

第2回東京大学学術資産アーカイブ化推進室主催セミナー
かわいい子には旅をさせよ — デジタルアーカイブとオープンデータ —
2018年11月22日(木) 東京大学 福武ホール

オープンサイエンスとオープンデータ

武田英明
国立情報学研究所

takeda@nii.ac.jp

ORCID: 0000-0002-2909-7163

オープンサイエンス

- 社会と科学のデジタル化・インターネット化に伴う科学の新しい形
- キーワード
 - 論文のオープンアクセス
 - 研究オープンデータ
 - データ・リポジトリ
 - 研究データ共有
 - 研究データ・アーカイブ
 - 科学プロセスのオープン化
 - 市民参画
 - ...

なぜオープンサイエンスか？ オープンとは何か？

- 特にインターネットのオープン性が強く反映
- インターネットのオープン性から科学のオープン性はいくつかの異なるルートから来ている
 - 科学コミュニティのオープン性
 - 科学のデジタル化・インターネット化(データ中心科学)
 - 学術出版のデジタル化・インターネット化(オープンアクセス)
 - 政府のデジタル化・インターネット化(オープンデータ)
 - 市民科学(citizen science)のデジタル化・インターネット化

データ中心科学

各分野でのデータ共有の進展

WDC (1950s-)

ICSU WDS (2008)

サイエンス

Budapest Open Access Initiative (2002)

シリアルズ・クライシス

オープンアクセス

SPARC

オープンサイエンス

大学図書館/STM出版

オープンアクセス出版

科学データ共有

科学データ・アーカイブ

機関リポジトリ

データリポジトリ

政府

Web 2.0

インターネット

オープンデータ

オープン

Creative Commons

研究のオープン化

デジタル化

オープンイノベーション

研究オープンデータ

オープンソース

WWW(Web)

オープンガバメント

Citizen Science

FTP

WWW

Cloud Computing

1990年代

2000年代

2010年代

科学のデジタル化・インターネット化

- 科学自身がデジタル化・インターネット化することにより、既存の研究方法を向上させたり、新しい研究方法が生まれている。
- 第四の科学
 - 経験科学
 - 理論科学
 - シミュレーション科学
 - データ中心科学
- 手段としてのオープン化

学術出版のデジタル化・インターネット化

- 論文の電子出版化とシリアルズ・クライシス
- オープンアクセス
 - 2002年 ブタペスト・オープンアクセス宣言
- 機関リポジトリ
- プリプリント・サーバ

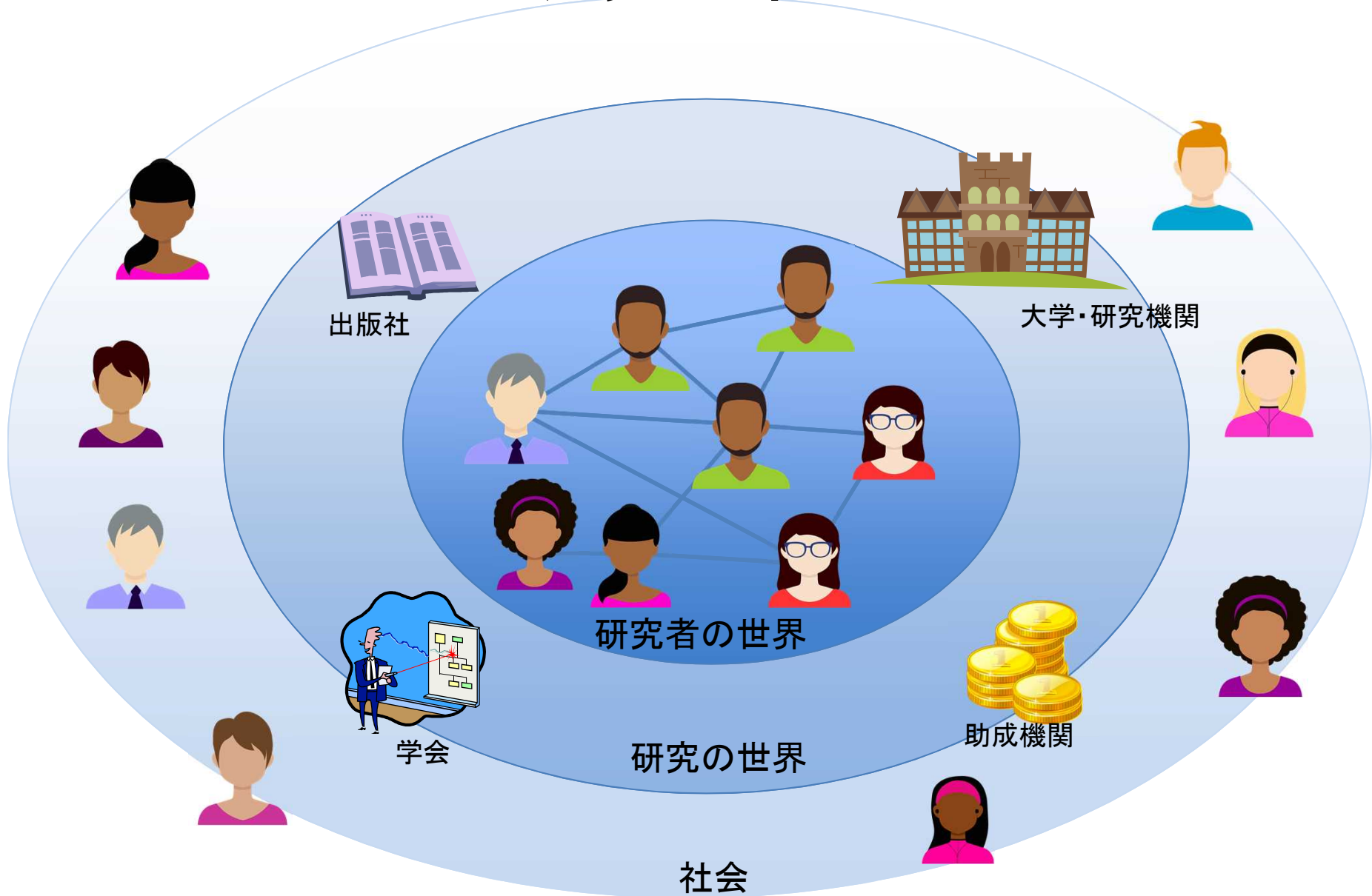
政府のデジタル化・インターネット化

- Web 2.0からGovernment 2.0へ
- オープンガバメント
- 公共としての研究
 - オープンデータ

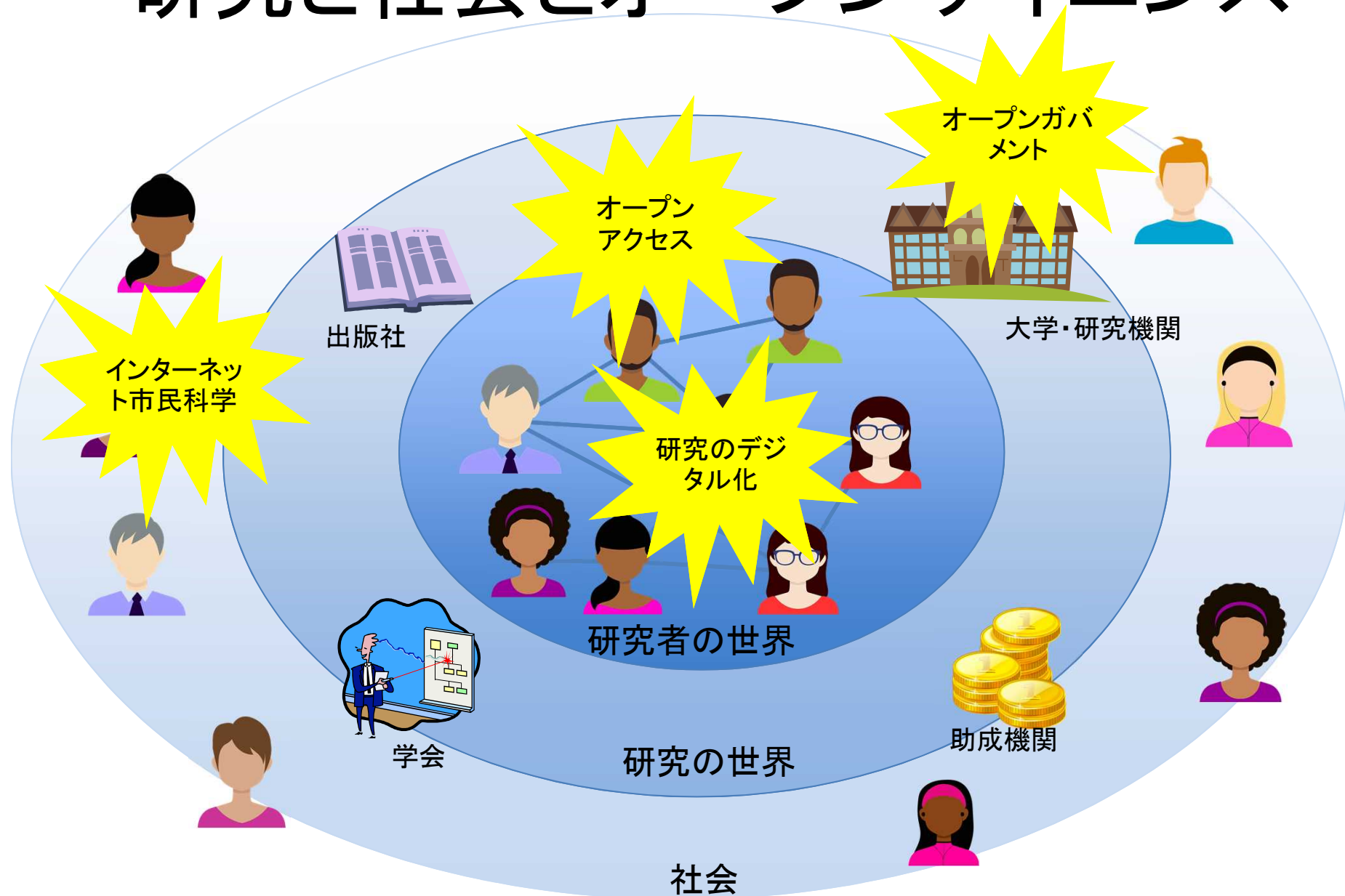
市民科学(citizen science)のデジタル化・インターネット化

- 市民の科学活動への参加、市民と研究者の共同
 - 命名は1990年代、しかしそういう活動は昔から存在
 - 例) 市民による観測データ収集
- デジタル化・インターネット化による新しい方法の発明
 - 例) Galaxy Zoo, ニコニコ学会
- 成り立ちからしてオープン化は自然な成り行き

研究と社会



研究と社会とオープンサイエンス



オープンサイエンス

オープンサイエンスとはデジタルメディアを用いた科学の無境界化(ボーダレス化)

– 何の境界か

- 科学者間、分野間、科学者・市民間の境界

– 無境界化するものは何か

- 科学のプロセス

– 多様な科学の担い手の出現、多様なコラボレーションの実施

- 科学の成果

– データのオープン化、共有

“オープンサイエンス”の例

『明月記』と『宋史』の記述から、平安・鎌倉時代における連発巨大磁気嵐の発生パターンを解明

国立極地研究所（所長：白石和行）、国文学研究資料館（館長：今西祐一郎）、京都大学（総長：山極壽一）等の研究者からなる研究グループは、『明月記』などの古典籍に残されたオーロラの記述と、樹木年輪の炭素同位体比を比較することなどにより、平安・鎌倉時代における巨大磁気嵐（注1）の発生パターンを明らかにしました。本研究は、総合研究大学院大学（学長：岡田泰伸）の複合科学研究科極域科学専攻と文化科学研究科日本文学研究専攻の学融合共同研究事業「オーロラと人間社会の過去・現在・未来」及び「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」の中で実施されました。

藤原定家（1162-1241）が残した『明月記』には、1204年2月21日と23日、京都でオーロラが見えたとの記述があります。これは1週間のうちに何晩も、京都のような緯度の低い地域でオーロラが観測された「長引く赤いオーロラ」の記録としては、これまでに調査されている中では日本で最古のものです。中国の歴史書『宋史』には、同じ2月21日に、太陽に大きな黒点が観測されたという記述があります。これらの記述は、現代的な観測データから推定する限り、太陽から噴き出たコロナ質量放出（注2）が何度も地球に直撃することで、大きな磁気嵐が単発で終わらず何度も発生する「連発巨大磁気嵐」が起きていたと考えられます。

研究グループは、さらに時代を遡って連発巨大磁気嵐の発生パターンを検討するため、『宋史』における900年代～1200年代の「長引く赤いオーロラ」の記録と、太陽活動の強弱を反映する樹木年輪の炭素同位体比の測定データを比較しました。その結果、太陽活動の極小期前後よりも、極大期付近に多く記録されていたこと、また、太陽活動が長期的に弱くなった1010～1050年には、そのような「長引く赤いオーロラ」の記述がないことが明らかになりました。

これらの知見は、今後の最悪の磁気嵐を予測する上で重要な情報となります。本研究は米国地球物理学会の発行する学術誌「Space Weather」にオンライン掲載され、同誌のEditors' Highlightに選ばれました。

“オープンサイエンス”の例

2016/11/24

江戸の文化を現代に取り込む「江戸料理レシピデータセット」を整備～江戸時代の料理本を「レシピ化」し、クックパッドでも公開～

大学共同利用機関法人情報・システム研究機構 国立情報学研究所(NII、所長:喜連川 優、東京都千代田区)と大学共同利用機関法人人間文化研究機構 国文学研究資料館(国文研、館長:今西 祐一郎、東京都立川市)は11月24日より、江戸時代の料理本に書かれた調理手順をデジタル化し、現代の人々でも使える形式に整備した「江戸料理レシピデータセット」の公開を始めました。また、現代語訳、および、現代の「レシピ」形式に編集したデータを、同日から料理レシピ投稿・検索サービスの「クックパッド」でも公開します。



毎日の料理を楽しみにする **cookpad** 282万レシピ

料理名・食材名 目的・用途 レシピ検索

バレンタイン 恵方巻 ひき肉 チキン南蛮 さば

クックパッド江戸ご飯のキッチン

クックパッド研究開発部が江戸時代のレシピを掲載していきます。食文化研究会様 (...)

フォローする

トップ レシピ 56 つくれぽ 献立

カテゴリ

すべて (56)

江戸時代のおやつ (4)

江戸の野菜料理 (1)

江戸時代の定番 (3)

レシピを検索 検索 1/6 ページ 次へ

【江戸の味】うずみ豆腐

ご飯、豆腐、わさび味噌がベストマッチ！絶妙な美味しさです！

材料：絹ごし豆腐、花かつお、わさび味噌、ご飯

【江戸の味】こおり豆腐(黒蜜かけ)

シンプルで、とても美味しい江戸時代のデザートです。豆腐なのにデザート。さらに豆腐の浮...

材料：絹ごし豆腐、かんてん、水、白砂糖、黒砂糖、水

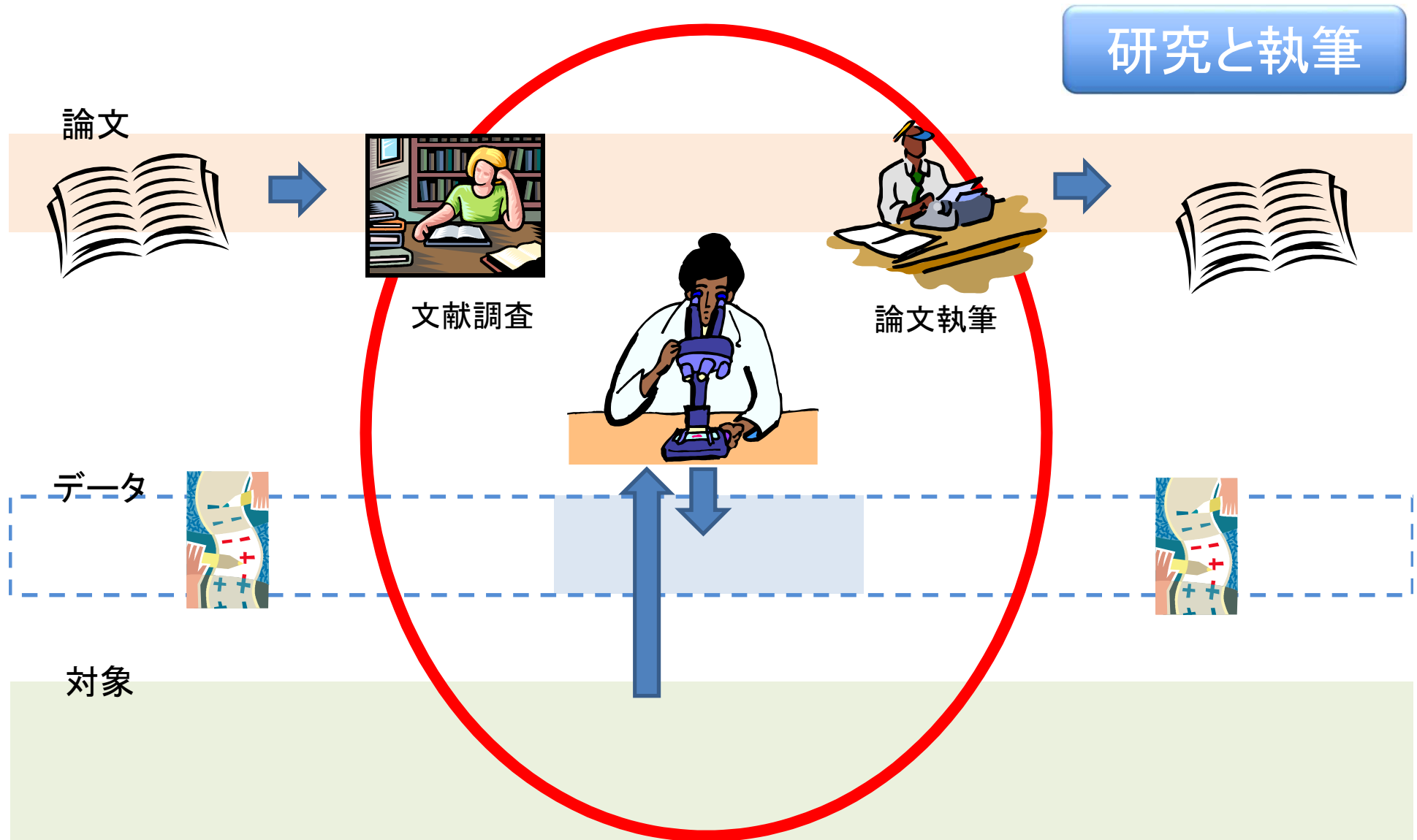
【江戸の味】煎酒 醤油以前からの調味料

さっぱり味のさわやかな味です。刺し身やサラダのドレッシングにぴったりです！

材料：日本酒、梅干し、鰹節

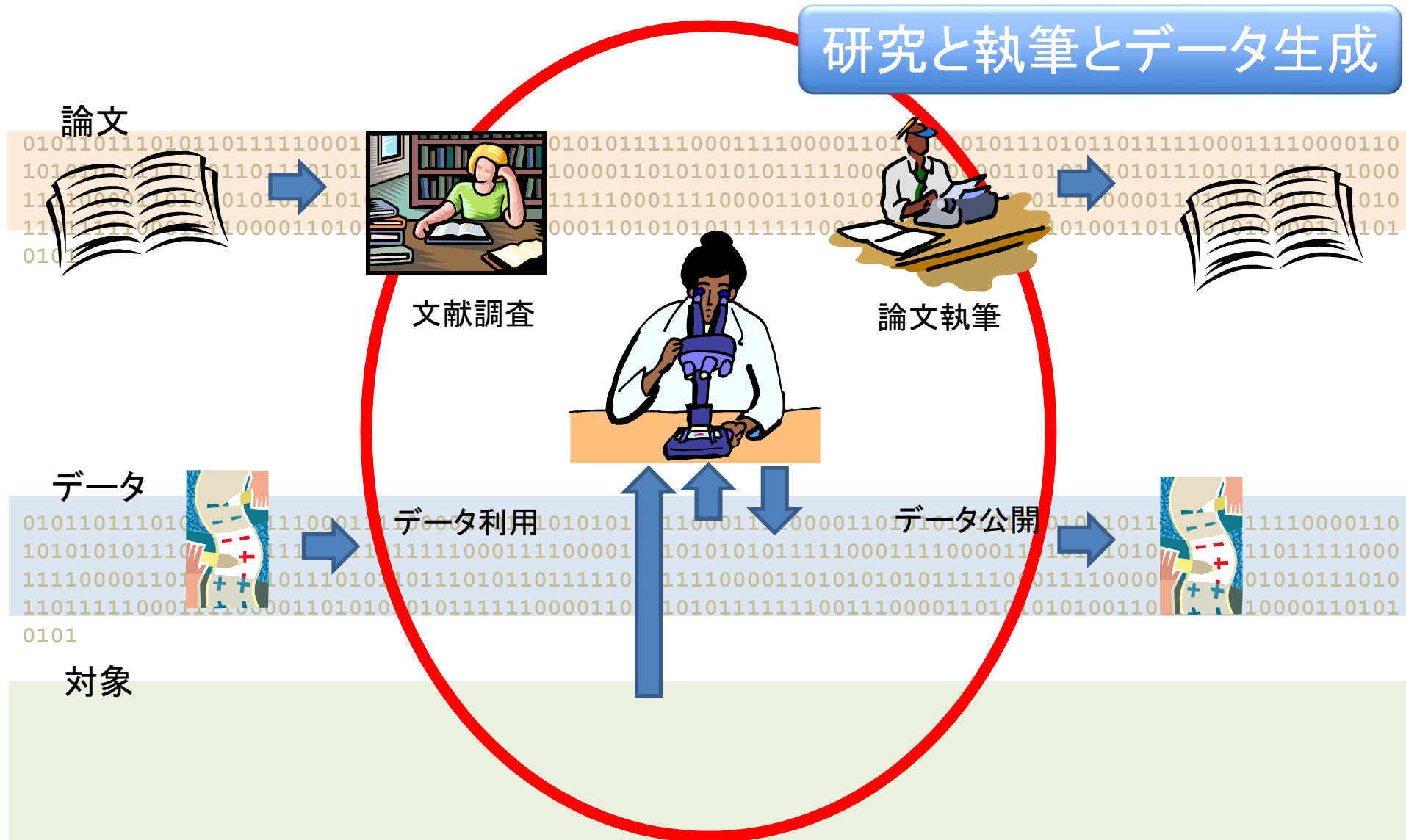
オープンサイエンスにおける研究スタイル

デジタル化以前の研究者



現在の研究者

研究と執筆とデータ生成



今後の研究者

研究成果＝データ生成

論文・データの一体化

論文



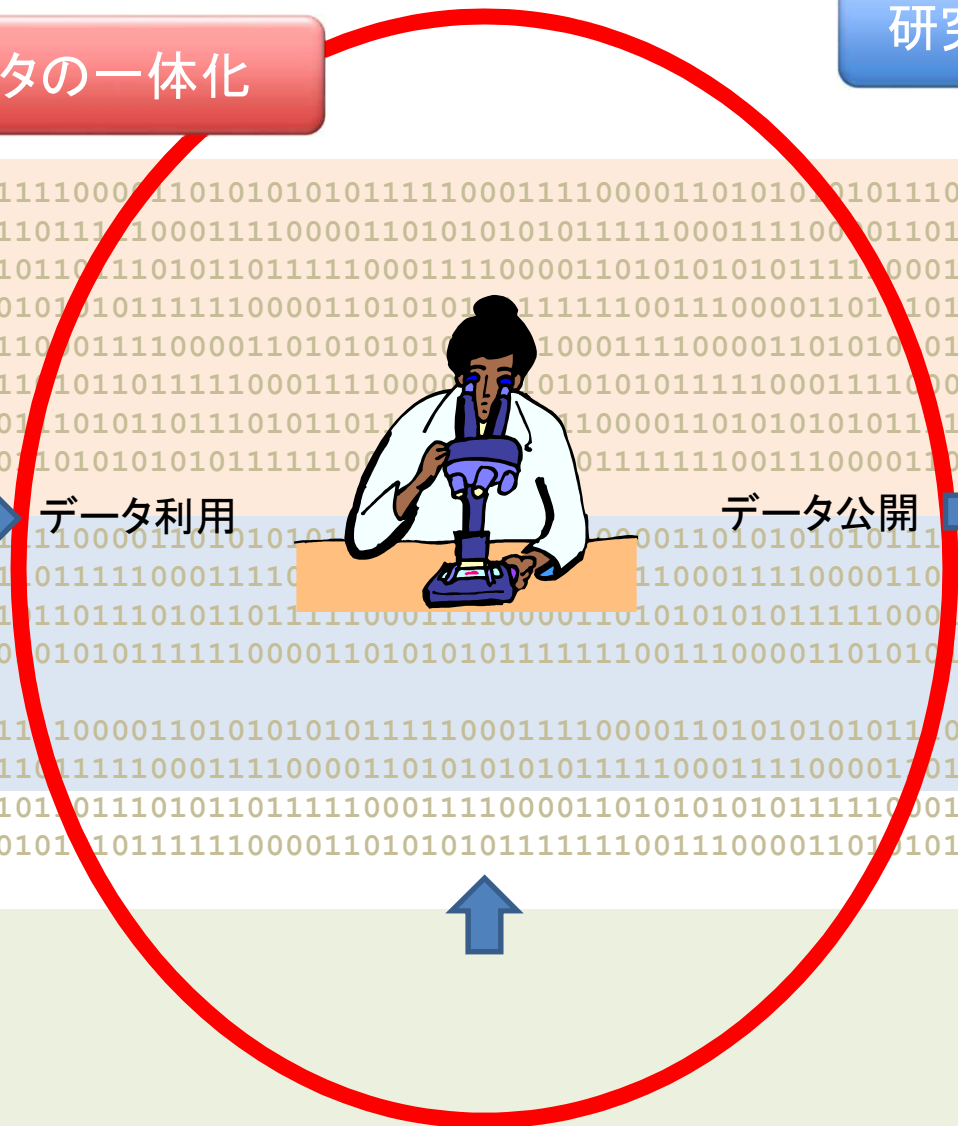
データ利用

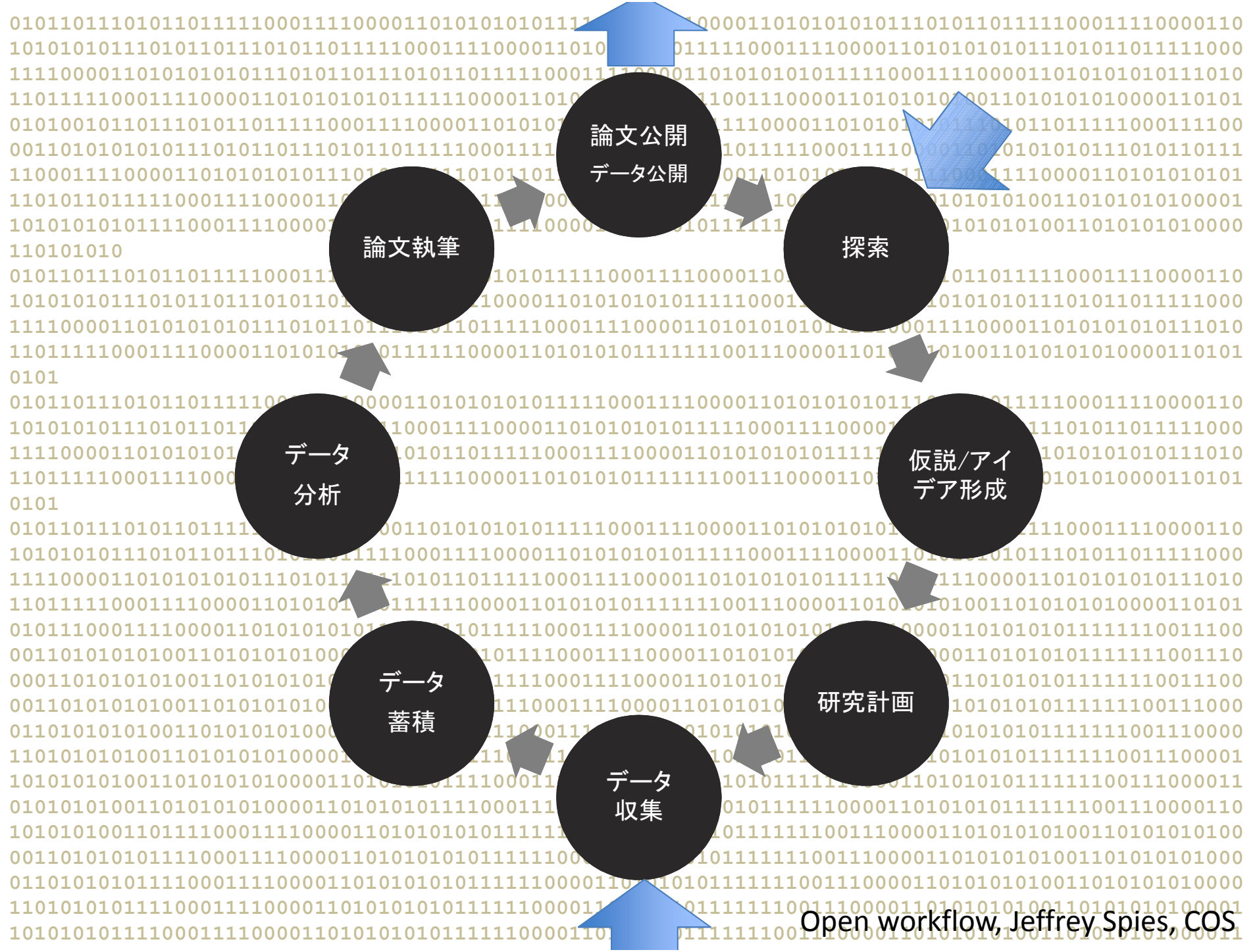
データ公開

データ

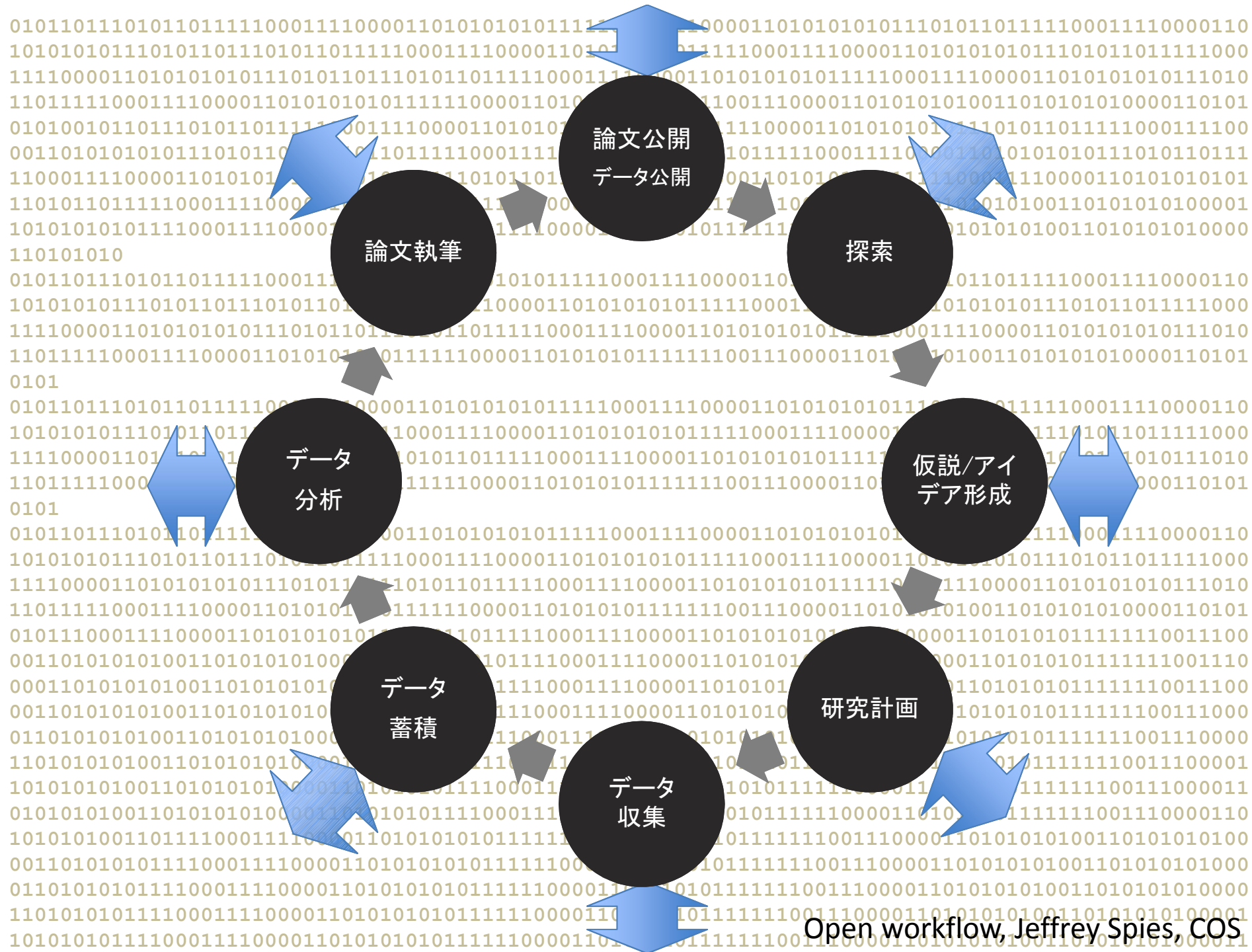


対象



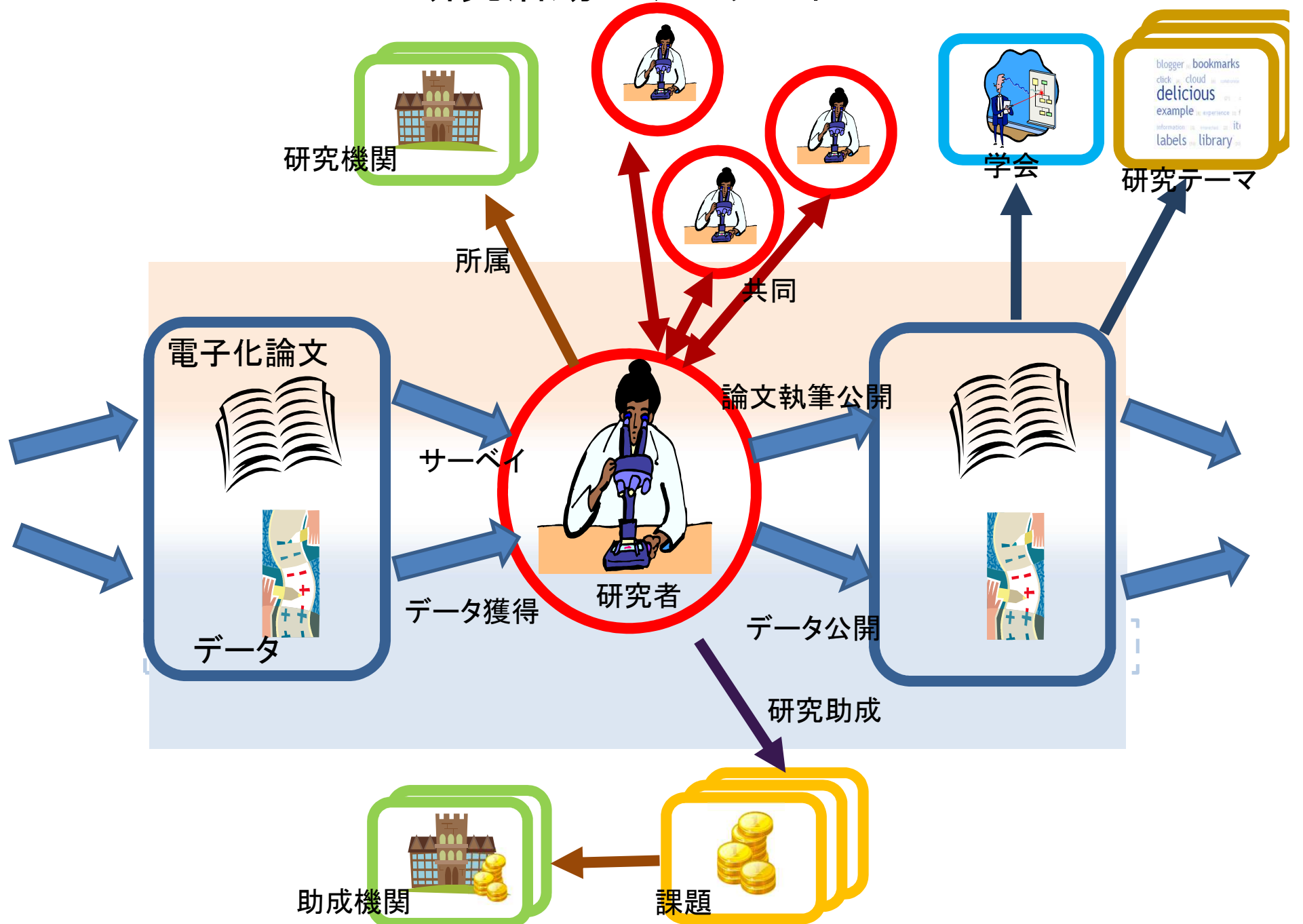


Open workflow, Jeffrey Spies, COS



Open workflow, Jeffrey Spies, COS

研究活動のデジタル化

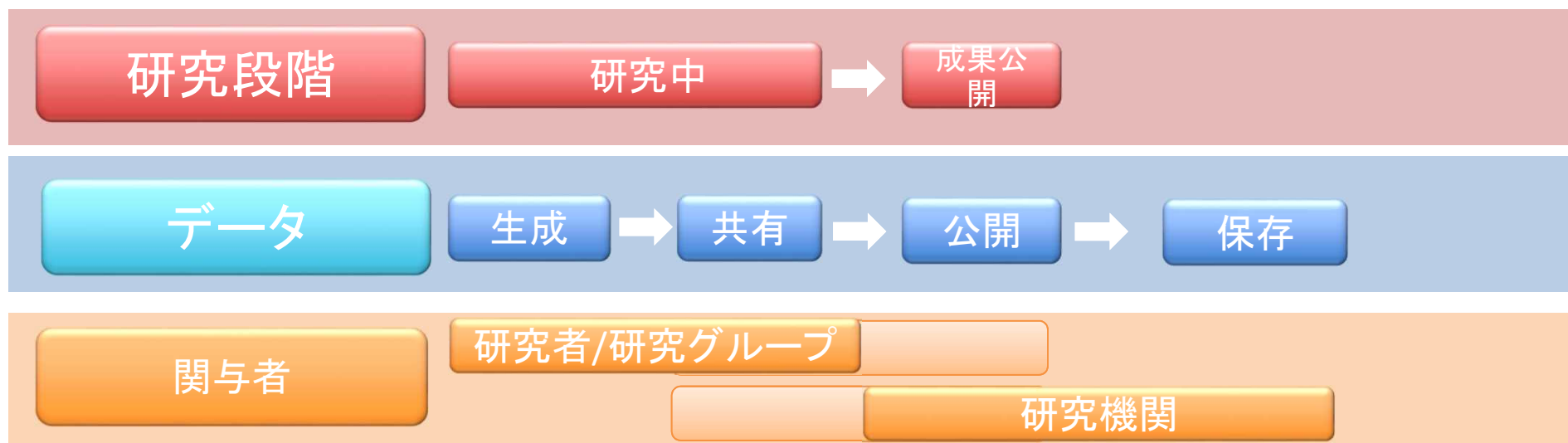


データの共有と公開

- 研究ワークフロー
- データのライフサイクル
- データ共有/公開のメリットとデメリット
- データ公開の原則
- データの公開のためのアーキテクチャー

データのライフサイクル

- データは作られ, **共有され**, **公開され**, 保存される
 - 多くは共有からオープンに公開へ
 - 一部(プラバシーデータ、セキュリティデータ)は共有のまま
- データのライフサイクルを通じたサポート



データ共有のメリット

- データの早期公開はよりよい成果が期待できる
 - エラーの早期発見、早いコミュニティ形成
- 一つのデータから多様な研究
- 再現可能性
- 他データとの結合
- 学際的研究の促進
- データの保全
- サイテーション
- 教育やアウトリーチ
- 社会や市民科学とのつながり

データ共有のデメリット

- 内部利用より高度な“標準化”の必要
- キュレーション
- 維持コスト
- 横取り研究の可能性

FAIR原則

- Findable 見つけられる

- (メタ)データはグローバルで永続的でユニークな識別子を持つべき
- データは豊富なメタデータで記述されるべき
- (メタ)データは検索可能な資源に登録あるいはインデックス化されるべき

- Accessible アクセスできる

- (メタ)データは標準的な通信プロトコルで識別子を使って取得できるべき
- プロトコルはオープンでフリーで汎用の実装可能であるべき
- プロトコルは必要であれば認証、認可の手順を持つべき
- メタデータはデータが入手不可になってもアクセス可能であるべき

- Interoperable 相互運用可能

- (メタ)データは知識表現として形式的かつアクセス可能かつ共有可能かつ広く適用可能な言語を使うべき
- (メタ)データはFAIR原則に沿った語彙を使うべき
- (メタ)データは他の(メタ)データへの適切な参照を持つべき

- Re-usable 再利用できる

- メタ(データ)は精度と関連性に関する属性を複数持つべき
- (メタ)データは明確でアクセス可能なデータ利用ライセンスを付与すべき
- (メタ)データは由来をつけるべき
- (メタ)データは領域に関連したコミュニティの標準に合致すべき

オープンサイエンスの可能性とリスク

- 新しい学術コミュニケーションの可能性
 - 新しい参加者、新しいコラボレーション
 - 研究者、研究グループ、分野を超えた、データの新しい利用、再利用、結合、変換
 - 研究データを通じた社会とのコラボレーション

- 新しい学術コミュニケーションのリスク
 - 既存のエコシステムの破壊、変容に伴う混乱
 - メディア: 学術ジャーナル、国際会議
 - 組織: 学術出版社、学会、大学
 - ルール: 評価方法
 - 新しいデジタルインフラストラクチャの確立

オープンサイエンスにおける学術コミュニケーション

- これまでの学術コミュニケーション(特に科学)においては、学術ジャーナルを中心に構築されていた。
 - 保存メディアとして
 - 公開、頒布メディアとして
 - 評価システムとして
- オープンサイエンスにおいては、もはやジャーナル(だけ)では担えない、担う必要はない。
 - 流通するものは、論文だけではなくデータも
 - インターネットがあれば、公開、頒布は誰でもできる。では誰が誰にどう公開、頒布するか
 - 評価はどうするか。
- 学術コミュニケーションのための新しいインフラストラクチャの誕生

オープンサイエンスにおけるデジタル インフラストラクチャ

- オープンサイエンスにおける学術コミュニケーションを支えるシステム
 - Post journal
 - Beyond journal
- 求められる特性
 - データが中心
 - 多元性
 - さまざまな視点からのデータの取り扱い
 - オープンであるが、同定可能
 - どんなエンティティ(データ、人、組織等)も参加可能であるが、それが特定できる必要
 - 分散性
- 課題
 - コスト
 - ガバナンス

学術コミュニケーションのための デジタルインフラストラクチャの要素

- 研究データマネージメントシステム
- 永続的識別子 (Persistent Identifier, PID)
- リポジトリ
 - 機関リポジトリ、サブジェクトリポジトリ、データリポジトリ
- プリプリント、OA publishing
- ディスカバリーサービス
 - 検索
 - 結合
- 評価システム

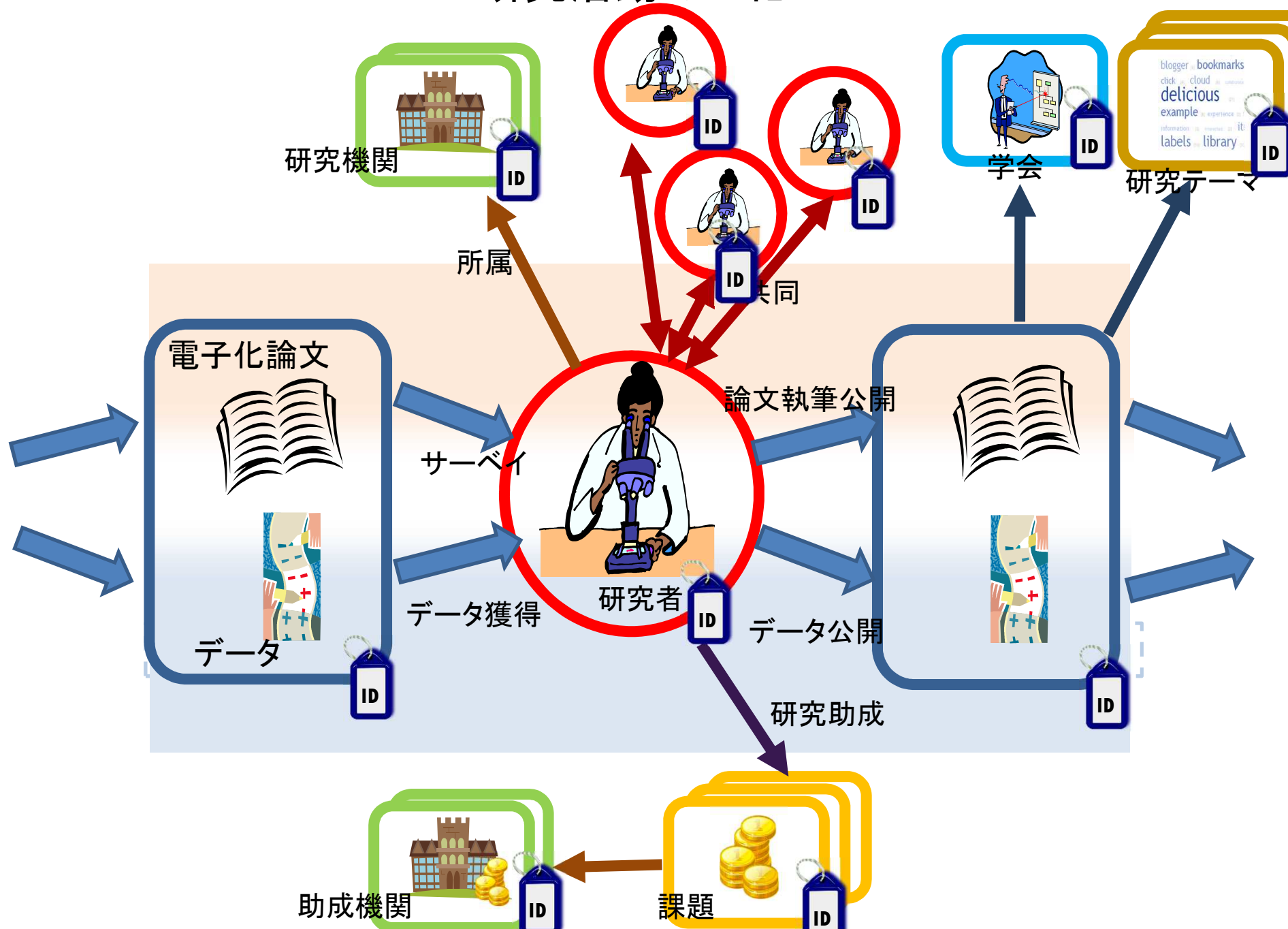
学術コミュニケーションのための デジタルインフラストラクチャの要素



Persistent Identifier (PID)

- PIDの重要性
 - オープンな世界であるこそ、同一性や識別性が重要
 - PIDはそれをシステムとして担保する
 - 多くの場合は、IDとそれのメタデータがセットで管理される。
- PIDの例
 - DOI (Crossref, DataCite, JALC, ...)
 - ORCID
 - Open Funder Registry
 - RAID
 - IGSN
 - OrgID
 - Instrument ID
 - ...

研究活動のID化



DOI (Digital Object Identifier)

- 識別子(DOI name)からデジタルオブジェクトが存在するURIに変換するサービス
- International DOI Foundation(IDF)が運営するサービス(1998年～)
- 元々は出版社が識別子を共有するにつくった制度であるが、現在はより広範に様々なデジタルオブジェクトの識別子を提供する制度に。
- 分散管理
 - IDFは登録業務を行う団体(Registration Agency, RA)に委譲
 - CrossRef, DataCite, JaLCはRA

IDF 登録機関(Registration Agency)

- CrossRef
 - 学術的な専門的な研究コンテンツ
 - 雑誌論文、図書、学会予稿など
 - 引用リンク、検索可能なメタデータデータベース



- mEDRA
 - インターネット上の文書のための永続的引用システム
 - 知的所有実体 (intellectual property entities) 相互の関係トラッキング
 - タイムスタンプや電子署名の認定



- OPOCE (Office des publications EU)
 - EU 配下組織の出版物のためのDOI 管理を通じ、EU の公式パブリッシャーとして機能



- China National Knowledge Infrastructure (CNKI)
 - 中国の政治・経済・人文科学・科学技術などの情報。



- DataCite
 - データセットに関する学術基盤構築。ベストプラクティスの共有、
 - 識別や名前解決の課題などへの取り組み



- EIDR
 - 映画、テレビ番組、商業音声／映像資料へのレジストリ
- Wanfang Data Co., Ltd.
 - 中国語雑誌のリンクサービス構築と、中国の科学データ管理



- Airiti, Inc
 - 繁体中国語資料へのDOI
- ISTIC (The Institute of Scientific and Technical Information of China)
 - 中国語学術雑誌・中国の科学データ。



- Japan Link Center (JaLC)



- KISTI

CrossRef

- STM出版における引用関係を管理する仕組み
- 1999年～
- DOIのRA
- 世界70カ国に会員を持ち、現在登録されているDOIの大部分がCrossRefのDOI。
- 機能：
 - DOIのメタデータ(書誌データ、引用関係)の管理
 - DOI登録
 - メタデータを用いたサービス
 - 書誌検索、引用検索
 - Reverse look up: メタデータを検索して対応するDOIを返す
 - この機能を用いて、出版社は個々の引用文献にDOIを付与

コスト: 会員(主に出版社)
ガバナンス: NPO(NGO)

Japan Link Center (JaLC)

- 平成24年3月に世界で第9番目に国際DOI財団(International DOI Foundation, IDF) から、DOI登録機関(RA)の認定を受けた。
- 同年5月に4国内学術機関による共同で運営を開始
 - 科学技術振興機構 (JST)
 - 国立情報学研究所 (NII)
 - 国立国会図書館 (NDL)
 - 物質・材料研究機構 (NIMS)

コスト: 国(JST)主体 + 会員
ガバナンス: 国系研究所


国文学研究資料館の例

国文研 大型プロジェクトのデータベースにもDOI

今後、国文研は30万点を目標とした古典籍画像データベースを構築していきます。2017年4月より、そのコンテンツにDOIを付与する予定です。このDOIを論文に明記すれば、論文に用いた古典籍をダイレクトに確認できるのです。

国文学研究資料館 古典籍画像 30万点(公開予定)

古典籍画像データベース



源氏小鏡

実際の画像とは異なります

デジタル画像データ(DOI付与) + 書誌データ



どこでも、いつでも、誰でもアクセスできます!



国文研は古典籍データへのDOI付与を推進します

国文学研究資料館

〒190-0014 東京都立川市緑町10-3

E-mail: cjinfo@nijl.ac.jp

URL: <http://www.nijl.ac.jp/pages/cjproject/>



検証可能な学問に向けて

論文に引用する古典籍に
DOIを明示しましょう



日本語の歴史的典籍の国際共同研究ネットワーク構築計画



<http://id.nii.ac.jp/1283/00003167/>

国文学研究資料館の例

DOI (ディー・オー・アイ) って何

Digital Object Identifier の略。

恒久的にデジタル情報を特定できる、国際的な識別子、それがDOIです。対象は、書籍や論文にとどまらず、研究データ、さらに映画やテレビ番組などの情報資産にも、広がっています。

インターネット上にあるデジタルコンテンツの所在情報は、一般的にURL (http//...) によって示されています。しかしページやコンテンツの場所などが変わるたびにURLも変更になり、わずかに数年で、対象のサイトに行き着けないということがしばしば。

そこで、デジタルコンテンツへの永続的なアクセスを実現するため考案されたのが、DOIです。

日本では、大学図書館や国立情報学研究所、国立国会図書館が古典籍のデジタル画像へのDOI付与に先進的に取り組んでいます。国立国会図書館では、博士論文(14万件)等への付与に加えて、2015年2月から、約9万件の古典籍画像にもDOIが付与されました。DOIは、いま研究者が知っておくべき識別子なのです。

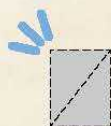
古典籍もインターネットの時代に→注目されるDOI

(参考) DOIハンドブック

http://www.doi.org/doi_handbook/translations/japanese/hb.html



DOI利用のメリット



リンク切れなし

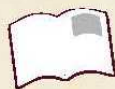
恒久的に、確実に、画像に到達できます。また、DOIはURLに変換が可能です。



画像の確認が簡単

DOIを入力するだけで画像にたどりつけるので、OPAC(蔵書検索システム)を利用して、所蔵機関の請求記号を調べるなどの手間が省けます。また、すぐに画像で確認できるので、論文内容の検証が容易になります。

たとえば、発表の途中でも資料に示されたDOIをタブレットやPCで検索すれば、参加者はその場ですぐに検証が可能となるのです。



書誌情報も同時に確認可能

国文研の場合、DOIは今後のシステムで書誌情報とともに示される予定ですので、DOIを入力することで、画像を含む書物全体の情報を得ることができます。

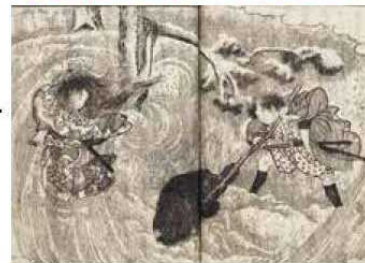


タイトルだけでは区別しにくい 多くの版種がある場合も便利

たとえばDOIをつけることで、国文研に2点ある『優曇華物語』を区別することができます

国文学研究資料館蔵『優曇華物語』

DOI:10.□□□□/○○○○○○○○



(文化元年 鶴屋喜右衛門刊)

DOI:10.□□□□/△△△△△△△△



(嘉永元年の後印)








薄墨色の背景がなくなる

<http://id.nii.ac.jp/1283/00003167/>

論文出版(公開)の多様性

- ジャーナル論文だけではない多様な出版(公開)手段の発生
- プリプリントサーバ
- 機関リポジトリ
- サブジェクトリポジトリ
- オンラインプロシーディングサーバ
- オーバーレイ出版

台頭するpreprint サーバ

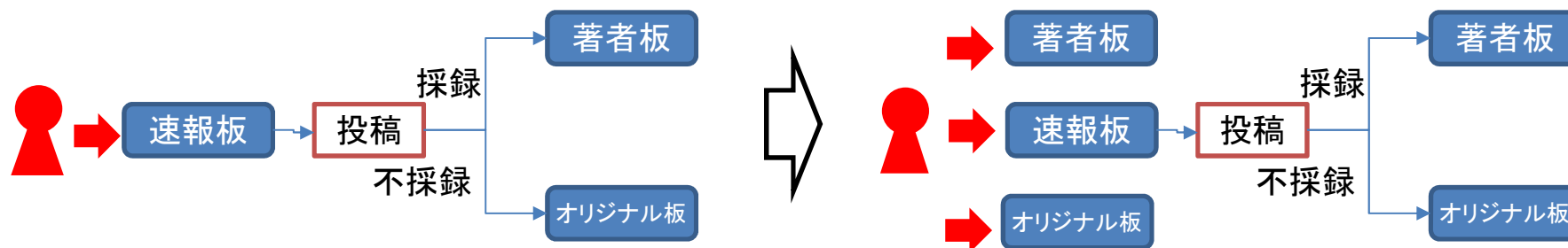
- 1991  **arXiv**: 物理学、数学、計算機科学、統計学、経済学
- 1993  **SSRN**: 社会科学 他 [Elsevier]
- 2013  **PeerJ Preprints**: バイオサイエンス [PeerJ]
- 2016  **bioRxiv**: バイオサイエンス [Cold Spring Harbor Laboratory]
- 2016  **engrxiv**: 工学 (OSF)
- 2016  **socARXIV**: 社会科学 (OSF)
- 2016  **PysArXiv**: 心理学 (OSF)
- 2017 **BioRN**: バイオサイエンス [Elsevier]
- 2017  **LawArXiv**: 法学 (OSF)
- 2017  **AgriXiv**: 農学 (OSF)
- 2018  **ChemRxiv**: 化学[ACS, GDCh, RSC](figshare)

プリプリントサーバの新しい役割

- 学術コミュニケーションのエコシステムの一部として機能することが求められている。
- すなわち、多様な出版（公開）手段の一つとしてのpreprintを位置付け、機能を再設計することが求められている。
 - 他の出版（公開）手段の役割分担
 - 他の出版（公開）手段との連携
 - オープンアクセスへ積極的な参画

プリプリントサーバの役割

- 速報板
 - 発見などを素早く報告。その後、ジャーナル投稿、出版される。
 - 著者版
 - OA版をプリプリントサーバにおく
 - 出版版との共存の一つの手段
 - 速報板もジャーナル論文が出版された後は著者版の役割
 - オリジナル版
 - プリプリントサーバの論文が原板となる論文
 - 時間が勝負の論文
 - Citationされる
 - 速報板もジャーナル論文が不採録の時はオリジナル版の役割
- これまで、プリプリントサーバの役割は速報板が第一義であったが、そのプロセスで生じる著者版、オリジナル版としての利用が出てきている。



速報性が重視される分野での利用例

arXiv.org > stat > arXiv:1610.02920

Statistics > Machine Learning

Generative Adversarial Nets from a Density Ratio Estimation Perspective

Masatoshi Uehara, Issei Sato, Masahiro Suzuki, Kotaro Nakayama, Yutaka Matsuo

(Submitted on 10 Oct 2016 (v1), last revised 9 Nov 2016 (this version, v2))

Generative adversarial nets from a density ratio estimation perspective. M Uehara, I Sato, M Suzuki, K Nakayama... arXiv preprint arXiv:1610.02920, 2016 - arxiv.org. Generative adversarial networks (GANs) are successful deep generative models based on a two-player minimax game. However, the objective function derived in the original motivation is changed to obtain stronger gradients when learning the generator.

☆ 99 引用元 23 関連記事 99

Learning in implicit generative models

S Mohamed, B Lakshminarayanan - arXiv preprint arXiv:1610.03483, 2016. Generative adversarial networks (GANs) provide an algorithmic framework for constructing generative models with several appealing properties: they do not require a function to be specified, only a generating procedure; they provide san

☆ 99 引用元 86 関連記事 99

arXiv.org > stat > arXiv:1610.03483v1

Statistics > Machine Learning

Learning in Implicit Generative Models

Shakir Mohamed, Balaji Lakshminarayanan

(Submitted on 11 Oct 2016 (this version), latest version 27 Feb 2017 (v4))

萩原 立雄

Learning in Implicit Generative Models

Shakir Mohamed and Balaji Lakshminarayanan

Google DeepMind, London

{shakir, balajiln}@google.com

Abstract

Generative adversarial networks (GANs) provide an algorithmic framework for constructing generative models with several appealing properties: they do not require a function to be specified, only a generating procedure; they provide san

[37] M. Sugiyama, T. Kanamori, T. Suzuki, M. C. Du Plessis, S. Liu, and T. Takouchi. Density-difference estimation. *Neural Computation*, 25(10):2734–2775, 2013.

[38] M. Uehara, I. Sato, M. Suzuki, K. Nakayama, and Y. Matsuo. Generative adversarial nets from a density ratio estimation perspective. *arXiv preprint arXiv:1610.02920*, 2016.

[39] J. Zhao, M. Mathieu, and Y. LeCun. Energy-based generative adversarial network. *arXiv preprint arXiv:1609.03126*, 2016.

どちらもconference論文になってないよう。arXiv論文が直接citeされている。

データ・リポジトリの重要性の増大



SCIENTIFIC DATA 



Search



E-alert



Submit



Login

Policies

[Editorial & Publishing Policies](#)

[For Referees](#)

[Data Policies](#)

[Recommended Data Repositories](#)

Recommended Data Repositories

Scientific Data mandates the release of datasets accompanying our Data Descriptors, but we do not ourselves host data. Instead, we ask authors to submit datasets to an appropriate public data repository. Data should be submitted to discipline-specific, community-recognized repositories where possible, or to [generalist repositories](#) if no suitable community resource is available.

View data repositories

- Biological sciences:
 - Nucleic acid sequence; Protein sequence; Molecular & supramolecular structure; Neuroscience; Omics; Taxonomy & species diversity; Mathematical & modelling resources; Cytometry & immunology; Imaging; Organism-focused resources
- Health sciences
- Chemistry & chemical biology
- Earth & environmental sciences
- Physics, astrophysics & astronomy
- Materials science
- Social sciences
- Generalist repositories
- Other repositories

Repository Name	Information on fees/costs	Size limits	Integrated with <i>Scientific Data's</i> manuscript submission system	Re3data / FAIRSharing entry
Dryad Digital Repository	\$120 USD for first 20 GB, and \$50 USD for each additional 10 GB	None stated	Yes ✓	view FAIRsharing entry
figshare	100 GB free per <i>Scientific Data</i> manuscript. Additional fees apply for larger datasets	1 TB per dataset	Yes ✓ - To qualify for the 100 GB of free storage, data must be uploaded to figshare via our submission system. Download instructions.	view FAIRsharing entry
Harvard Dataverse	Contact repository for datasets over 1 TB	2.5 GB per file, 10 GB per dataset	No	view re3data entry
Open Science Framework	Free of charge	5 GB per file, multiple files can be uploaded	No	view FAIRsharing entry
Zenodo	Donations towards sustainability encouraged	50 GB per dataset	No	view re3data entry
Mendeley Data	Contact repository for datasets over 10 GB	10 GB per dataset	No	view FAIRsharing entry

データのディスカバリーサービス

re3data.org

Search

Browse ▾

Suggest

Resources ▾

Contact



Data sharing made easier: use Repository Finder to find the right repository for your data

More and more funders and publishers require research data to be made available in appropriate repositories, but determining which repository to choose or what counts as an “appropriate repository” can take up a lot of time. What is a researcher to...

[Read more](#)

2,000 Data Repositories and Science Europe’s Framework for Discipline-specific Research Data Management

By offering detailed information on more than 2,000 research data repositories, re3data has become the most comprehensive source of reference for research data infrastructures globally. Through the development and advocacy of a framework for discipline...

[Read more](#)

Three new DOI Fabrica features to simplify account management

Last month we launched DOI Fabrica, the modernized version of the DataCite Metadata Store (MDS) web frontend. It is the one place for DataCite providers and their clients to create, find, connect and track every single DOI from their organization...

[Read more](#)

Google Dataset Search ベータ版

データセットを検索できます



お試しください [boston education data](#) または [weather site:noaa.gov](#)

[Learn more](#) about including your datasets in Dataset Search.



日本の東京都の駅乗降客数

opendata-web.site



日本の東京都の引っ越し料金相場

hikkoshizamurai.jp



東京都の市区町村別人口数、人口密度、面積

opendata-web.site



日本の東京都の事業会社別、路線別、駅乗降客数

opendata-web.site



東京オリンピック会場一覧

日本の東京都の駅乗降客数



opendata-web.site

データセットの提供元
オープンポータル

プロバイダからの利用可能なダウンロード形式
CSV

対象期間 2017/01/01 - 2017/12/31

対象地域
東京都

説明

日本の東京都の駅一日あたり乗降客数のデータを一覧表示およびCSVデータのダウンロードが可能です。
出展：国土数値情報ダウンロードサービス 2018年7月5日 駅別乗降客数(2017年 平成29年度)

新たな評価システムを求めて

Wellcome Open Research

Wellcome Open Research

[SUBMIT YOUR RESEARCH](#)

[BROWSE](#)

[GATEWAYS & COLLECTIONS](#)

[HOW TO PUBLISH](#)

[ABOUT](#)

[BLOG](#)

[MY ACCOUNT](#)

[SIGN IN](#)

[Check for updates](#)

RESEARCH NOTE

REVISED Importance of human demographic history knowledge in genetic studies involving multi-ethnic cohorts [version 3; referees: 2 approved]

Previously titled: H3Africa: crucial importance of knowledge on human demographic history in strategies for data exploitation – an analysis of the *Luhya in Webuye, Kenya* population from the 1000 Genomes Project

[✉ Benard W. Kulohoma](#) 

[+ Author details](#)

Abstract

Paucity of data from African populations due to under-representation in human genetic studies has impeded detailed understanding of the heritable human genome variation. This is despite the fact that Africa has sizeable genetic, cultural and linguistic diversity. There are renewed efforts to understand health problems relevant to African populations using more comprehensive datasets, and by improving expertise in health-related genomics among African scientists. We emphasise that careful consideration of the sampled populations from national and within-continental cohorts in large multi-ethnic genetic research efforts is required to maximise the prospects of identifying and fine-mapping novel risk variants in indigenous populations. We caution that human demographic history should be taken into consideration in such prospective genetic-association studies.

Keywords

Africa, GWAS, Population substructure, H3Africa


METRICS

221

 VIEWS

24

 DOWNLOADS

 Get PDF

 Get XML

 Cite

 Export




 Track

 Email

 Share

Open Peer Review

Referee Status: 

Version(s)	Invited Referees	
	1	2
REVISED Version 3 published 31 Oct 2018	 read report	
REVISED Version 2 published 18 Sep 2018	 read report	 read report
Version 1 published 06 Jul 2018	 read report	 read report

- 1 **Michèle Ramsay** , University of the Witwatersrand, South Africa
- 2 **Nicola J. Mulder** , University of Cape Town, South Africa

[All reports \(5\)](#), [Responses and comments \(2\)](#)

Comments on this article

[All comments \(0\)](#)

<https://wellcomeopenresearch.org/articles/3-82/v3>

F1000Prime

F1000Prime

Article Recommendations

Search

Advanced

ARTICLE RECOMMENDATIONS | RANKINGS | F1000PRIME REPORTS | F1000 FACULTY | BLOG

MyF1000 | SIGN IN | SUBSCRIBE

ARTICLE RANKINGS

All Time Top 10
Current Top 10
All Time Most Viewed
Current Most Viewed
Hidden Jewels

F1000Prime »

Rankings

When Faculty Members recommend articles, they rate them as 'Good', 'Very Good' or 'Exceptional' (1, 2 or 3 stars, respectively). The total scores for each article are used to create **Top 10** lists in each Faculty (**All Time** and **Current**). To find other top-ranking articles in your specialty, browse and filter by Total Score.

Hidden Jewels lists contain the top-ranking articles from more specialized journals that you may have missed in your general reading.

Check the **Most Viewed** articles (**All Time** and **Current**) to see what articles your peers are reading on this site.

Current Top 10

All Time Top 10

Hidden Jewels

12



Volitional social interaction prevents drug addiction in rat models.

Venniro M, Zhang M, Caprioli D, Hoots JK ... Heins C, Morales M, Epstein DH, Shaham Y. [+](#) Nat Neurosci. 2018 Nov; 21(11):1520-1529



These authors present an innovative approach, involving multiple methods, to highlight the importance of social factors in cocaine self-administration.

4 Recommendations | Klaus Miczek 06 Nov 2018 | M Foster Olive | Christopher Evans | Markus Heilig

CONFIRMATION | CONTROVERSIAL | GOOD FOR TEACHING | INTERESTING HYPOTHESIS | NEW FINDING | TECHNICAL ADVANCE

12



Sensory neuron diversity in the inner ear is shaped by activity.

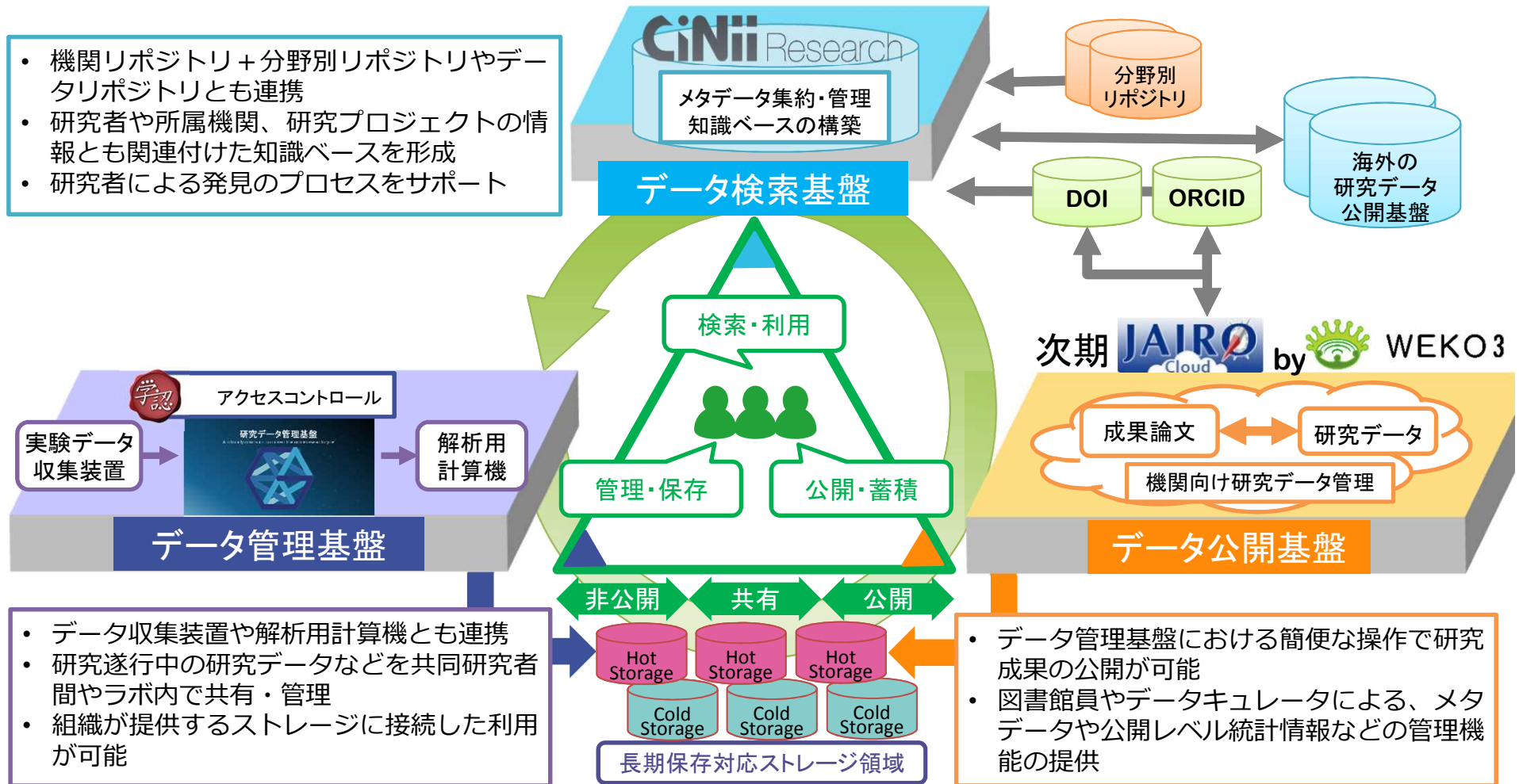
Shrestha BR, Chia C, Wu L, Kujawa SG, Liberman MC, Goodrich LV. Cell. 2018 Aug 23; 174(5):1229-1246.e17



Type I spiral ganglion neurons (SGNs) convey the auditory signal from the sensory inner hair cells (IHCs) towards the brain. Diversity of their afferents is believed to underlie...

4 Recommendations | Tobias Moser with Tina Pangršič 23 Oct 2018 | Peter Heil | Donna Fekete [ci](#) | Bernd Fritsch

オープンサイエンス対応 – 研究データ基盤



まとめ

- オープンサイエンス
 - 様々な境界がなくなる科学の世界
 - デジタル/インターネット/Webの技術の発展とともに拡大
- (研究)オープンデータ
 - オープンサイエンスの重要な基盤の一つ
 - 研究のサイクルにあわせて、データは共有、公開、保存
 - データ公開の原則(FAIR原則)
 - データ共有のアーキテクチャー
 - 識別子とメタデータの重要性(DOI, JaLC)
 - リポジトリ、プリプリントサーバの世界の拡大
 - 新たな評価への模索