、ギフトオーサーの問題が浮上。
研究成果の記述を主体にせざるを得ず、研究データに関する記述がおろそかになる。
研究データ

データ論文を含む学術出版



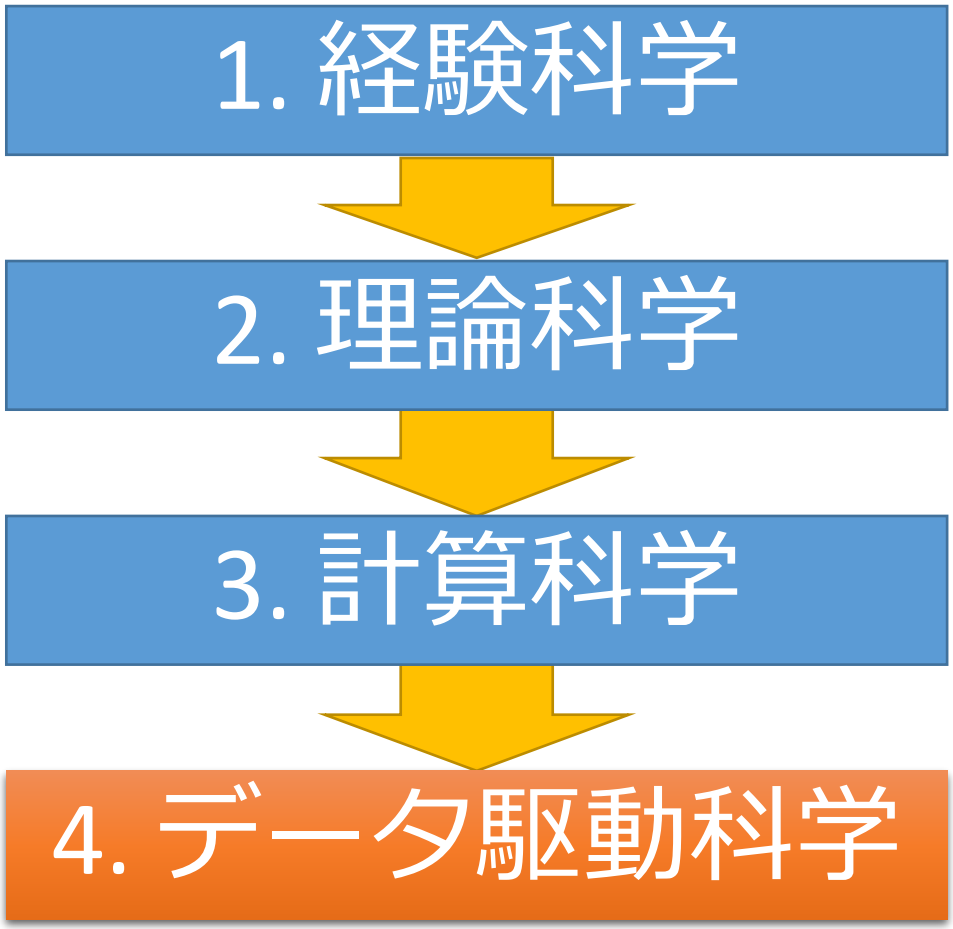
研究貢献バッジ

https://badges.mozillascience.org/

- 研究への貢献の評価を多様化・明確化することで、関係者全員に適正なクレジットを付与。

今後の課題

データ駆動科学



- センサ：ネット経由で実世界からデータ収集できる。
- 情報基盤：データを処理・管理・保存・共有できる。
- 機械学習：ビッグデータを一般化・転用できる。

研究データの重要課題

- インセンティブ：データ駆動科学基盤の構築に取り組むインセンティブを与える。
- キャリアパス：従来型の論文（＋査読）以外の評価システムで人材を育成する。
- キュレーション：データ出版のワークフローに図書館員のスキルを活用する。
- コラボレーション：データのオープン化を通して、超学際的な協働を促進する。

コミュニティの形成

- **研究者**：研究データ・オープンサイエンスの新しい事例、その学術的議論。
- **図書館員（データライブラリアン）**：研究データを扱う方法の情報交換、提案。
- **ソフトウェア開発者**：研究データ基盤の構築に関する事例紹介、ツール開発。
- **「オープン」を梃子に、新しい研究システムの実現に向けて動き出しましょう**。

関連サイト

http://agora.ex.nii.ac.jp/~kitamoto/research/open-science/

1. オープンサイエンスへのコンバージェンス～同床異夢から共通認識を醸成するコミュニティの形成～
2. 巨人の肩とデータサイテーション
3. 「研究のバリア」を打破する研究基盤デザインと研究データ利活用
4. オープンサイエンスのジレンマ～研究者共同体のインセンティブと結果オープン性へのモチベーション～