提言

オープンサイエンスの深化と推進に向けて

令和２年（２０２０年）５月２８日

日　本　学　術　会　議

オープンサイエンスの深化と推進に関する検討委員会

この提言は、日本学術会議オープンサイエンスの深化と推進に関する検討委員会の審議結果を取りまとめ公表するものである。

日本学術会議オープンサイエンスの深化と推進に関する検討委員会

委員長	喜連川　優	（連携会員）	情報・システム研究機構国立情報学研究所所長 東京大学生産技術研究所教授
副委員長	引原　隆士	（特任連携会員）	京都大学図書館機構長　附属図書館館長　教授
幹　事	安達　淳	（連携会員）	情報・システム研究機構国立情報学研究所 副所長・特任教授
幹　事	村山　泰啓	（連携会員）	国立研究開発法人情報通信研究機構ソーシャルイノベーションユニット戦略的プログラムオフィス研究統括
	久留島　典子	（第一部会員）	東京大学史料編纂所教授
	溝端　佐登史	（第一部会員）	京都大学経済研究所長・教授
	澁澤　栄	（第二部会員）	東京農工大学卓越リーダ養成機構特任教授
	高木　利久	（第二部会員）	富山国際大学学長
	相澤　清晴	（第三部会員）	東京大学大学院情報理工学系研究科教授
	木村　学	（第三部会員）	東京海洋大学海洋資源環境学部特任教授
	渡辺　美代子	（第三部会員）	国立研究開発法人科学技術研究機構副理事・日本学術会議副会長
	永井　良三	（連携会員）	自治医科大学学長
	林　和弘	（特任連携会員）	文部科学省科学技術・学術政策研究所科学技術予測センター上席研究官

本提言の作成にあたり、以下の方々に御協力いただいた。

山極　壽一	（第二部会員）	日本学術会議会長、京都大学総長
上村　みどり	（連携会員）	帝人ファーマ株式会社生物医学総合研究所上席研究員
川合　眞紀	（連携会員）	自然科学研究機構分子科学研究所所長
小林　良彰	（連携会員）	慶應義塾大学法学部教授
細野　秀雄	（連携会員）	東京工業大学栄誉教授・特命教授（兼、元素戦略研究センター長）
青木　学聡		名古屋大学情報連携推進本部情報戦略室教授
赤池　伸一		内閣府政策統括官（科学技術・イノベーション担当）付参事官（エビデンス担当）
梶田　将司		京都大学情報環境機構 IT 企画室教授
山地　一禎		情報・システム研究機構国立情報学研究所教授

本提言の作成に当たり、以下の職員が事務を担当した。

事務局	犬塚　隆志	参事官（審議第二担当）
	五十嵐　久留美	参事官（審議第二担当）付参事官補佐
	薦田　有紀子	参事官（審議第二担当）付審議専門職付
	大澤　祐騎	参事官（審議第二担当）付審議専門職付

要　　旨

１　作成の背景

2009 年、4th paradigm なる書が公開された。コンピュータ界のノーベル賞であるチューリング賞受賞者 Jim Gray への追悼の書は、スーパーコンピュータによる計算科学（第３の科学）の次のステージとして、支配方程式のない科学の領域に向けて Data-Intensive Scientific Discovery、すなわちデータ駆動科学を提唱した。その後、2012 年に開催された画像認識コンペでは深層学習が従来手法に対し圧勝し、本格的な第４の科学の時代へと突入していく。若干遅れて、我が国は Society5.0(データ駆動社会、あるいは、データ駆動サイバーフィジカルシステム)を提案した。

オープンサイエンスとは、ICT を活用して科学を変容させることであり、現在研究データを含めた研究成果をインターネットの上で広く共有する科学の進め方に注目が集まっている。こうした新しい科学の在り方は、科学者の自発的な取り組みとして展開されるのが自然であるが、国内では政策的な観点からの議論が先行している。本課題別委員会では、これからの科学の作法をも変えうるオープンサイエンスを、各研究分野におけるデータ駆動科学の現状も加味した総合的な視点でとらえ、研究データ共有の促進と共有のためのプラットフォームの重要性を明らかにすることを目的として審議を進めた。

２　現状および問題点

新型コロナウイルス感染症（COVID-19）が猛威を振るう中、2020 年１月 31 日に NIH、AMED を含む国内外の研究資金配分機関や論文誌出版社が協力し、論文投稿前に実験データを WHO に提供することの合意形成がなされた。論文が出版されるまで待っている間に失われる命を救うための策である。これを受けて２月 11、12 日に WHO 本部で開催された COVID-19 緊急対策会議では「論文投稿時に WHO とデータを共有しないと論文を受理しない」ことが確認された。これは疫病と戦うための究極の国際的なデータ共有であり理想的な施策である。一方、この事象が意味する深層も理解する必要があろう。すなわち、学術分野によってデータ共有の意識や実態が異なる中、これほどの危機的状況にならないと、データの共有が進展しないことも意味する。また、データが極めて重要な資産と見なされる時代において、その共有と公開に関しては、そのデータの特性に応じた、協調と競争のバランス取りが必要となる。データを国際的に共有するかは慎重に検討する必要があり、公開から産業展開を踏まえた非公開まで協調と競争のバランスをどのように取るかは深い課題ともいえる。

研究データの重要性が著しく高まる中で、データ利用の法制度も十分に整っているとはいえない。データは著作物ではないため、著作権法による直接的保護はない。データベースの権利は規定されているものの実施例は少ない。この状況に鑑み、我が国は、2019 年不正競争防止法に「限定提供データ」と名付けたフレームワークを導入した。これは法制度により、データを他者と共有するとともに、データの権利を保護する画期的なものである。しかしながら現時点では、我が国固有の法制度でしかないという課題もある。グーグルや

アマゾンの存在を念頭に、公正取引委員会では、デジタル社会における競争政策を検討している。旧来のビジネスと異なりデジタルプラットフォーマーにおいては、市場支配力が極めて強い事業者が相対的に出現しやすく、不当なデータ収集、不当なデータの囲い込みは独占禁止法上問題となりうる。加えて、データに個人情報が含まれる場合には、個人情報保護法が適用される。EU は GDPR（General Data Protection Regulation：一般データ保護規則）なる我が国より厳しい制度を有し、2020 年米国はカルフォルニア州で CCPA(カリフォルニア州 消費者プライバシー法)なる同様の法制度が導入されつつある。

このような現状において、研究者が多層に関係する種々のルールに配慮することは大きな負担になり、また不適切なデータ利用ではないかという懸念により研究活動を萎縮させてしまいかねない。本提言では、データに対する現状の認識を多様な学術の専門家から聴取し整理することを試みたが、分野やセクターによる温度差は大きい。データ駆動科学を推進し、科学を変容するデータ共有の精神が進展すべく、日本学術会議で丁寧な議論を今後も重ねてゆくことが必須といえる。

３　提言の内容

(1)データが中心的役割を果たす時代のルール作りの必要性:データを扱う作法の混乱を避けるべく、政府は、不正競争防止法、個人情報保護法、著作権法等のデータに関連する法規制を集約・整理し、データを安心して活用できるルールを、国際的なコンセンサスを得ながら明確化する必要がある。研究データの特性と社会との関係性を踏まえた着実な取り組みを実践するための、適切なガイドラインの作成も必要である。

(2)データプラットフォームの構築・普及の必要性：膨大なデータを収集、キュレート、アノテート、メタデータ付与、保存等を推進すべく、学術界およびその研究活動を支える機関は、国の支援を得てプラットフォームの普及に努め、さらに機械学習や高次解析との接続を可能とする次世代機能の提供を実現していくべきである。このプラットフォームでは、公開から非公開までのバランスが取れたデータ共有を支援し、分野間やセクター間など多様なデータの潜在的融合を許容してさらに高い価値を生み出すことを指向すべきである。現時点においては、データを共有する意図があろうとも、経済的理由あるいはスキルの欠如等からデータ保存がなされない事態も散見される。職場の異動時や定年時を含め、データ散逸を防ぐために誰でも容易に利用できるプラットフォームが必須である。

(3)第１次試料・資料の永久保存の必要性：研究成果を直接もたらした第１次試料（岩石、堆積物、土壌、流体、生物、物質、遺構、遺物など）の永久保存体制の構築および第１次資料（文書記録、書籍、景観、技術、生活様式、生産様式など）の維持保存体制を強化する必要がある。また、それらを抽出選択する背景となった第０次試料（未研究の採集試料・資料）の選択的保存について基本方針を確立する必要がある。保存のための設備、整理、管理運営と公開方針について国際・全国連携体制で進める必要がある。

目　　次

１　はじめに..... 1
２　研究データに着目したオープンサイエンスに関する環境整備の国内・国際動向.... 1
2.1　国際動向..... 1
2.2　国内動向と課題..... 4
2.3　データと学術情報の流通..... 5
３　各学術分野における状況..... 6
3.1　材料系、化学系、製薬系..... 6
3.2　生物系、医療系..... 8
3.3　地球科学系、農業系、生態学系..... 10
3.4　人文学系、経済学系、社会科学系..... 12
４　データ基盤とガバナンス..... 15
4.1　国内の基盤整備の状況..... 15
4.2　基盤整備の方向性..... 16
4.3　ガバナンスの重要性..... 16
５　提言：オープンサイエンスの深化と推進..... 17
5.1　オープンサイエンスの深化と推進に向けて..... 17
5.2　提言..... 18
＜参考文献＞..... 21
＜参考資料１＞審議経過..... 26
＜参考資料２＞日本学術会議主催学術フォーラム開催..... 27
＜付録図表１＞研究データ基盤システムと関連体制..... 28
＜付録図表２＞学術情報の位置付け..... 28
＜付録図表３＞研究のライフサイクル..... 29
＜付録図表４＞３つの掘削プラットフォームで運用される国際深海掘削画（IODP）組織図
..... 29
＜付録図表５＞海洋の３区分と３箇所のデータ・試料永久保管施設..... 30
＜付録図表６＞高知コア研究所における試料保存例..... 30
＜付録図表７＞１ｈａ以上の大規模水田のばらつき..... 31
＜付録図表８＞精密農業の考え方..... 31
＜付録図表９＞農業用語の共通化の取り組み例..... 32
＜付録図表 10＞異なるメーカー製品の情報共有例 32

１　はじめに

近年注目を浴びている「オープンサイエンス」という用語は、学術の新たな方法論や規範を目指す種々の活動を包括する、複合的な概念として用いられることがある。例えば、欧州委員会では、オープンサイエンスは、デジタルテクノロジーを用いた相互協力および知識の伝搬に基づく、科学研究への新しいアプローチとしている[1]。そして、科学におけるよりオープンな議論・再検証を可能にしながら、新たな研究機会を増やし、科学と社会の相互信頼を確保しながら研究成果やイノベーションを望めるビジョンと考えられている。より具体的には、研究論文のオープンアクセス（OA）[1]や、データその他、市民科学、論文査読プロセスのオープン化などの学術研究活動やその成果発信・管理など全てにおいて新たなスタイルや方法論を模索するさまざまな活動を包括的に捉える概念といえ、既存の学術情報流通をすでに変革し始めている[2]。その中で、研究データに着目し、欧州・米国などで国際的にデータ保全・共有やデジタルデータ基盤整備の必要性などが合意され、科学政策検討・立案および必要な基盤投資の議論が進められてきた。

２　研究データに着目したオープンサイエンスに関する環境整備の国内・国際動向

2.1　国際動向

国際的な文脈、特に科学技術政策においてオープンサイエンスが重視されるようになったのは、2013 年の G8 科学技術大臣会合（英国）以降となる。この国際政策論議の場で初めて研究データのオープン化が大臣合意され、日本も対応を開始した。また、2016 年 G7 同大臣会合（つくば）において、研究データは社会と科学の相互信頼の上からもまた今後の社会の成長の点からも重要な情報アセット(資産)であると位置づけられた[3]。この大臣会合においてオープンサイエンスに関する部会を設置することとなり（部会議長国：日本および欧州連合）、以降毎年この部会は活動を継続している。部会は科学政策としての研究データの取り扱いに主眼をおき、２つのテーマ、研究者によるオープンサイエンス活動の評価・報酬・キャリア形成、および国際的なデータ基盤の整備、を軸に活動している。

学術・科学技術研究開発における研究データの整備共有公開の問題は、普遍的な課題として国際科学会議（ICSU）、世界科学データシステム（WDS）、科学技術データ委員会(CODATA) などによって戦後より長らく対応が進められてきたが、欧米諸国では Web が本格的に浸透し始めた 20 世紀後半から ICT の進展をふまえた議論が加速している。例えば、それぞれの議論の結果や合意事項として、全米研究評議会から 1997 年[4]、OECD（経済協力開発機構）からは 2007 年[5]、英国研究会議　(RCUK) が 2011 年[6]、英国王立協会から 2012 年に[7]、それぞれ研究データへの自由なアクセス利活用による研究の加速、および科学の進展を目指す旨の文書やガイドラインが発表されている。また、OECD は「オープンサイエンスは公的資金による主な研究成果（論文出版と研究データ）

[1] 学術論文に対して、誰もがインターネットを介して経費負担無く自由にアクセスして利用できるようにすること。

を一般に対して、最低限の制約をもってデジタルフォーマットでアクセス可能とすること」として、各国の状況をとりまとめた[8]。

学術・科学技術研究活動およびその関連活動の多様性を考慮しつつも、科学政策・公的投資の視点から研究データのオープン化を中心とした取り扱いが主に重視されてきた理由は、科学研究の発展と加速、経済政策、国際社会の動向など様々である。その上で、さらにマクロレベルで考えると「オープンサイエンス」と商業的な研究開発は、制度的に異なるサブシステムの補完関係にあり、この２つのサブシステムを用いた適切なバランス形成が公共政策としての課題となる[9]。

いずれにせよ、「オープンサイエンス」の概念は非常に幅広く、一つの定義で定められると言い切るのは困難である一方、このキーワードのもとで、例えば研究データについてだけでも、欧州・米国などで国際的にデータ保全・共有やデジタルデータ基盤整備の必要性などが合意され、科学政策検討・立案および必要な基盤投資の議論が進められている。

北米の状況

米国の研究データ管理、オープンサイエンスを大きく進展させた取り組みとして、National Institutes of Health（アメリカ国立衛生研究所, NIH）、およびNational Science Foundation（アメリカ国立科学財団, NSF）によるデータ共有のための方針の制定があげられる。特に、NSFが2011年より研究データ管理計画(Data Management Plan, DMP）の提出を義務付けたことにより、大学等研究機関では図書館を中心に研究データ管理と公開に関する組織、システムの整備が進められるようになった[10]。

California Digital Libraryでは、長年にわたり、データの公開基盤の開発と州内の教育研究機関への展開に取り組んできた。データキュレーション[2]・アーカイブ[3]システムMerritt、汎用データリポジトリ[4]基盤Dryad、様々な研究成果物に向けたPID[5]（DOIまたはARK）のリゾルバサービス[6]であるEZID等を開発し、研究データ公開を全体的に支援するためのシステムを提供している[11]。この他にも研究データの公開、保管、転送に関わる情報基盤が広く利用されるようになっている。

研究プロセスの終盤にあたる研究成果の公開から、研究開始時のデータ収集や研究途上における研究グループ内でのデータ管理に関しては、例えば、パデュー大学では目的別に複数のストレージソリューションを標準提供している[12]。また、Center for Open Science（COS）が開発、提供するOpen Science Framework（OSF.io）は、研究室や研究プロジェクトで簡便に研究データが管理できるツールとして、機関単位で採用する事例も増えている[13]。情報基盤の整備はさらに発展し、図書館組織が電子ラボノート(ELN)

---

[2] データに限定した場合、データを収集し、整理し、分類して利活用しやすい状態にする作業の一部または全部。

[3] データを記録し、保存し、将来の利用に供すること。

[4] オープンアクセスを実現するために、論文等の研究成果を蓄積し公開するためのウェブサービス。

[5] Persistent Identifier（永続的識別子）の略。デジタルオブジェクトを一意に、かつ恒久的に識別するID。

[6] デジタルオブジェクトのPIDとインターネット上で保存されるURLを関連付け、相互参照を可能にするサービス。

や研究情報管理システム（Laboratory Information Management System）の選定支援を行う[14]等、デジタル化を前提としたデータの生成から公開までをカバーする包括的な解決策を研究者が求め、研究機関がこれに応じる動きがみられる。

欧州の状況

英国では、1998 年に政府機関がとりまとめた研究者の行動規範に関する報告書をきっかけに、大学での研究データ管理の意識が高まり始め[15]、2011 年にはエジンバラ大学が英国で初めて研究データ管理の方針を策定した。同大学では、研究データの保存のストレージや公開のためのリポジトリサービスのみならず、最近では、研究のプロセスに沿ったサービスの充実化がはかられている[16]。

エジンバラ大学の先駆的な取り組みは、欧州の他の大学の研究データ管理の取り組みにも影響している。スイス連邦工科大学チューリッヒ校では、図書館と IT センターが役割分担を果たしながら、研究データのライフサイクルに沿ったサービスを提供している。IT センターでは、単にストレージを提供するだけではなく、電子ラボノートシステムを開発し[17]、教員や学生へのサービスとして提供している。その他にも、研究中にはデータを解析する基盤とともにストレージサービスを提供し、研究後には解析プログラムとデータをセットで公開できるサービスもある[18]。

各大学で個別に取り組まれてきた先進的な成果は、国や欧州委員会が推進する研究インフラの整備事業の支援を受けながら、よりスケールメリットを生かしたサービスに展開されている。特に、2008 年にポータルサイトが正式に公開された European Open Science Cloud [19]は、欧州だけではなく、他の地域や国における研究基盤の在り方に、大きく影響を及ぼしている。

豪州の状況

豪州政府は 2007 年に、研究公正の観点から研究者の行動規範に関する報告書をとりまとめた[20]。その報告書の中で、研究データ管理の環境整備とリテラシー教育は、組織の責任として実施すべきであることが言及されている。外部資金を受けた際には、その実施状況が国の研究公正委員会の調査対象となるために、大学として研究データ管理を支援する環境整備が進んでいる。欧米と比較すると、豪州の研究費助成機関は詳細な DMP の提出は要求していないが、研究プロジェクトを実行する上での DMP の効果を大学として認識し、研究データ管理システムを活用する先端的な機能として提供されている。

クイーンズランド大学では、UQ Research Data Manager（UQ RDM）という研究データ管理システムを開発し、教員や学生に提供している。研究プロジェクトを開始する際に、研究テーマなどの基礎情報を入力すると、取得するデータの量や機密性に応じて適切なデータの保存領域が自動的に確保され共同研究者間で利用できるようになる。それらは DMP の基礎情報となり、研究の進捗に応じて逐次情報の追加や更新を行う。UQ RDM では DMP を研究活動の実践時に随時参照し見直す計画書として扱い、研究者自身が、研究の管理や推進のために積極的に利活用できる環境が整っている。研究データ管理システム

の開発に積極的なニューサウスウェールズ大学やシドニー大学でも、類似した考え方のもとで開発されたサービスが提供されている。

2016 年に政府が発表した国レベルでの研究インフラ開発のロードマップに基づき[21]、大学による個別の自助努力から、より効率的な基盤の整備体制が構築されつつある。ロードマップの実現に向けて、2018 年には Australian Research Data Commons と呼ばれる研究データに関連したサービス提供機関の連合体が発足し、大学と連携した国レベルでのサービス開発が具体的に進められている。

2.2　国内動向と課題

日本のオープンサイエンス政策は、先に述べた 2013 年の G8 ロンドンサミットに始まる国際的な流れに呼応する形で始まり、内閣府において 2014 年 12 月に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が設置され検討が進められてきた。2015 年３月に報告書「我が国におけるオープンサイエンス推進のあり方について～サイエンスの新たな飛躍の時代の幕開け～」を公表し[22]、オープンサイエンスを急速な電子情報通信基盤の普及と発展に基づき、学術ジャーナル論文のオープン化（OA 化）と「データのオープン化」からなると定義づけている。2016 年に始まる第５期科学技術基本計画にオープンサイエンスの推進が明記されると[23]、統合イノベーション戦略 2018 および 2019 にてその具体的な施策が目標とともに示された[24]。その狙いは我が国の Society5.0 の実現に向けて研究データの利活用促進による新たな研究の創出とイノベーションの加速であり、中核をなす方策は、研究プロセスで生まれるデータをオープン・アンド・クローズ戦略[24][7]に基づき、管理・保存・共有し、必要に応じて公開できる研究データ基盤整備である。また、研究開発法人向けの研究データポリシー策定ガイドラインや研究データリポジトリ運用・整備ガイドラインが内閣府によって策定され、研究助成団体は研究助成に対する DMP の提出を求めるなど、研究データの管理・利活用に関する具体的な取り組みが進んでいる（付録図表１参照）。さらに、具体的な事例づくりの一環として、ムーンショット型研究開発制度において、国立情報学研究所（NII）が開発している研究データ管理基盤である NII Research Data Cloud（NII RDC）を積極的に活用するなどして、先進的な研究データ管理を行うこととなっている。

このように、政策主導で進められている研究データの利活用の推進に対して、科学者コミュニティの主体的な関与がまだ少ないことが課題である。特にオープンサイエンスが目指す知の開放による新たな価値発見やデジタル・ネイティブな知識体系の構築を推し進めるため、また、研究データの流通およびその基盤に着目した、より健全な学術情報流通と研究評価の仕組みを生み出すために科学者の主体的な関与は欠かせない。日本学術会議は、学協会を含む科学者コミュニティを先導する担い手として政策関係者との対話を繰り返すことで、科学者の最小限の負担で研究データの利活用を推し進め、国際

[7] データの特性から公開すべきもの（オープン）と保護するもの（クローズ）を分別して公開する戦略。

的にもまだ具体の像がみえていないオープンサイエンスパラダイムの構築に主体的に取り組むことが求められている。

2.3　データと学術情報の流通

我が国の学術研究の情報流通は、この 15 年ほどの間に書籍、論文誌等の紙媒体から、電子的資料によるインターネットを介した流通への転換が進み、特に自然科学系ではほぼ電子ジャーナルに置き換わっている。この過程で、世界規模の学術出版社による電子ジャーナルの寡占が進むとともに、購読費用の問題が顕在化し、大学の財政基盤および研究費の圧迫、ひいてはパッケージ契約を中心とした大規模な購読中止による学術情報流通の崩壊に直面する事態に陥っている。一方、世界的には学術界と出版社の学術情報流通をめぐる主戦場は、プレ、ポスト論文のオープンアクセス、データジャーナル出版、研究のライフサイクルの源流となる研究データの占有に移行している。この状況に呼応し、G7/G8 科学技術大臣会合では次世代の科学技術を担う施策として、研究データの共有・公開に向けたコンセンサスの形成と方針の明確化が進められてきた。

学術情報流通の仕組みは、歴史的には研究者が自発的に貢献して構築してきたものである。しかしながら、研究と流通の分業が進む中で、学術論文データベースに基づく個々の研究者や研究機関の評価が重視されるようになるとともに、学術団体や出版社の活動が世界の学術市場の覇権を巡って競争が激化する状況にある（付録図表２参照）。また、国や研究助成団体による研究評価指標の過度な利用および産業界の危機感が、学術の未来への投資ではなく、短期的研究課題と成果を重視する雰囲気を作り出している。このように、研究者、研究機関、学術分野、国が分割統治され、学術の本質からはほど遠い競争と搾取にさらされている（付録図表３参照）。今こそ人類のための、データ取得を含めた研究活動そのものに価値を置いた評価が望まれる。

現在、我が国の学術界は、個々の研究者や研究機関の個別努力による出版社への対応を脱し、学術研究の理念に基づく明確な将来像を示すことが求められる。欧州で進められている OA2020 等による新しい雑誌契約は、従来の購読モデルからオープンアクセス出版モデルへの転換を足がかりとした研究成果のオープン化への動きである[25]。複数の国での大規模パッケージ購読の契約中止、米国カリフォルニア大学連合による商業誌編集プロセスからの撤退宣言は、研究活動とその学術流通の主権の学術界への回復を目指している。我が国においては、購読モデル維持の中でリポジトリ基盤やデータ管理のための基盤の構築が進み、オープンリサーチデータ[8]環境の基盤が整いつつある。オープン化は学術活動に抜本的な変容をもたらす。研究者の所属や環境で制約されることが少ないオープンな学術情報流通基盤へと発展させ、大学や研究機関、研究助成機関が明確にその方向へ向かう意思表示をすることにより、国際的な学術環境の中で我が国の学術の寄与を一層拡大していくことが望まれる。

---

[8] オープンデータとは、誰でも自由に使え、再利用でき、かつ誰でも再配布できる公開されたデータのことを指す。国や地方公共団体がもつデータを民間に公開する動きが有名である。それと区別するために、研究データの公開を指す場合は、オープンリサーチデータという。

我が国は、学術情報流通において、研究者が商業出版に搾取される状況の改善に向けた根本的な対応が求められている。オープンアクセスによる流通容易性を謳うAPC[9]価格の高騰の抑制、商業出版のピアレビューへの対価支払いなど、研究者と研究を守る活動では、データを含む学術情報流通の在り方の見直しが迫られている。オープンサイエンスの枠組みにおいて、我が国のデータプラットフォームを利活用してくための方針を明確にし、商業出版を含む学術流通の場において、先導的に国の代表機関による交渉と諸国との連携を行うことが不可欠である。

３　各学術分野における状況

2014年以降、日本学術会議からオープンサイエンスや研究データに関する多くの提言が表出されている。その後、多くの研究分野においては、データ駆動型の研究への変容が急速に進み、研究環境を含めて新たな理解と取り組みが喫緊の課題となっている。本提言をまとめるにあたり、すべての学術分野を俯瞰することが難しいことから、研究データやデータ駆動科学に関連して個別学術分野で起こっている動きから、研究現場で生じている新たな課題の把握に注力する[10]。その上で、オープンサイエンスの重要課題の措定を試みる。

3.1　材料系、化学系、製薬系

材料系

新規物質の開発においては、画期的物質系が現れるたびに学問体系が変わってしまうことが生じる。その学問の発展には、サイエンスとして研究データの公開が不可欠となる。現在の学術界が前提とする論文による研究成果の優先権（あるいは先取権）の確保と、査読に基づく妥当性の検証は、このような学問の転換点において成立しなくなる場合がある。また、論文編集システムが正当性の拠り所とする研究の審査過程で、研究コミュニティの研究倫理が崩れることが生じる。ここに寡占を極める学術誌および出版社が扱うことができない、逆に着目すべき研究成果の公開を改革する手がかりがある。

一つの例として高温超伝導材料開発では、研究データの公開と検証が論文誌による査読、出版において、論文が受理された日ではなく、プレプリントサーバarXiv.org [26]に登録された日が尺度とされた例が見られる。このオープン化が、研究者の論文投稿と研究成果への優先権を守ることとなっていた事実がある。

成功した結果や新規性を主張点とする既存ジャーナルは、ネガティブデータ[11]に価値を置かない。しかしながら、それらのデータの公開にオープンサイエンスとしての要点がある[27]。材料系の試みの一つに、超伝導材料のネガティブデータのデータベースがあり、共通の測定技術、評価手法で条件を明確にしたデータの公開は、新しい材料探索にも寄与することが確認された。オープンサイエンスのデータ戦略において、隔離した

---

[9] Article Processing Charge（論文掲載料）の略。論文をオープンアクセス誌などに掲載する際に必要となる費用。

[10] 各系の動きは、特徴的なケースのヒアリングに基づいており、その系の全ての考え方を必ずしも代表しているわけではない。ここでは、その意見から見出される一般論への手がかりを記載するに留めている。

[11] 研究の場においては、通常は論文で報告されない作業仮説に反するデータ。

ネガティブデータの公開は、次世代の研究者育成およびオープンな競争を生み出すことができる。この位置づけの明確化や研究者への教育が、材料系においてもますます必要になるであろう。

化学系

化学分野は、研究データの集積と共有では長い歴史を持つ非常に進んだ分野である。米国化学会の Chemical Abstracts Service（通称ケミアブ）は、1907 年に始まる論文中の情報収集、体系化を経て、1.3 億件を超える化学物質に関するデータを共有している。このデータを検索できる SciFinder [28]は、アカデミアだけでなく企業にも必須のツールとなっている。我が国の日本化学会も 1968 年以降化学情報協会を通じてこの情報収集に国際的に協力してきた。

その他に、結晶学や無機材料系の国際的なデータベースもあり、既知の物質か否か、あるいは、既知だと思われている物質の中で新しい特性が見つかったとき、まずはこれらのデータベースの情報との比較が、新規性を判別する拠り所となる。これらのツールで共有される物質群は、論文に条件が書かれている範囲での信頼性が担保されていることが必須である。

このような国際的な基盤が整っている状況において、我が国だけで何か物質情報を単に収集・整理する意味は見いだせない。しかしながら、安全性などの物性やその他属性データについて信頼性が担保されたデータベースを作ることには非常に意味がある。例えば、我が国では物質・材料研究機構（NIMS）が材料科学の信頼性のあるデータ収集と複合的なサービス開発を積極的に推進していることから、我が国におけるデータベース整備のあり方として参考にすべきであろう。

製薬企業は、国内での研究開発と研究所運営を中止する方向にあり、今後データの入手自体に資金投入する可能性がある。このとき、公的および企業のデータの相互の流通の可能性が生まれる。現在データを持っている主体である各大学や企業においては、データが相互にリンクされておらず、また、その流通をコントロールできる権限がない。投入する資金などのリソースとの関係に依存した構図をどのようにするかが課題である。一方で、情報科学と化学の融合を目指し、人間の作業に基づく経験則づくりや、理論に基づいた新規物質開発を、AI[12]とロボット化で効率化するためのデータ利用が加速化している。我が国としても、こうした次世代の研究活動への後押しが必要であると考えられるが、欧米と同じことを日本が繰り返すだけでは優位性を確保することは難しい。この課題に答える方策は現時点で明確ではなく、今後、日本の特色を踏まえて強化すべき道筋を冷静に考える必要がある。

製薬系

化学分野、創薬化学分野で、特にデータのオープン化に対する取り組みが積極的にな

[12] Artificial Intelligence（人工知能）の略。推論・判断などの知的な機能を人工的に実現するための研究。また、これらの機能を備えたコンピュータシステム。

されている背景には、単にデータを共有化するだけでなく、機械学習が律速段階をより早く、より安く、より正確に、あるいは今まで専門家だけがやっていたことをいわゆる初心者にも可能にし、既存プロセス自体を根本から変える試みが米国のベンチャー企業などで進められていることがある。病気の原因を病理学的な判断で見つけ、目標を決め、それに対して評価系を立て、効果のある化合物を最適化し、動物への効果、毒性などを確認後、人への投与を行うには、大変な時間とプロセスを必要とする。この過程を、AIと組み合わせたDe novo designと呼ばれる手法により、既存の同種の薬品の化合物構造からの誘導体ではなく全く新たな構造をもち、しかも動物の病態モデルにおいて効果がある化合物を直接見いだせるようになってきた。

我が国では、化学界が必要とする信頼性の高いデータベースと、それを構築し維持する環境が重要視されていない。そのため、個々のデータベースはともかく、それらの統合に至る道筋が見えない。データベースを統合する統括機関がないことで、データベースの重複や関連付けが困難となっている。結果的に、公共のデータという認識が醸成されていない。一方で、論文執筆と関係なく、人が使いやすいデータを提供するというデータサイエンティストの育成もままならず、製薬系としても戦略的な取り組みが不可欠である。さらには、結局ランダムに研究を推進していることにより、メガトライアル[13]をすると研究推進に有効なデータが他の情報に隠れてしまうという矛盾を生んでいる。

データが支配する世界の中での存在感を確保することが不可避であって、日本としての価値のあるデータを持っていくことにより、世界の公的データベースを使えるオープンリサーチデータ環境を維持し進めていく必要がある。また研究者は、データを登録し公開しなければ、論文発表で研究を完成し、優先権を維持することが困難になるという意識を、学会内で高めていくことが重要である。さらに、ネガティブデータを収集し、それも合わせることで、効率的に目的にたどりつける可能性は製薬系においても多分にあり、そのためにデータの共有を促す環境を、論文を課さない専門職としてのデータキュレーター、データサイエンティストの育成が急務となっている。

3.2　生物系、医療系

生物系

生物学の最近の目覚ましい発展はオープンリサーチデータ抜きでは語れない。ヒトをはじめとした様々な生物のゲノムデータはもとよりオミクスと呼ばれる遺伝子発現、タンパク質、代謝産物などの網羅的なデータの幅広い共有が基礎生物学を大きく前進させてきた。その結果、生物学はこれらのデータを基盤としたデータ駆動科学へと変貌を遂げつつある。オープンサイエンスの波は、基礎生物学に止まらず、医学、薬学、農学などの応用科学にも押し寄せてきており、生命科学やバイオ産業を大きく変えようとしている。収集されたデータは医薬品やバイオ素材などの研究開発に利用されるだけでなく、それらのデータの速やかな共有によって、感染症の検出・モニタリング、感染源特定な

[13] 大規模なランダム化比較試験のこと。ランダム化比較試験とは、対象の集団を無作為に複数の群に分け、その試験的操作の影響や効果を測定する比較研究。

ど、ヒトや動植物の感染症対策や食品安全にも利用されている。また、多種多様な生物を識別できるデータが蓄積されているため、生物種の特定、外来生物の検出など、生物資源の保護にも利用されている。このように、生命科学分野のオープンサイエンスはすでに大きな成果を生み出しており、それらを推進するための国際的な提携なども多数できており活発に活動している。

しかしながら、データそのものが大きな価値や意味をもつだけに、データ共有にはまだまだ障害も多い。例えば、詳しくは後述するように医学に応用するには臨床情報などのいわゆる表現型データが欠かせないが、これらは個人情報であることもあり、共有があまり進んでいない。また、ヒトゲノムそのものも個人情報であるため、法的、技術的に様々な対応策が必要である。一方、生命科学分野ではデータだけでなく、ヒト試料（バイオバンク）やその他生物の試料（バイオリソース）の保存・共有も国際的に進みつつある。ヒト以外の試料に関しては生物多様性条約（CBD）[29]が定める「遺伝資源の利用から生ずる利益の公正かつ衡平な配分」を義務化した名古屋議定書[30]があるが、近年途上国側は、この利益配分を遺伝子配列等のデジタル情報にも拡張すべき、と主張している。このように、生物系は国際条約などの法的な側面を考慮したデータのオープン化の検討が避けられない。

医療系

医学は人間の生老病死を研究し、医療はその情報をもとに実践する学術である。分析技術の進歩により膨大なデータが産生される時代となったが、その臨床的・社会的意味を明らかにするためには、巨大データベースの構築とデータ共有が必須である。

これまで医学・医療の進歩は、多くの疾患を克服してきた。しかし癌や心血管病などの慢性疾患の克服には長期間の観察が必要であり、さらに集団の状況の把握だけでなく、個別の患者に対する予測が求められる。そこでゲノムをはじめとするオミクス情報、健診・検査データ、日常の身体データ、環境因子などを統合し、医療をできるだけ個別化しようという動きが生まれた。これが Precision Medicine[14]である。

治療法の厳密な評価には、一般にはランダム化介入研究[15]が必要であるが、数十億円規模の研究費と数年間の期間が必要となり現実的でない。このため「次善の策」としてビッグデータ[16]解析が注目されている。その分析結果は、介入研究の倫理的根拠ならびに科学的基礎データとしてだけでなく、医療資源の配分などの医療制度の面からも重要である。

医療ビッグデータは産業的にも大きな価値を生む。薬剤の有効性が実証されれば、その医薬品は世界の巨大市場を獲得することができるからである。医療ビッグデータは、

---

[14] 精密医療、個別化医療、先制医療と訳される場合がある。

[15] ランダム化比較試験と同義語。

[16] 近年の ICT（情報通信技術）、特にセンサーの飛躍的発展によって、地球物理、気象、地震、天文、生命科学、マーケティング、ファイナンスなど多くの研究分野や社会で出現した大量・大規模のデータ。

自国の産業を守るためにも重要である。曖昧な根拠に基づいて批判されたときに、これに反論するためである。

しかしながら、我が国では、ビッグデータの構築が困難な面がある。例えば、電子カルテは主要ベンダーが診療データを囲い込んでいるため、データ互換には膨大なコストが必要である。今後データの標準化研究をさらに進める必要がある。NII では、日本医療研究開発機構（AMED）からの支援を受け、医療系６学会（放射線学会、消化器内視鏡学会、病理学会、眼科学会、皮膚科学会、超音波学会）と連携して医療画像データプラットフォームの構築と画像認識 AI の開発を進めており、画像データをいかに効率よく集めるかが最重要課題となっている。また、SINET5 [31]等の高速なネットワーク環境を利用することでこれまで活用されて来なかった CT のような巨大な画像が収集可能となり、１億弱の集積された画像データを利用して開発された AI は高い成果をあげている。既に病理学会は福島県で実証実験を済ませ、眼科学会の取り組みなど社会実装を検討しているものもあり、このように ICT の発展と共に多種で膨大なデータ群を得ることで、医療系においても、世界的にみて強力なデータ駆動科学に対応する研究環境が構築可能となる。

3.3　地球科学系、農業系、生態学系

地球科学系

地球科学系は、太陽惑星系を含み地球惑星科学とも称する総合科学分野である。環境科学や自然災害の自然的背景を理解する基礎科学としても位置付けられる。対象によって宇宙惑星科学、大気海洋水圏科学、固体地球科学、地球生命圏科学、地球人間圏科学として分類されている[32]。この各分野において国際的に展開されるビッグデータの国際的統一基準による管理と運用は、科学推進のために欠かせない。

ビッグデータを国際的に管理するためには、国際的なデータ記載、そのアーカイブ管理、公開原則など、細部に至る国際的合意形成が大きな課題である。地政学的事情も考慮した調整など科学的データ管理に止まらない事柄については、G7/G8やG20、国際連合などの積極関与が避けられない。例として宇宙天気にも絡む超高層データ、各国領空海も絡む大気現象データ・海洋物理化学データ・採集試料の国際共有、海洋底物理的地質学的データ・採集試料の国際管理、地下生命圏など極限環境下の生物学的試料とデータなどがある。国際基準に基づく研究機関における管理分担も明確にする必要がある。日本の地球科学系各分野それぞれにおいてオープンデータシステムの構築のために積極的な基準作りなどにおいて国際貢献し、拠点形成を促していく必要がある。

地球科学系では、観測される数値データとしてのデジタルデータのみならず、研究に供された第１次物質的試料（岩石、堆積物、土壌、流体、生物など）の永久保存が、研究結果の検証と将来における研究手法の発展に伴う再研究のために欠かせない。500 年に及ぶ近現代科学の発展を進めてきた西欧諸国に比し、西洋産業革命以降に合流した明治以降の 150 年程度の短い歴史しかない日本においては、永久保存体制の脆弱性は歴然としている。デジタルデータのみならず、研究成果をもたらした第１次物質的試料の永

久保存体制の構築、それらの背景となった第０次試料の選択的保存について基本方針を確立する必要がある。また、そのための設備、整理と公開のための体制と方針を確立し、全国連携体制で効率的な施策を確立することが望まれる（国際深海掘削の例：付録図表４～６参照）。国際深海掘削計画では、世界の３箇所で運用されているデータ・試料永久保管施設（付録図表５参照）の経費や利用状況が参加国間で共有され、効果的な運用が実現されている[33]。

農業系

精密農業とは、作物や生育環境に関する高解像の時空間情報を収集し、環境保全と収益向上を同時に実現するための営農管理体系であり、1990 年代に世界でほぼ同時に始まり、現在では先端技術を利用したスマート農業やデータ駆動農業ともいわれている。精密農業の特徴は、ばらつきの存在（付録図表７参照）とばらつきの管理（付録図表８参照）にあり、従来の平均値や偏差値による農学や農業施策と異なる。ばらつきの要因は、地質学的な土壌形成や農地作成の長周期履歴、長期・短期の農作業履歴、気象環境、作物の遺伝的特性、雑草や野生動物の履歴など多様である。

精密農業を実行すると、個々のほ場から膨大なデータが創成される。例えば、可変施肥田植機による田植え作業の場合、RTK-GPS（Real Time Kinematic GPS）によるほ場内の位置と肥沃度および施肥量のデータが得られる。空間解像度はセンチメートルのオーダで、時間解像度は秒のオーダである。

日本にはおよそ 4300 万枚のほ場が農地台帳に登録されており、最近、デジタルデータとして閲覧可能な状態になった。これらのほ場に精密農業データが蓄積されるようなことになると、生産計画や地域計画に有用な農地の価値が顕在化し、ベンダーロックイン[17]の不安や地域を越えた共通語の未整備などのため、大規模な混乱が予想される。そこで、農作業用語の共通化やデータ形式の共通化などの政府事業が進められ（付録図表９参照）、ユーザーである農家の要望を反映したモデル事業も行われた（付録図表 10 参照）。農業生産現場のみならずフードチェーン全体を対象にした食の安全保障のためのデータ共有化の取り組みが分野横断・業界横断で推進される必要がある。

生態学系

生態系は多様な生物と物理的な環境との相互作用によって成り立っており、全体としてのエネルギー収支や循環システムは算定できるが、それぞれの種間、種内の個体間の関係を読み解くことは難しい。そのため、個々の種の働きは個別に実験室で再現し、それを野生下に想定してシステムをモデル化する方法を取っている。この際は、それらの実験がオープンにされて検証可能なものであることが必要となるが、野生下の現象は複雑で再現不能である場合が多いので、細部にわたる予想を立てることは難しい。しかし、地球環境が悪化の一途をたどるなかで、生物多様性や絶滅危惧種の算定、さらには生態

[17] 特定のベンダー（メーカー）の製品や独自技術を使って利用者を囲い込み、他社の製品やサービスへの切り替えができなくなる状態。

系サービスといった数値に基づく評価が重要性を増している。環境保全に対する取り組みが国の政策、さらには複数の国にまたがる国際関係をも左右するので、生態学的な指標の役割を再認識する必要がある。とくに、最近は地球環境の変動予測に生態学的指標（昆虫や海洋生物の動向）が用いられることが多くなっている。信頼できるデータに基づく分析が指標づくりには求められており、国際的なプラットフォームによるデータ共有が不可欠となっている。

動物行動学では、近年観察者の恣意が入らないデータの取り方が推奨されるようになった。例えば、データロガーを動物の体に装着して、その動物の位置情報や活動を逐一計測する方法が取られるようになっている。また、これまでは調査者が観察して得たアナログ情報をデジタル情報に変換し、シークエンスとして分析する手法がとられていたが、映像を資料として用いることによってより検証しやすくしている。また、糞などの試料から抽出したDNAを解析し、個体群における繁殖構造を分析して、個体の移動を推測したり、毛髪や歯石から安定同位体を抽出して過去の食性を調べるなど、マクロな分野にミクロな分野の手法が導入される傾向も顕著になっている。

動物行動学や霊長類学では、特異事象のケーススタディでは統計分析ができるほど多くのデータを集められないことがあるため、客観的な記録データとしてのビデオを証拠として提出する傾向にある。また、訓練を積んだプロの研究者だけではなく、一般の人々が撮影しYouTubeなどに掲げた画像や映像を資料として分析する手法も登場しつつある。シチズンサイエンスとしてのオープンサイエンスの重要性が高まっている。以上の通り、同分野は、データとその利活用による研究活動の変化が最も進んでいる分野の１つである。

3.4　人文学系、経済学系、社会科学系

人文学系

人文学分野では、この20年ほどで、著作とその草稿、文書・記録などの基盤的人文学データが急速に「デジタルアーカイブ」として共有されるようになり、かつ学術的使用に耐えうる高度な資料提供がデータベース等でなされるようになった[34]～[45]。こうしたデジタル化をうけて、「デジタル・ヒューマニティーズ」とよばれる分野が特に欧米で拡大し、デジタル化された大規模な学術データについて、人間の手作業では実現し得なかった、コンピュータプログラムによる分析によって、新たな知見が発見されるようになった[46]。さらにはそうした学術成果を学界や社会に提供する手段としても、情報学の技術や技法が応用され、これまでの紙媒体中心であった学術の形を大きく変え始めている[47]。日本では、資料や書籍のデジタル化、電子書籍普及の遅れ等によって、現在も個々の資料に対してデジタル化を進めつつ、オープンデータ化によるその高度共有利用の実現、あるいは解析ツール開発などを並行的に行っている段階である。しかし、データ利用条件を明示したオープンデータ化の流れのなかで、必要なデータを自由に入手し、それらを組み合わせて、新たな価値を発見するような動向が、日本の人文学でも

次第に拡がっている[18]。画像についてはIIIF（International Image Interoperability Framework）形式[48]の普及、メタデータ[19]についてはRDF（Resource Description Framework）などを含む機械可読形式による記述が標準化しつつあり、データが個々のシステムを超えて繋がり、研究に活用される状況が目指されている（Liked Open Data）。人文学にとって要ともいえるテキスト自体も、TEI（Text Encoding Initiative）[49]によって構造化され、関連データとのリンケージ形成がなされている。そのなかで、従来の大規模デジタルアーカイブを超えた連携検索用ポータルサイトの構築が様々な領域で進みつつある。その代表的存在は、内閣に置かれた知的財産戦略本部を中心として企画され、国立国会図書館が開発した日本の文化資源情報の総合的検索サイトをめざす「ジャパンサーチ」[50]であり、2020年1月時点では連携機関数21、連携データベース数61、メタデータ件数19,664,870件にとどまっている。他にも東アジア漢字圏の字形データを対象に国際連携検索サイトを目指す動きなど[51]、様々な試みが生まれており、それらが繋がる時、人文学研究の大きな変革が期待される[20]。

一方、デジタルデータが利活用されるほど、そのもととなった原物（原本）資料であり、古文書・書物・遺物などの重要性がより強く認識されるようになった。急速に進む技術革新によって、原物資料から取り出せるデータの質と量が拡大してきた事実は、原物資料が無限のデータを内在させていることを改めて示したからである。また、遺構、景観、言語、技術、生活様式など、時間とともに常に変化していく資料を過去に記録した媒体、例えばガラス乾板等の古写真や音声テープ、映像フィルム[21]なども、原物資料と同じく位置付けられる。こうした広い意味での原物資料を、組織的・体系的に永久保存していく確かな体制構築は、現在喫緊の課題といえよう。

経済学系

経済学は、方法論においてモデル学化、計量・数理化を強め、政策科学としての役割を重視しており、それゆえにデータの充実とその広範囲の利用は、経済学の発展、国際的な競争に不可欠の条件と見なされている。統計データは経済学研究にとってインフラの役割を果たしており、その整備・オープンサイエンス化の水準が当該国の経済学の研

---

[18] オープンデータでは利用条件の明示が重要だが、人文学系データベースにおいても、クリエイティブ・コモンズ・ライセンス表示が拡大している。またオープンデータをより有効に利活用するために重要なデータの機械可読性についても、パーマネントリンクの付与が推奨されるとともに、RDFなど の Web 標準技術を駆使して、Web 上に公開されるさまざまなデータを相互に関係づけるLinked Data化も進められている。

[19] データを利用するにあたって、データそのものではなく、属性等を示すデータ。

[20] 従来文字データを扱うことが多かった人文学は、数値データ中心の分野と比較して、ビックデータ解析は難しいとみなされていた。しかし、オープン化された大量の機械学習可能なテキストデータが生み出されつつある現在、人文学研究の中身自体も大きく変容していく可能性が指摘されている。例えば大量の歴史資料テキストデータについて、ビックデータ解析が可能な状況になると、人間が行うべき研究は、その解析によって発見された一見反復的な歴史事象群を、いかに異なる歴史的コンテキストに位置付け、差異を考察していくかといった、意味の解釈へと比重を移していく可能性がある。なお、こうした人文学の変化に対応するための大学・大学院レベルの教育プログラムは、日本においては十分に構築されているとはいえないが、東京大学大学院人文情報学拠点における人文情報学概論など、いくつかは存在し、オープンサイエンスとしての人文学研究を見据えた若手研究者も少数ながら育ちつつある。

[21] ここでは原物資料を永続的に保存していく体制構築に言及しているが、実際の保存や特に公開に際しては、肖像権の問題等についても十分に考慮する必要がある。

究水準・体制を規定するといっても過言ではない。特に、同一対象を継続的に観察した経年のパネル（コホート）データは緻密な実証・理論研究を可能にするものであり、公共オープンデータ化が世界的にも研究教育において重視されている。我が国では、データ構築が個別に行われ、科学研究費など変動的な外部資金に依存してきた経緯から、データ構築が断絶するリスクも大きく、そのことが日本経済研究の魅力を損なう要因にもなっている。また、日本学術会議がこれまで提言してきた「エビデンスに基づく政策研究」(EBPM)手法が政府（各省庁）において推進されており、2017 年統計改革推進会議が設置されている。このこともデータベースの構築と経済学におけるオープンサイエンス化を促進する契機になっている。

本分野に関連した多様なデータのうち、質量とも中軸を占める公的統計は「政府統計の総合窓口(e-Stat)」を通じて広く利用されている[52]。さらに、学術研究を目的として個票形式のミクロデータもまた統計法に基づいて、オンサイト利用施設で情報セキュリティーを確保した上で利用されている[53]。公的統計以外では、家計パネル調査、生活の質に関する調査や社会意識調査などが、アーカイブ機能を有する共同利用・共同研究拠点に相当する経済系研究機関や財団に保蔵され、公開されている。また、健康・医療情報と社会経済情報を統合した文理融合研究も実施されており、データ構築と利活用が新しい研究を生み出している。相対的にオープン化になじまない実験データや文理融合型データも構築・利活用されており、新しい研究領域の開拓も進んでいるが、研究データの保存・管理はオープン化の途上にある。

経済学は、人文社会系のなかで最も国際共同研究が盛んな分野であり、研究成果・研究データのデジタル化・保存、オープンアクセスが進んでいるが、研究データに関しては、個別データの分散化、データ分析の専門家の不足などの問題がある。今後、我が国の研究水準を高め、国際共同研究を拡大する上で、拠点間連携の強化などで共用資産として安定的にデータ保存・管理・活用を行うとともに、データの保存・利活用・オープンサイエンス化に対する研究者コミュニティでの意識向上、コンセンサスの形成が不可欠である。

社会科学系

政治社会現象に対する実証分析は年々、増加し、現在では国際ジャーナルに掲載される政治学や経済学、社会学の論文の約 70%がデータ分析によるものである。今後、さらに日本の社会科学を発展させるためには実証分析に必要なデータを誰もが利用できるデータ・アーカイブの構築が必須である。その理由として、①一流の国際ジャーナルへの投稿に際して、論文の追試が可能なデータ公開が条件となっていることが多い、②研究期間終了や研究者の退職に伴いデータが散逸し、研究費が非効率的に使われている、③市町村合併による自治体行政資料や一定期間経過後の判例など貴重なデータが散逸している、④各国が急速に学術の国際化やデータのアーカイブ化を推進する中で、日本に関連するデータが利用しにくいことなどがあげられ、海外における日本研究衰退の一因となっている。

さらに、⑤研究データの公開の推奨ならびに研究健全化のためのデータ保存義務化があげられる。これは近年、FFP（ねつ造、改ざん、盗用）等の研究不正事案が多発しているためであり、事後対応および事前防止のためにデータ保存義務化が求められるようになったことに起因する[54]。このため、自然科学や人文学では自然科学研究機構基礎生物学研究所 IBBP センターや人間文化研究機構国文学研究資料館などでデータの蓄積が進捗している。これに対して、日本では社会科学だけが大学共同利用機関法人を持たないために、これまで意識調査データを部分的に集める程度に留まり、包括的なデータ・アーカイブが構築されずにきた。

ここで、諸外国の現状を概観すると、米国のミシガン大学社会調査研究所[55]やハーバード大学データベース[56]、欧州の GESIS[57]や CESSDA[58]、ESRC[59]、アジアの中国統計局[60]や台湾・中央研究院 SRDA[61]がデータ集積と公開に積極的に取り組んでおり、日本を除く多くの国でデータ・イニシアティブ競争が行われている。これに対して、日本ではサーベイデータの集積はあってもアグリゲートデータを管理・可視化して誰もが利用できるデータ・アーカイブが存在しないために、データ・イニシアティブ競争の時代から日本が取り残されている。

このため、次の要件を満たす社会科学データ・アーカイブの構築が必要となる。まず、データ・アーカイブの要件として、①データ・ユーザーではなくデータ・サプライヤー、②単なるデータ集積ではなくデータシステムの構築、③多様なデータに対応できるプラットフォームの構築、④コンテンツの追加を容易にするシステム管理、⑤オープンアクセスに耐えるサーバーと回線の設置、⑥永久的な受け皿としての組織主体の一元化があげられる。つまり、データを作成、収集する泥臭い努力をせずに、システムや解析の開発といった構想を繰り広げても現実にコンテンツとなるデータがなければ「絵に描いた餅」で終わることになる。

4　データ基盤とガバナンス

4.1　国内の基盤整備の状況

2016 年に日本学術会議オープンサイエンスの取組に関する検討委員会が公開した「オープンイノベーションに資するオープンサイエンスのあり方に関する提言」では、日本としてナショナルレベルでの研究データ基盤を整備する必要性がまとめられた[62]。それを受けて、内閣府や文部科学省等の委員会でも多角的な検討が重ねられた。令和元年６月に閣議決定された統合イノベーション戦略 2019 では、基盤整備の主体としての NII の位置づけや開発する NII RDC の活用についても言及されている[24]。

研究データ基盤として整備が進められている NII RDC は、研究データのライフサイクルに沿った管理・公開・検索を支援する基盤から構成されている。研究プロセスの終盤にあたる研究データ公開や検索機能の整備のみならず、研究開始時のデータ収集や研究中の研究グループ内でのデータ管理に発展していることが大きな特徴となっている。従来は、オープンアクセスを推進するために、文献を主とした公開の場の「箱」をリポジトリ等で提供し、それらを横断検索できるサービスが提供されてきた。NII RDC では、

従来のサービスの研究データ対応に加え、研究中のデータ管理や共有を支援し、オープンサイエンスやデータ駆動型の研究スタイルをさらに加速することを主眼としている。

4.2　基盤整備の方向性

NII RDC は、現在大学や研究機関が組織として導入するための試行実験が進められており、2020 年度中から開始される本格運用の中で、基礎的なデータ基盤整備が具体的に実践されていくことになる。この優位性を最大限に生かし、日本の研究力の強化に繋げていくためには、利用機関と連携しながらさらなる機能の高度化を推進していく必要がある。

第３節では、オープンサイエンスやデータ駆動科学に関連したデータ共有の現状についてまとめたが、各分野において期待される今後の展開を基盤整備の視点から整理すると以下のようになる。

- 学術だけではなく産業界も利用でき産学連携を図れる基盤
- 従来の査読論文だけではなくプレプリントも含めた幅広い成果公開のための基盤
- 通常は論文で報告されない作業仮説に反するデータも含めたデータ共有を実現できる基盤
- 散逸しているデータやデータベースの情報を統合する基盤

こうしたニーズを満足するために機能開発が必要な場合もあるが、概ね現状の NII RDC が提供する機能により実現可能である。今後の課題としては、まず、2020 年度中に運用を開始する NII RDC の利用者コミュニティを、いかにして拡大していくかがあげられる。既存の JAIRO Cloud や実証実験が進められている研究データ管理基盤 GakuNin RDM は、大学や研究機関といった組織を対象として普及を進めている。この組織の軸を研究分野の軸へと転換していかなければ、各研究分野からあげられたようなニーズを解決していくことはできない。各研究分野としての活動を牽引する大学共同利用機関や学会等と連携し、新しい利用のための普及の展開が実現されていくべきである。

先進的な研究データ管理を実現するために基盤が実現すべき機能や役割については、欧米の先行機関の取り組みに見られるように発展途上にある。その中でも、データ解析システムや電子ラボノートといった周辺サービスとの連携機能や、研究プロジェクトにおける研究データのやりとりがさらに円滑になるツールの導入など、様々な模索が続けられている。こうした研究開発による機能の高度化は NII RDC にも不可欠であり、研究機関や研究分野との密な連携を介して、ニーズに基づいた方向性を定めていく必要がある。現在 NII と NIMS では、特定分野における研究データ基盤の機能開発の連携を進めている。こうした連携を他分野ともさらに展開することで、共通基盤としての NII RDC が高度化していくというサイクルの実現が望ましい。

4.3　ガバナンスの重要性

オープンサイエンスを志向した研究データ管理のインフラ整備としては、まず、以前より推進されてきた研究プロセスの終盤にあたる研究成果公開における研究データ整

備のさらなる拡充があり、それに加えて、研究開始時のデータ収集や研究途上における研究グループ内でのデータ管理に発展していることが現在の大きな特徴である。前者は、主に産業を含む幅広い研究データの利活用を促進し、その利活用を定量的に把握して研究インパクトの測定ならびに研究評価につなげることが見込まれる。一方、後者では、研究プロジェクト内での共有を進め、研究の効率化に加えて、新しい気づきや発想の機会を増やすことや、研究公正に役立つことも見込まれている。これらの違いを踏まえたガバナンスが重要となる。

さらに、各分野における事例からもわかるように、高度な研究データ基盤の整備とデータプラットフォームの構築は、純粋な研究力の強化に留まらず、イノベーションを生み出す機会を牽引し、新たな科学知の発見と産業が生まれることを加速することにも繋がっていく。また、研究開始時から幅広く研究活動のモニタリングも可能となり、研究データ駆動型の学術情報流通のエコシステムが醸成されることとなる。

このことは、研究データ基盤整備と研究データ流通における、研究機関、研究助成団体、国等の各組織のガバナンスが非常に重要になることを意味し、日本の研究者が生み出した研究データが適切に管理・共有・公開され、その貢献が国際的に認められる体制を開発していくことになる。また、不正競争防止法、個人情報保護法、著作権法等の法規制の再整備を進めることで、研究データを扱う作法の混乱を避ける必要もある。

公的データについては行政機関の縦割りなどが障害となっていて責任者は二次利用に慎重であるが、現場の責任者に負担を与えない制度作りが必要である。また、情報時代の日本の学術のあり方を、国際的なコンセンサスを得ながら議論することも重要である。そして、政府とアカデミアの双方から、データ流通におけるルール作りや標準化に向けた主体的な取り組みが求められる。

5　提言：オープンサイエンスの深化と推進

5.1　オープンサイエンスの深化と推進に向けて

オープンサイエンスの深化と推進のためには、大学を中心とした研究機関がどういうデータを共有し、どういうデータを守るかが重要になっている。例えば、大学はオープンな場であって、原則的に、全ての学位論文をオープンにしなくてはいけないという姿勢そのものが意思表示であり、大学というものの在り方に関わってくる。それゆえ、産業界と連携する場合には、秘匿性が高く非公開とすべきデータの取り扱いについて、関係者の合意によって守る仕組みを考えなければならない。

一つの参考にすべき例としてドイツの研究機関であるフラウンホーファーでは、大学とは違う産学連携で製品を生み出す組織を外付けにしている。その組織では論文の生産が目標ではなく、製品の生産や特許が目標にされている。したがって、その過程で出てくるデータ上の秘密は、そこでは守られる。しかしながら、大学では学生とともに論文を書く場合、論文が原則オープンである以上、論文の生産過程を秘匿することはできない。ここに大学の限界がある。

学問の分野を超えて科学のデータをどう共有するかについては、まだその分野では価値の分からないものに対して、他の分野の関心をどう誘引するかの問題になる。そもそも自然科学の中でも、データで表すことのできる現象、検証可能な現象の捉え方が、とりわけ生物学と物理学の間、あるいは生物学と社会学の間で大きく異なる。一方で、研究を抽象化した考え方やフレームワークは利用可能でもある。例えば、人間の社会学と動物の社会学のはざまで、ゴリラの研究成果を人間に利用するだけでなく、また、その逆も可能であって、用語もかなり重複している。これを、用語の問題だけではなくて、現象の問題として共有していくということがこれから起きれば、生物、社会だけでなく、いろいろなイノベーションが生まれる可能性がある。また、「エコシステム」という言葉は、元々は生態学で生まれた用語で、それを現在、組織や社会を論じる際にも広く使っているが、各々の内容は全く違う。これは、その定義が転用可能だからであり、エコシステムという言葉を通じて現象が共有されたことになる。しかるに、データに関しても、同様に重複しているものや転用可能なものがあり得る。特に、社会学のエコシステムに、生物学として使われているエコシステムの内容（データ）を適用して検証することを支援するデータベースやプラットフォームが望まれる。

現代の科学が細分化し、サイロ主義が進む中で、自身の研究分野で突出した評価を受けることに加えて、イノベーションの視点から、他の分野でその発見なり考え方が利用されたり、関心が持たれたり、引用されることも重視し、その活動を促していくことに意味がある。

言葉は時代とともに変化し、変化しながら本来の定義が変わっていく。そして、また元に戻ってみたときに、違う世界や価値を発見する。古文書や歴史の中に探ることも同様だと思われる。オープンサイエンスの本質は、いろいろな分野のデータと定義を考え直し、あるいは逆照射して新たな価値を認識する機会になる。より多くのデータ収集で精度を上げるという表面的な話だけでなく、新たな価値発見に向けてどのようにそのデータを構成し、共有するか、また経済性を保つかは、今の技術では明らかではない。少なくとも現時点で、データを安易に廃棄し、その価値を認めないという態度は、改めねばならない。オープンサイエンスは、単なる一過性の言葉では無く、しっかりとその本質と実態を踏まえ、今後どうすべきかの指針を示すことが重要である。学術だけではなく、同時に産業を中心とした学術に関わる各方面が動きだせば非常に大きな価値を持つ。まず、学術がそれを示さなければならない。

5.2　提言

令和２年の始まりとともに明らかになった新型コロナウイルス感染症（COVID-19）の世界的な感染拡大は、世界で初めてリアルタイムにウイルスの変異と感染のパンデミックを計測した事例ともいえる。この脅威に対して世界の国々が個別の対策を行い、同じ検査手法・統計尺度で議論できず相互不審を抱く中で、情報の共有化と適切な対策の実行が重要な要素となっている。国家間の人の往来によって我が国に到達した第二波の感染に対して、伝搬までの猶予を有効に生かすことができなかった。信頼できるデータが、

同一精度で、デジタル形式で、国家・民族の壁を越えて解析されることが、２週間の時間スケールで拡大する状況の中で、不可欠となった。この経験は、収束後に実データから洗い直すことが求められる。同様に、収束に至る過程をも、それぞれの対策の可否も含めて、オープンデータ化することが、抗体と匹敵する重要な資産となる。

ワクチン開発、抗ウイルス薬の探索・開発、感染メカニズムの解明、数理モデル、それらが全てオープンなデータで進められる状況により、あらゆる分野の研究者から企業の開発者が貢献できる世界が生まれている。さらには、過去の抗体との関係などのデータに基づく疫学的な議論から、当面の対策も実施されている。まさに、現場では信頼できるデータにより、多くの命を救う処置に繋げる活動がなされている。

2020年１月31日にNIH､AMEDなど世界の研究資金配分機関や論文誌出版社が協力し、COVID-19に関する論文投稿前に実験データをWHOに提供することの合意形成がなされたが、これはすなわちプレプリントが最前線における処置の実践データとして最も重要であることを示したことになる。感染症への対策の試行をリアルタイムで伝え、結果のデータを公開する機能は、既存の出版社を中心とするジャーナルが持ち得なかったものである。当然、今後のキュレーションを経て、論文化することが求められるが、データが共有化され証拠資料が示されなければ、論文に意味はなく、将来の同様の感染症の蔓延に対して有益な情報となり得ない。

我々はまさに、現実として経験しているこのような苦い状況を通じて、オープンデータとキュレーションを進め、管理運営する体制を取ることが求められる。そして、状況の収束が得られた暁には、過去から現在に至る人文学・社会科学における人類の営みの記録の蓄積と、資料のデジタルトランスフォームによって見いだされる新たな知を、人類の共通の財産とし将来につなげ、社会をより強固なものにすることが望まれる。

オープンサイエンスの深化と推進に関する検討委員会では、研究データやデータ駆動科学に関連して、個別学術分野や基盤整備に関する国内外の動向調査を進めてきた。各学術分野における研究データの管理や共有の状況には差異はあるものの、それぞれがもつ課題は、以下の３つの必要性に集約される。これらを、オープンサイエンスの深化と推進のための重要課題として取りまとめる。

(1)　データが中心的役割を果たす時代のルール作りの必要性

データを扱う作法の混乱を避けるべく、政府は、不正競争防止法、個人情報保護法、著作権法等のデータに関連する法規制を集約・整理し、データを安心して活用できるルールを、国際的なコンセンサスを得ながら明確化する必要がある。難病に関するデータや天文データ等は共有が進んでいるものの、商用価値の高いデータの効果的共有は重要であるが容易ではない。個人情報の取り扱いも明確にする必要がある。研究データの特性と社会との関係性を踏まえた着実な取り組みを実践するための、適切なガイドラインの作成も必要である。

(2) データプラットフォームの構築・普及の必要性

第３節でまとめたように、オープンサイエンスやデータ駆動科学が進んでいるとされる研究分野においても、より幅広い研究者に対し、膨大なデータを収集、キュレート、アノテート[22]、メタデータ付与、保存等が容易に実現されるデータプラットフォームが必要とされている。NIIでは現在、研究データのライフサイクルに沿った管理・公開・検索を可能とする基盤の開発を進めているが、これを学術界およびその研究活動を支える機関と協調しながら多方面で普及していく必要がある。第２節にまとめた欧米諸国の動向からわかるように、先進的なデータプラットフォームが備えるべき機能については、常に研究開発が進められている。単にファイルの共有や公開を実現するプラットフォームに留まらず、機械学習や高次解析との接続などを可能とする先鋭的な次世代機能を、利用者のニーズに即座に応えられる強力な開発体制を備えながら実現していくべきである。オープンサイエンスに期待されるイノベーティブな活動の一つとして、研究分野のサイロを超えた学際的な研究への展開があげられる。このプラットフォームでは、公開から非公開までのバランスが取れたデータ共有を支援し、分野間やセクター間など多様なデータの潜在的融合を許容してさらに高い価値を生み出すことを指向すべきであることも忘れてはならない。現時点においては、データを共有する意図があろうとも、経済的理由、あるいはスキルの欠如等からデータ保存がなされない事態も散見される。職場の異動時や定年時を含めデータ散逸を防ぐために誰でも容易に利用できるプラットフォームが必須である。

(3) 第１次試料・資料の永久保存の必要性

研究成果の基礎となる数値データを直接もたらした第１次試料（岩石、堆積物、土壌、流体、生物、物質、遺構、遺物など）、および数値データ化されてはいないが、人文学や社会科学における研究成果の基礎となった第１次資料（文書記録、書籍、景観、技術、生活様式、生産様式など）の永久保存体制の構築は重要である。将来の解析分析技術の進展を待って研究成果の検証に供することのできる試料･資料の保存は、科学サイクルの出口と入口に対するオープンサイエンス構築の鍵でもある。オープンサイエンスの進展が成果の出口、研究過程を軸に発展した後には必ず入口のあり方が問われることになろう。また、第１次試料・資料を抽出選択する背景となった第０次試料・資料（未研究の採集試料・資料）の選択的保存について基本方針を確立する必要がある。このことは大学・研究機関・博物館・資料館などにおいて保管庫確保が大変困難な現状にある中で急務である。また、保存のための設備、整理、管理運営と公開方針について国際・全国連携体制で進めることが重要である。

[22] データにそのデータの由来、測定条件、日時等の解釈に必要な情報を埋め込むこと。

＜参考文献＞

[1]European Commission: “Open innovation, Open Science, open to the world. A vision for Europe”, 2016 [Online]. Available: https://doi.org/10.2777/061652

[2]林和弘: “オープンサイエンスの進展とシチズンサイエンスから共創型研究への発展”, 学術の動向, Vol.23, No.11, pp.12-29, 2018.

[3]G7 茨城・つくば科学技術大臣会合: “つくばコミュニケ（共同声明）” [Online]. Available: https://www8.cao.go.jp/cstp/kokusaiteki/g7_2016/2016communique.html

[4]US National Research Council: “Bits of power, Issues in Global Access to Scientific Data”, 1997 [Online]. Available: https://www.nap.edu/catalog/5504/bits-of-power-issues-in-global-access-to-scientific-data

[5]OECD: “OECD Principles and Guidelines for Access to Research Data from Public Funding”, OECD Publishing, Paris, 2007 [Online]. Available: https://doi.org/10.1787/9789264034020-en-fr

[6]UK Research and Innovation: “Common Principles on Data Policy”, 2011 [Online]. Available: https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/

[7]Royal Society: “Science as Open Enterprise”, 2012 [Online]. Available: https://royalsociety.org/topics-policy/projects/science-public-enterprise/

[8]OECD: “Making Open Science a Reality”, OECD Publishing, Paris, 2015 [Online]. Available: https://doi.org/10.1787/5jrs2f963zs1-en

[9]David, P. A.: “The historical origins of ‘Open Science’:An essay on patronage, reputation and common agency contracting in the scientific revolution”, Capitalism and Society, Vol.3, No.2, Article 5, 2008.

[10]Pasek, E. J.: “Management Plan Requirements for National Science Foundation Grants: A Review”, Science and Technology Librarianship, No.87, 2017.

[11]California Digital Library ウェブサイト: “University of California Curation Center (UC3)” [Online]. Available: https://cdlib.org/services/uc3/

[12]Purdue University ウェブサイト: “Data Storage at Purdue: Purdue Supported Storage” [Online]. Available: https://guides.lib.purdue.edu/DataStorage

[13]Open Science Framework ウェブサイト: “OSF INSTITUTIONS” [Online]. Available: https://osf.io/institutions

[14]Harvard University ウェブサイト: “Harvard Biomedical Data Management” [Online]. Available: https://datamanagement.hms.harvard.edu/electronic-lab-notebooks

[15]Jones, S.: “Developments in Research Funder Data Policy”, The International Journal of Digital Curation, Vol.7, No.1, pp.114-125, 2012.
[16]Rice, R. and Haywood, J.: “Research Data Management Initiatives at University of Edinburgh”, The International Journal of Digital Curation, Vol.6, No.2, pp.232-244, 2011.
[17]openBIS ウェブサイト: “openBIS Laboratory Notebook & Inventory Manager” [Online]. Available: https://labnotebook.ch/
[18]AiiDA ウェブサイト: “AiiDA Automated Interactive Infrastructure and Database for Computational Science” [Online]. Available: http://www.aiida.net/
[19]European Open Science Cloud ウェブサイト: “EOSC Portal” [Online]. Available: https://www.eosc-portal.eu/
[20]National Health and Medical Research Council, the Australian Research Council and Universities Australia: “Australian Code for the Responsible Conduct of Research 2007”, 2007 [Online]. Available: https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2007
[21]Department of Education, Skills and Employment, Australian Government: “2016 National Research Infrastructure Roadmap”, 2016 [Online]. Available: https://www.education.gov.au/2016-national-research-infrastructure-roadmap
[22]内閣府: “我が国におけるオープンサイエンス推進のあり方について～サイエンスの新たな飛躍の時代の幕開け～”, 2015 [Online]. Available: https://www8.cao.go.jp/cstp/sonota/openscience/index.html
[23]内閣府: “第５期科学技術基本計画”, 2016 [Online]. Available: https://www8.cao.go.jp/cstp/kihonkeikaku/index5.html
[24]内閣府: “統合イノベーション戦略 2019”, 2019 [Online]. Available: https://www8.cao.go.jp/cstp/tougosenryaku/index.html
[25]The Open Access 2020 Initiative ウェブサイト: “OA2020” [Online]. Available: https://oa2020.org/
[26]Cornell University ウェブサイト: “arXiv.org” [Online]. Available: https://arxiv.org/
[27]Hosono, H., Tanabe, K., Takayama-Muromachi, E., Kageyama, H., Yamanaka, S., Kumakura, H., Nohara, M., Hiramatsu, H. and Fujitsu, S.: “Exploration of new superconductors and functional materials, and fabrication of superconducting tapes and wires of iron pnictides”, Science and Technology of Advanced Materials, Vol.16, No.3, pp.1-87, 2015.
[28]American Chemical Society ウェブサイト: “SciFinder” [Online]. Available: https://scifinder.cas.org/

[29]環境省自然環境局生物多様性センターウェブサイト: “生物多様性条約（日本語）” [Online]. Available: http://www.biodic.go.jp/biolaw/jo_hon.html
[30]外務省: “生物の多様性に関する条約の遺伝資源の取得の機会及びその利用から生ずる利益の公正かつ衡平な配分に関する名古屋議定書（略称：名古屋議定書）”, 2017 [Online]. Available: https://www.mofa.go.jp/mofaj/ila/st/page23_001940.html
[31]国立情報学研究所ウェブサイト: “学術情報ネットワーク SINET5” [Online]. Available: https://www.sinet.ad.jp/
[32]公益社団法人日本地球惑星科学連合: “公益社団法人日本地球惑星科学連合定款” [Online]. Available: http://www.jpgu.org/wp-content/uploads/2018/07/jpgu_teikan.pdf
[33]Texas A&M University, International Ocean Discovery Program ウェブサイト: “IODP Publications” [Online]. Available: http://iodp.tamu.edu/publications/reports.html
[34]国立国会図書館ウェブサイト: “国立国会図書館デジタルコレクション” [Online]. Available: https://dl.ndl.go.jp/
[35]国立公文書館ウェブサイト: “国立公文書館デジタルアーカイブ” [Online]. Available: https://www.digital.archives.go.jp/
[36]国立公文書館アジア歴史資料センターウェブサイト: “国立公文書館アジア歴史資料センターデジタルアーカイブ” [Online]. Available: https://www.jacar.go.jp/
[37]国立国語研究所ウェブサイト: “日本語コーパス” [Online]. Available: https://pj.ninjal.ac.jp/corpus_center/chj/
[38]国文学研究資料館ウェブサイト: “日本古典籍総合目録データベース” [Online]. Available: https://base1.nijl.ac.jp/~tkoten/
[39]国立歴史民俗博物館ウェブサイト: “国立歴史民俗博物館資料・データベース” [Online]. Available: https://www.rekihaku.ac.jp/education_research/gallery/index.html
[40]東京大学史料編纂所ウェブサイト: “史料編纂所データベース” [Online]. Available: https://wwwap.hi.u-tokyo.ac.jp/ships/
[41]奈良文化財研究所ウェブサイト: “木簡庫” [Online]. Available: http://mokkanko.nabunken.go.jp/ja/
[42]京都府立京都学・歴彩館ウェブサイト: “東寺百合文書 WEB” [Online]. Available: http://hyakugo.pref.kyoto.lg.jp/
[43]東京大学大学院人文社会系研究科次世代人文学開発センターウェブサイト: “SAT 大藏經テキストデータベース” [Online]. Available: http://21dzk.l.u-tokyo.ac.jp/SAT/index.html
[44]情報・システム研究機構人文学オープンデータ共同利用センターウェブサイト: “人文学研究データリポジトリ” [Online]. Available: https://codh.repo.nii.ac.jp/

[45]国立文化機構ウェブサイト: “国立博物館所蔵品統合検索システム” [Online]. Available: https://colbase.nich.go.jp/
[46]北本朝展, カラーヌワットタリン, 宮崎智, 山本和明: “文字データの分析—機械学習によるくずし字認識の可能性とそのインパクト—”, 電子情報通信学会誌, Vol.102, No.6, pp.563-568, 2019.
[47]国立歴史民俗博物館監修　後藤 真, 橋本雄太編: “歴史情報学の教科書　歴史のデータが世界をひらく”, 文学通信, ８章, 2019.
[48]IIIF Consortium ウェブサイト: “International Image Interoperability Framework” [Online]. Available: https://iiif.io/
[49]Text Encoding Initiative ウェブサイト: “TEI: Text Encoding Initiative” [Online]. Available: https://tei-c.org/
[50]デジタルアーカイブジャパン推進委員会・実務者検討委員会ウェブサイト: “ジャパンサーチ（試験版）” [Online]. Available: https://jpsearch.go.jp/
[51]井上聡: “新たな字形連携データベースの構築について”, 画像史料解析センター通信, Vol.88, pp.6-7, 2020.
[52]総務省統計局ウェブサイト: “政府統計の総合窓口（ｅ－Ｓｔａｔ）” [Online]. Available: https://www.e-stat.go.jp/
[53]独立行政法人統計センターウェブサイト: “統計センター業務案内” [Online]. Available: https://www.nstac.go.jp/services/index.html
[54]文部科学大臣決定: “研究活動における不正行為への対応等に関するガイドライン”, 2014 [Online]. Available: https://www.mext.go.jp/b_menu/houdou/26/08/__icsFiles/afieldfile/2014/08/26/1351568_02_1.pdf
[55]Institute for Social Research, University of Michigan ウェブサイト: “ISR Data Science” [Online]. Available: https://datascience.isr.umich.edu/
[56]Harvard University ウェブサイト: “Harvard Dataverse” [Online]. Available: https://dataverse.harvard.edu/
[57]Leibniz Institute for the Social Sciences (GESIS) ウェブサイト: “GESIS Data Collection” [Online]. Available: https://www.gesis.org/en/services/data-collection
[58]Consortium of European Social Science Data Archives (CESSDA) ウェブサイト: “DESSDA Data Catalogue” [Online]. Available: https://datacatalogue.cessda.eu/
[59]Economic and Social Research Council (ESRC) ウェブサイト: “UK Data Service” [Online]. Available: https://esrc.ukri.org/research/our-research/uk-data-service/
[60]National Bureau of Statistics of China ウェブサイト: “National Data” [Online]. Available: http://data.stats.gov.cn/english/

[61]Center for Survey Research, RCHSS, Academia Sinica ウェブサイト: “Survey Research Data Archive” [Online]. Available: https://srda.sinica.edu.tw/index_en.php
[62]日本学術会議オープンサイエンスの取組に関する検討委員会: “提言 オープンイノベーションに資するオープンサイエンスのあり方に関する提言”, 2016 [Online]. Available: http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-t230.pdf

＜参考資料１＞審議経過

平成 30 年

12月18日　オープンサイエンスの深化と推進に関する検討委員会（第１回）
役員の選出、本委員会設置について、オープンサイエンスに関する意見交換

平成 31 年

２月５日　オープンサイエンスの深化と推進に関する検討委員会（第２回）
話題提供（梶田将司氏）

３月 28 日　オープンサイエンスの深化と推進に関する検討委員会（第３回）
話題提供（澁澤栄委員、青木学聡氏）

４月５日　オープンサイエンスの深化と推進に関する検討委員会（第４回）
話題提供（木村学委員）、報告書構成と執筆分担に関して現状と課題

令和元年

５月 16 日　オープンサイエンスの深化と推進に関する検討委員会（第５回）
話題提供（細野秀雄氏、林和弘委員、赤池伸一氏）

６月７日　オープンサイエンスの深化と推進に関する検討委員会（第６回）
話題提供（久留島典子、上村みどり氏）

６月 24 日　オープンサイエンスの深化と推進に関する検討委員会（第７回）
話題提供（溝端佐登史委員、川合眞紀氏）

７月 24 日　オープンサイエンスの深化と推進に関する検討委員会（第８回）
話題提供（山地一禎氏）

９月４日　オープンサイエンスの深化と推進に関する検討委員会（第９回）
話題提供（永井良三委員）、提言とりまとめについて

10 月 31 日　オープンサイエンスの深化と推進に関する検討委員会（第 10 回）
話題提供（山極壽一氏、小林良彰氏）

12 月 13 日　オープンサイエンスの深化と推進に関する検討委員会（第 11 回）
提言構成及び内容検討

12 月 26 日　オープンサイエンスの深化と推進に関する検討委員会（第 12 回）
話題提供（高木利久委員）

令和２年

３月２日　オープンサイエンスの深化と推進に関する検討委員会（第 13 回）
提言「オープンサイエンスの深化と推進に向けての提言」（案）
学術フォーラム「オープンサイエンスの深化と推進に向けて」企画書

５月 21 日　オープンサイエンスの深化と推進に関する検討委員会（第 14 回）
提言「オープンサイエンスの深化と推進に向けての提言」（案）の修正
学術フォーラムのテーマ変更について

５月 28 日　日本学術会議幹事会（第 291 回）
提言「オープンサイエンスの深化と推進に向けての提言」の承認

＜参考資料２＞日本学術会議主催学術フォーラム開催

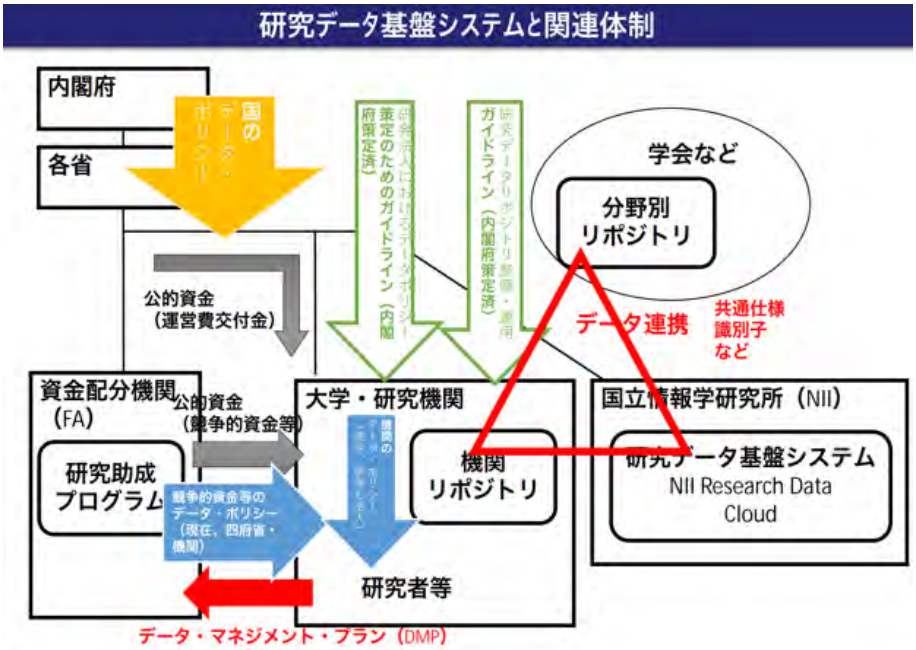
テーマ「COVID-19 とオープンサイエンス」
企画趣旨
　学術の成果をオープン化して広く共有することにより、研究の進展を加速し、学術的知見の導出の拠り所となる研究資料・データと研究成果の再現性を高めることを目的とした「オープンサイエンス」の方向性が世界的に注目されている。世界的なウイルス感染の拡大の中、オープンサイエンスの深化と推進の方向性を議論し、新しい科学の推進の在り方を問う。

日時　　：2020 年 6 月 3 日（水）10:30-16:30（講演時間は変更の可能性あり）
開催地　：インターネット公開

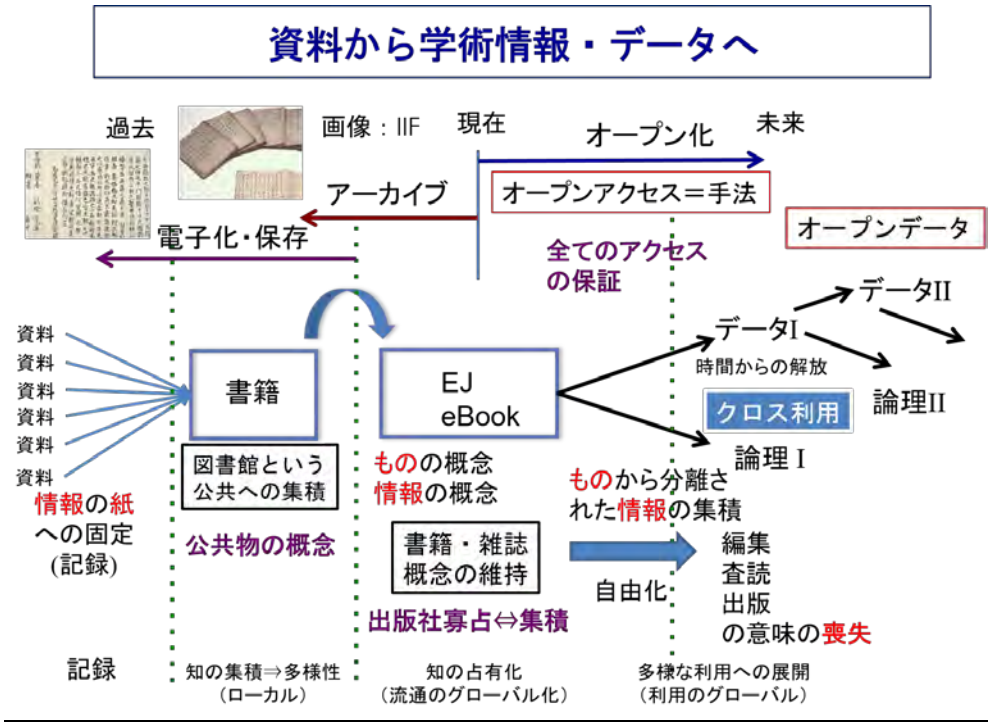
プログラム

10:30	開会宣言
10:35-11:00	喜連川　優（国立情報学研究所所長・東京大学教授） 問題提起
11:00-11:25	Prof. Wen Gao (Peking University, Boya Chair Professor and Dean of EE&CS School and Pengcheng Labs) 「Open Source is going to be a footstone of information Industry in China」
11:25-11:50	安浦　寛人（九州大学理事・副学長） 「教育データの収集と分析ー遠隔講義を契機にしてやるべきこと」
― 休憩 ―	
13:00-13:25	山極　壽一（日本学術会議会長・京都大学総長） 「フィールドワークから見た COVID-19 感染拡大に関して」
13:25-13:50	末松　誠（慶應義塾大学医学部医化学教室教授） 「医学領域のデータシェアリング: One for all, all for one」
13:50-14:15	木村　学（東京海洋大学特任教授・東京大学名誉教授） 「パンデミクス時代を超えてー地球惑星科学の国際公開連帯」
14:15-14:40	小池　俊雄（国立研究開発法人土木研究所水災害リスクマネジメント国際センター長） 「包摂的なアプローチによる水災害レジリエンスの強化と SDGs への貢献」
14:50-15:15	澁澤　栄（東京農工大学卓越リーダ養成機構特任教授） 「災害社会と農業の課題（仮）」
15:15-15:40	溝端　佐登史（京都大学経済研究所教授） 「新型コロナの打撃とコロナ後の世界－経済学的検討（仮）」
15:40-16:20	パネル討論
16:20-16:30	総　　括　渡辺　美代子（国立研究開発法人科学技術研究機構副理事・日本学術会議副会長）
16:30	閉会挨拶　引原　隆士（京都大学教授・図書館機構長）

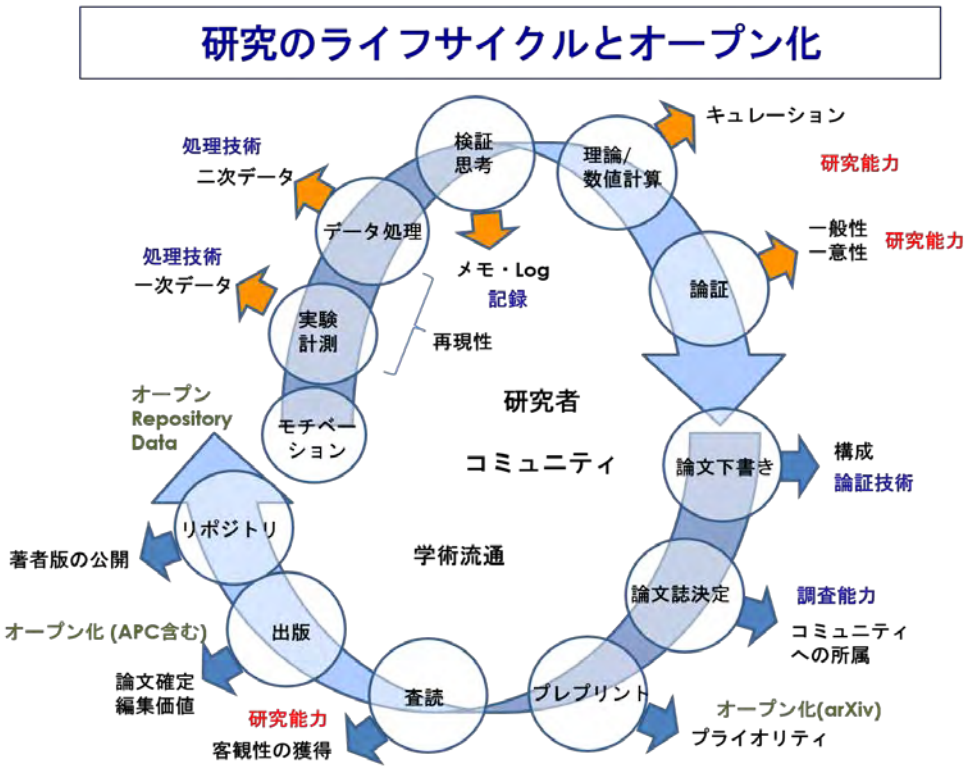
＜付録図表１＞研究データ基盤システムと関連体制



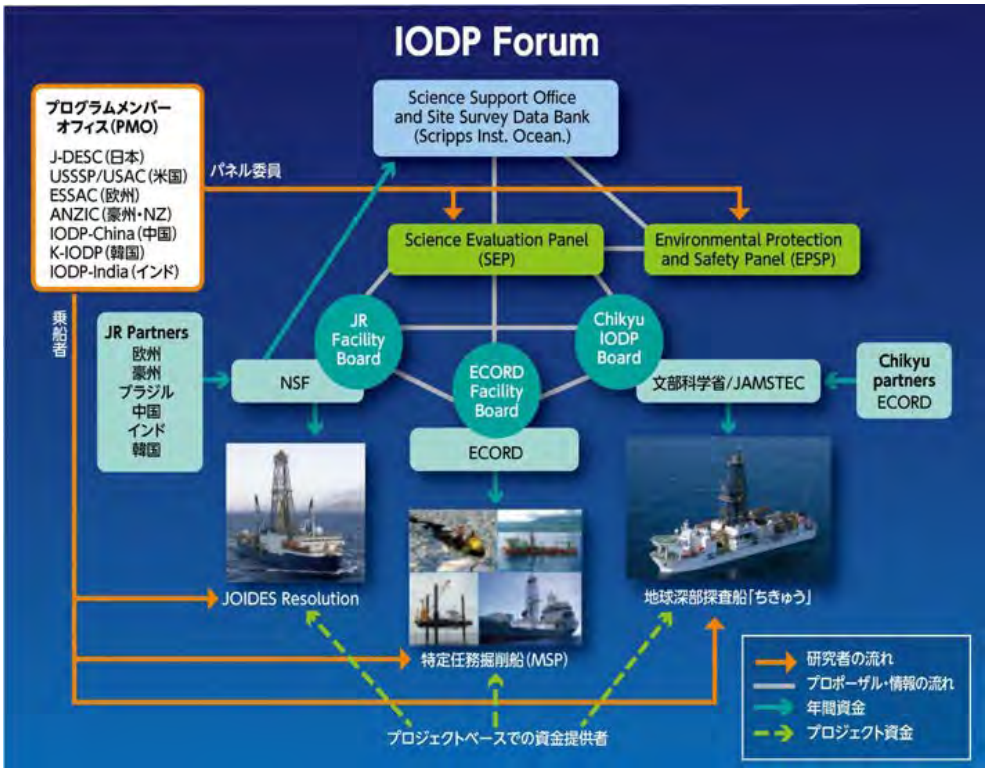
＜付録図表２＞学術情報の位置付け



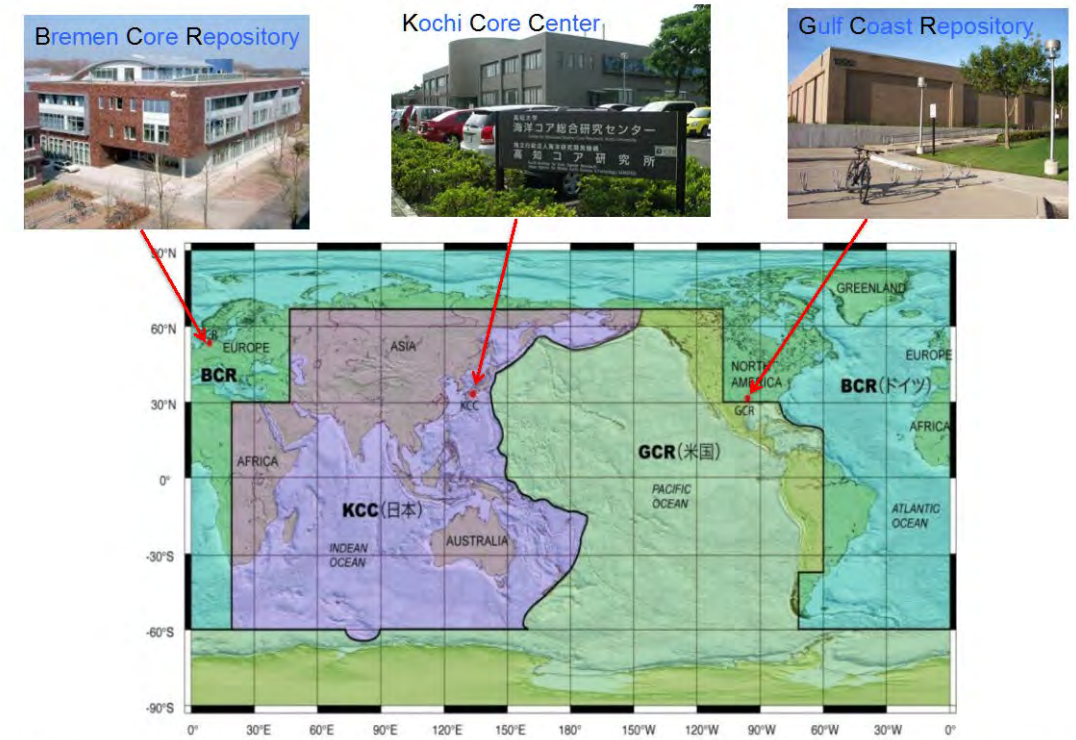
＜付録図表３＞研究のライフサイクル



＜付録図表４＞３つの掘削プラットフォームで運用される国際深海掘削画（IODP）組織図

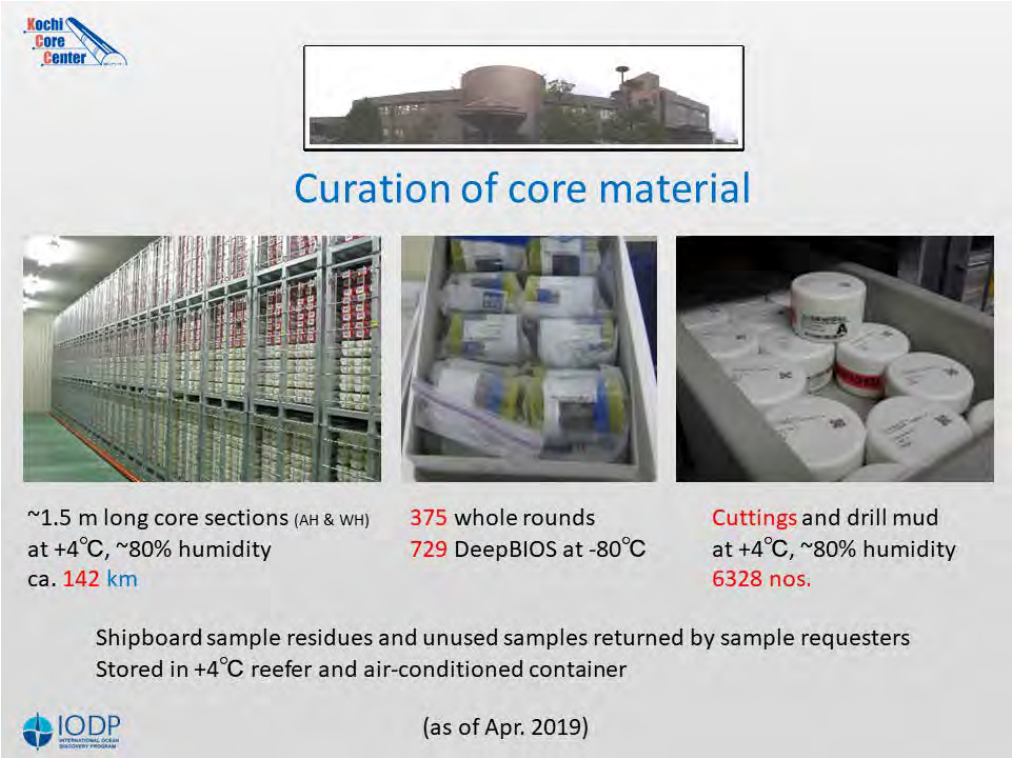


＜付録図表５＞海洋の３区分と３箇所のデータ・試料永久保管施設

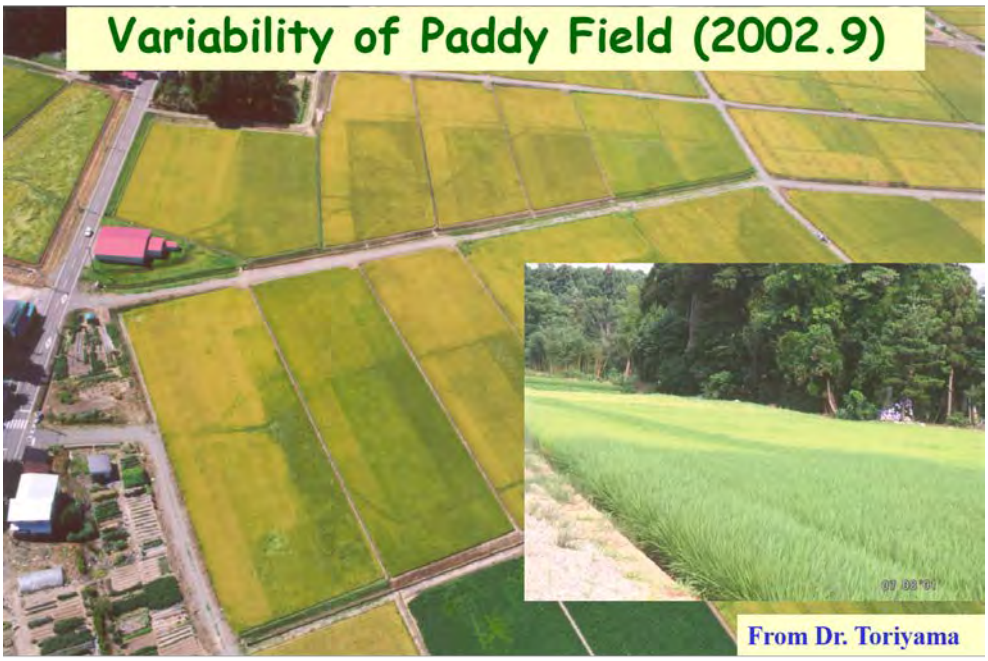


Geographic model: 3 oceanic regions, 3 IODP core repositories

＜付録図表６＞高知コア研究所における試料保存例

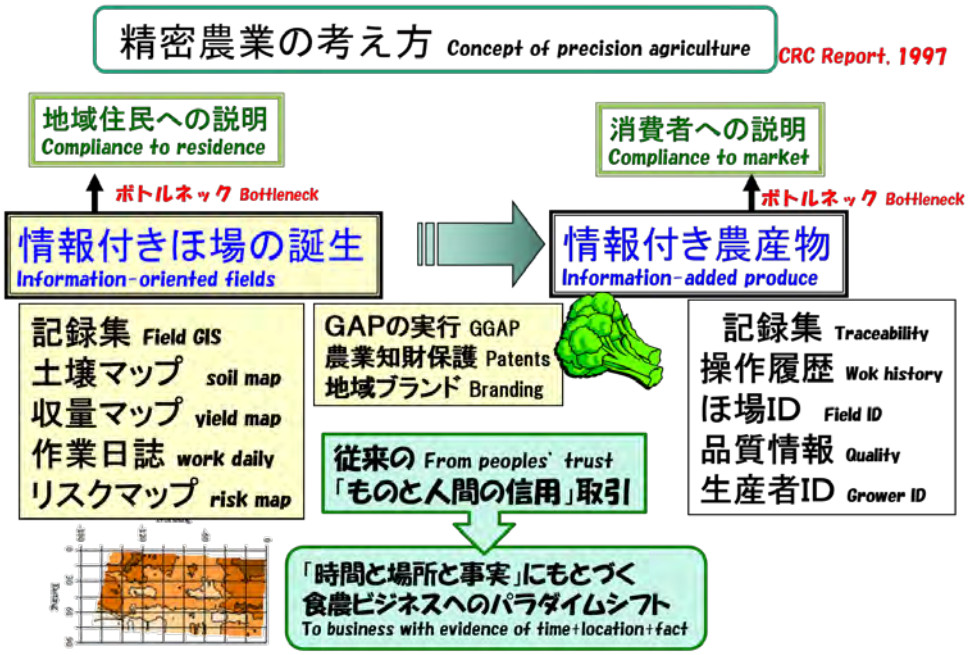


＜付録図表７＞１ｈａ以上の大規模水田のばらつき



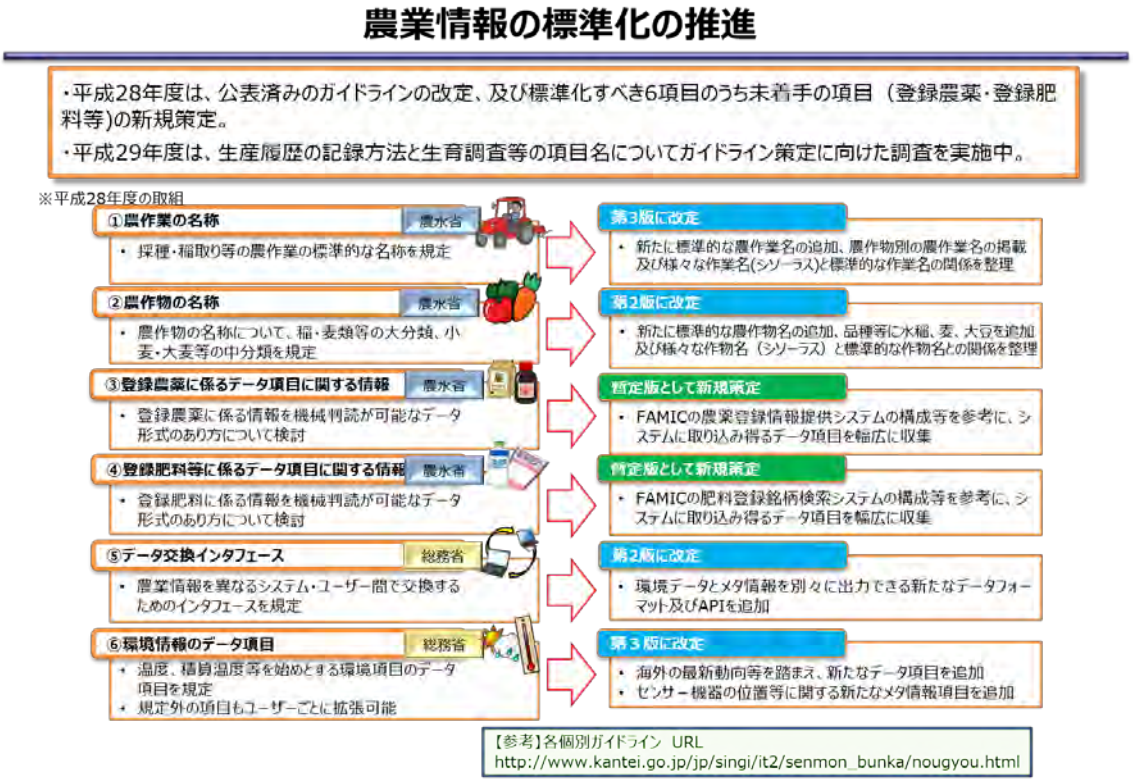
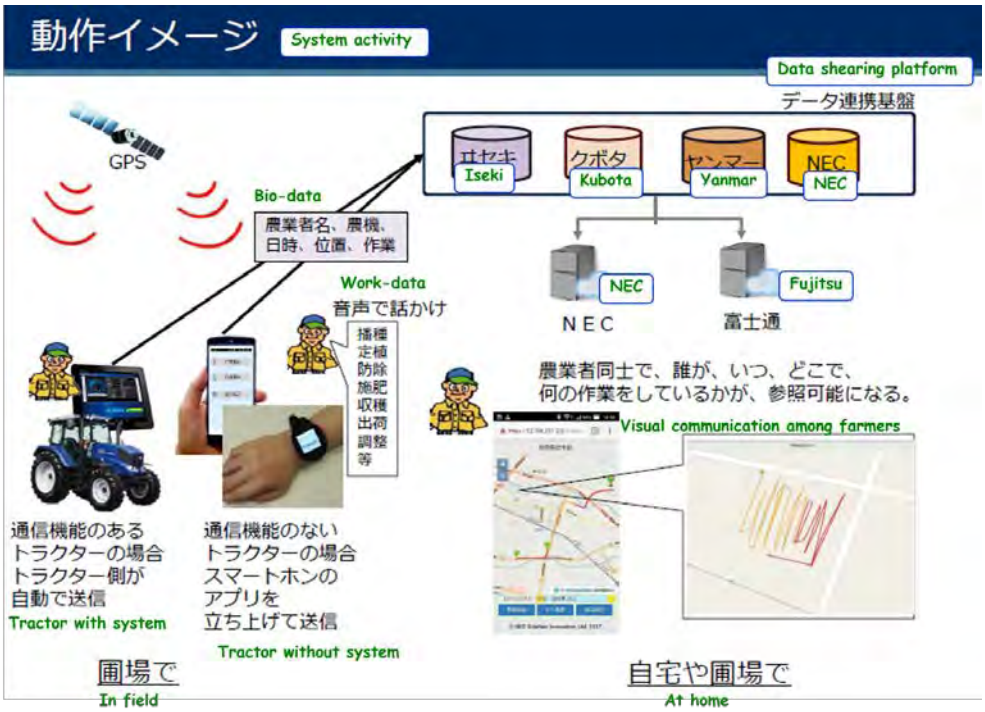
小規模水田を合筆して大きくしても、元の小規模水田の生産特性が反映して、一律の均等栽培は困難である。

＜付録図表８＞精密農業の考え方



ほ場の特性を高解像の時空間データとして記録し、ほ場ばらつきのデータにもとづいた複数の営農判断シナリオを提供する。

＜付録図表９＞農業用語の共通化の取り組み例



＜付録図表 10＞異なるメーカー製品の情報共有例