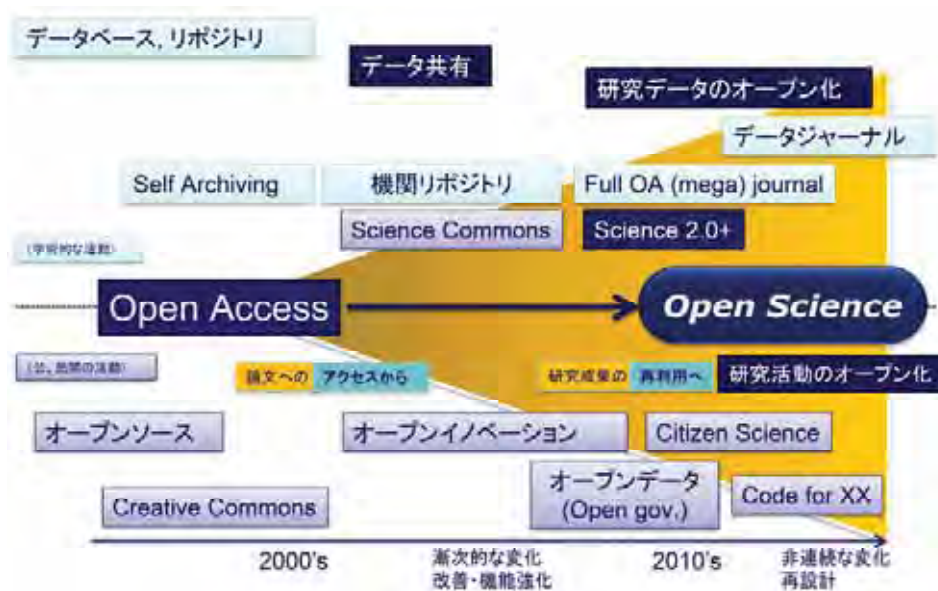


その一方で、現代科学は、ICT の発展・普及に伴い、インターネットが加速度的に発展してきた結果、研究のスタイルにも、その中身にも大きな変化が生じてきている。集積された膨大な量のデータは、科学的実験や研究の基礎となり得るものであり、インターネットを駆使して低コストの「研究データ」の活用・再利用といった専門分野を超えた知識生産のアプローチが登場した。これまでの雑誌論文による「成果の公開と流通」といった昔ながらの古典的なオープンネスの概念から、「研究データを中心とする成果のオープン化」といった研究のプロセス自体をオープンにするといった概念へと変化し、研究者間、あるいは専門分野を超えた知の創造が加速し、新しい共同研究スタイルや新しい研究方法が誕生している。これにより、新たな知の発見やこれまでの研究スタイルでは想像もできないような研究概念を生み出すことが期待され、社会に対する研究プロセスの透明化による理解の促進と、社会的認知の広がりによる新たな研究資金の獲得などが期待されている。

図 オープンアクセスからオープンサイエンスへ



#### 4. オープンサイエンスの主たる構成要素

オープンサイエンスを推進するための中核的役割を果たすオープンアクセス及びオープンデータについて、その基本的な考え方や国際的な検討状況も含めて、簡単に説明しておく。

##### (1) オープンアクセスについて

###### ① 基本的考え方

「オープンアクセス」とは、査読済み論文が『インターネット上で自由に入手でき、その際、いかなる利用者に対しても、論文の閲覧、ダウンロード、コピー、配信、印刷、検索、全文へのリンク付け、検索ロボットによる索引付け、データとしてソフトウェアに転送すること、その他、合法的な用途で利用することを財政的、法的、技術的な障壁なしで許可する』ものとされている。

出典：2002年4月のブダペスト・オープンアクセス・イニシアチブ(BOAI: Budapest Open Access Initiative)

## (オープンアクセスの誕生の背景)

### 1) 雑誌の危機（シリアルズ・クライシス）

世界的な論文数の増加、ニーズに対し代替品が存在せず競争が成立しない市場の特殊性、研究者の研究発表の場である学術雑誌（ジャーナル）の刊行が商業出版に依存している状況、利用者（研究者）と購入者（図書館等）が異なることから、費用を直接負担しない利用者が需要を超える利用環境を求めてしまうモラルハザードなど、様々な要因が複雑に影響し、結果として、ジャーナルの価格が上昇し続け、大学図書館や研究機関等の予算で購読を維持することが困難になり、研究者が研究成果を自由に主体的に発表・収集できなくなった。この危機的状況を特に顕著に表した例として、ハーバード大学図書館の諮問委員会が 2012 年に発表した文書がある。ハーバード大学は、ある出版社 2 社のオンラインジャーナルの購読料が 2006 年から 2012 年の間に 145% 上昇したことに對して懸念を示しており、このようなシリアルズ・クライシスを打開するためには、オンラインのオープンアクセスジャーナルに切り替えるのが最善策であると提案している。

日本においても、シリアルズ・クライシスに対応し、電子ジャーナルをはじめとした学術情報を国内の大学向けに、安定的・継続的に確保、提供を目的とする大学図書館コンソーシアム連合（JUSTICE）が設立されている。このほか物性物理分野の団体である物性グループが新たな情報アクセス体制の確立など 3 つの提言からなる「電子ジャーナルへのアクセス環境の整備に関する緊急アピール」を発表するなどの動きがある。

### 2) 電子環境の発展と情報伝達の拡充

インターネットの普及とそれに伴う学術雑誌の電子ジャーナル化によって、科学技術情報流通の伝達限界費用が限りなくゼロに近くなり、世界中に研究成果を知らしめる環境が整った。研究者自身が論文を web サイトで発表することも増え始めている。

## ② オープンアクセス概念の広がり（ジャーナルの論文からデータを含む研究成果へ）

オープンアクセスは、議論の背景となった課題を超える概念として展開しており、幅広い領域にまたがる研究概念へのアクセスの機会を格段に増大させ、多分野間の協調を促すことにより科学研究の多様化をもたらすとともに、公的資金による基礎研究の成果をイノベーションや新しい産業の創出に繋げると期待されている。つまり、オープンアクセスが拡大することは、ICT の発展と相まって、電子ジャーナルだけではなく、データを含む研究成果へのアクセスを可能とし、そうした世界規模での知的共有基盤の形成により、これまでにない発想による研究の展開やイノベーション創出が期待されている。

## (2) オープンデータ

### ① 行政データと研究データのオープン化

これまでオープンデータに関する活動は、政府が保有する行政データ等の公開活動と

して捉えられることが多かった。本報告書では、行政データ等のオープン化とは別に、研究データのオープン化について主に議論する。前者はオープン・ガバメントとも呼ばれるのに対して、研究データのオープン化においては、科学研究活動上の便益・損失や研究途上の情報の取扱い及び機微の判断など、関与する専門家、研究者又は科学コミュニティが判断すべきものが含まれる。したがって、国としての基本方針の策定に当たっても、行政データのオープン化との違いを理解した上で検討すべきである。

## ② オープン研究データに関する 5 原則

本原則は、G8 の下に置かれている政府高官グループ（GSO）において合意されたものである。

### (a) 容易に探せること（discoverable）

現在広く用いられている検索手法を使って、必要とするデータ（もしくは、データセット）が容易に見つけられることである。

具体的な技術的手段としては、

- 適切で持続的な識別子フレームワークの実現
- 記述メタデータ標準の採用
- 適切なデータフォーマット
- データの分類
- メタデータ検索とエキスパート API（活用インターフェース）

### (b) 容易にアクセスできること（Accessible）

倫理的、法的、商業的利益の保護などの制約は認めつつも、研究データはアクセス可能でなければならない。

アクセス可能となったデータの利用者は、論文などの引用と同じく引用元を明らかにする義務がある。この引用により、研究データの貢献が記録される。データのユーザーは、使用条件又は特定のデータと関連した条件に同意することが必要となる場合がある。例えば、被験者のプライバシーの保護に注意すること、知的所有権の保護又はエンバゴなどがそれに当たる。

併せて、研究データの引用として DOI（Digital Object Identifier）のような世界標準仕様に沿ったデータの同定と所在地（Landing Page）を特定する用意が必要である。

### (c) 容易に理解できること（Understandable）

数値、テキスト、イメージ、動画などは、コンテキスト、セマンティクス、データ分析ツール、アルゴリズムなど一体となって提供されることで、理解可能なものとなる。

観測、実験データは、データが生成された際の条件（時、場所、観測装置、実験手順など）に関するメタ情報によって理解可能となる。データ品質については、データの完全性と由来を確実にするために、生データから公開データに至るまでの加工プロセスについて、当該データが属する研究分野で広く承認された一定の規約が確立されなければ

ならない。

(d) 容易に管理できること (Manageable)

研究データが効率的・永続的かつ効果的方法で管理されるためには、データ管理方針と計画が、研究プロジェクト及び組織のレベルで全てのデータについて定められていなければならない。研究データを第三者によって効果的に利用できるようにすることは、その研究データの元来の目的に使えることを保証するよりも、かなりの継続的な努力を必要とする。

データ管理方針と計画は、誰がデータの有効性を維持する役割を果たすか、そして、キュレーション、保管と利用サービスの維持等に必要な経費がどのように支払われることになっているかという運営と、利用者の負担を明らかにしなければならない。

計画とプロセスは、学際研究を含むあらゆる潜在的用途を考慮に入れ、かつすでに必要性があきらかになっている長期保存も含めて、将来の研究のために利用可能な状態を保つようにしなければならない。

(e) 人材 (People) の確保

以上述べたようなグローバルな研究データインフラは、高いスキルとともに新しい状況に適応力がある人材と、入手可能なデータを収集し、それを有効利用できるユーザーに利用できるようにすることに積極的な組織文化を必要としている。多様でかつ相互に複雑な関係にある数多くのデータセットやプロトコルに精通した専門人材が必要となる。また、研究データ管理に関して、研究コミュニティの文化を変えなければならない。

出典：White Paper: 5 Principles for an Open Data Infrastructure Draft v2.0, 21 May 2013: G8 Global Data Infrastructure WG

③ 電子行政オープンデータの意義・目的

我が国の電子行政オープンデータの意義・目的は以下のとおりとなっている。

● 透明性・信頼性の向上：

公共データが二次利用可能な形で提供されることにより、国民が自ら又は民間のサービスを通じて、政府の政策等に関して十分な分析、判断を行うことが可能となる。それにより、行政の透明性が高まり、行政への国民からの信頼を高めることができる。

● 国民参加・官民協働の推進：

広範な主体による公共データの活用が進展し、官民の情報共有が図られることにより、官民の協働による公共サービスの提供、さらには行政が提供した情報による民間サービスの創出が促進される。これにより、創意工夫を活かした多様な公共サービスが迅速かつ効率的に提供され、厳しい財政状況、諸活動におけるニーズや価値観の多様化、情報通信技術の高度化等我が国を取り巻く諸状況にも適切に対応することができる。

● 経済の活性化・行政の効率化：

公共データを二次利用可能な形で提供することにより、市場における編集、加工、分析等の各段階を通じて、様々な新ビジネスの創出や企業活動の効率化等が促され、我が

国全体の経済活性化が図られる。また、国や地方自治体においても、政策決定等において公共データを用いて分析等を行うことで、業務の効率化、高度化が図られる。

出典：電子行政オープンデータ戦略(2012年7月4日高度情報通信ネットワーク推進戦略本部決定)

(「オープンデータ」と言えるための条件)

- 1) 機械判読に適したデータ形式で、
- 2) 二次利用が可能な利用ルールで公開されたデータ

**「オープンデータの5つの段階(出典:★)」と、データ形式**

段階	公開の状態	データ形式 例	参考) Linked Open Data 5star
1段階	オープンライセンスの元、データを公開	PDF、JPG	OL - Open License (計算機により参照できる(可読))
2段階	1段階に加え、コンピュータで処理可能なデータで公開	xls、doc	RE - Readable (Human & Machine) (コンピュータでデータが編集可能)
3段階	2段階に加え、オープンに利用できるフォーマットでデータ公開	XML、CSV	OF - Open Format (アプリケーションに依存しない形式)
4段階	Web標準(RDF等)のフォーマットでデータ公開	RDF、XML	URI - Universal Resource Identifier (リソースのユニーク化、Webリンク)
5段階	4段階が外部連携可能な状態でデータを公開	LoD、RDF スキーマ	LD - Linked Data (データ間の融合情報が規定。検索可能)

人が理解するための公開文書(編集不可)

公開文書(編集可)

機械判読可能な公開データ

**オープンデータの5つの段階**

出典：★ Open Dataのサイト (<http://5stardata.info/>) およびTim Berners-Lee氏のLinked Dataに関する提言ページ (<http://www.w3.org/DesignIssues/LinkedData.html>) を参考に作成。

## II. 国際動向を踏まえたオープンサイエンス推進の必要性

オープンサイエンスは、従来の科学研究活動の仕組みを大きく変える可能性を持つ概念である。しかし、それは、これまでの研究方法を代替するものではなく、従来の研究方法に対して新しい研究方法を提示し、サイエンスの新たな進展を可能にするものである。

科学コミュニティ、産業界、一般国民などあらゆるユーザーが研究成果をデジタル形式で広く利用できることにより、科学技術情報の更なる活用が促される。しかも、その波及効果は知の創造プロセスに留まらず、社会全体のイノベーションシステムの変革にも及ぶものと考えられる。

科学コミュニティにおいては、研究者間あるいは研究分野を越えたデータ駆動型の取組が加速することにより、新たなコラボレーション、新たな研究方法が広まることが期待でき、企業や個人においては、科学的成果を活用・再利用して新しい製品や新しいサービス（市場）を生み出すことが期待される。

特に、天然資源の乏しい我が国において、持続的な発展を続けていくためにも、科学技術イノベーションにより常に新たな価値を創出する環境を構築することが不可欠であり、オープンサイエンスの推進は、そのための環境整備にほかならないということを、ステークホルダーで共通認識を形成した上で、推進体制を構築する必要がある。

### 1. 研究成果の活用・再利用によるイノベーションの創出基盤づくり

- (1) 従来の科学研究活動の枠組みが変わることにより、科学的なデータへのアクセスが増加すれば、科学研究活動の効率化と生産性の向上をもたらし、国内外からの研究過程への参加の機会が増加することでデータの共有（統合）が進み、これまで取り組むことができなかったより複雑多岐にわたる研究を可能とすることが期待される。これらを確実に機能させるための仕組みを構築し、有効化する必要がある。また、次世代の研究者が同じ研究を繰り返すことを避け、成果（論文、研究データ等）の活用・再利用ができる基盤を構築する必要がある。
- (2) データ生成者との直接的なつながりがなくとも、データの存在を公開することで異分野での利活用を進展させる（新規分野開拓）ことにより、新たな知見やイノベーションを創出する仕組みとする必要がある。
- (3) 研究成果としてのデータが、どこにあるかを把握しやすくし、あるいは（プロジェクト終了後）アクセスできない現状を改善することにより、あらゆるユーザーが活用・再利用できるようなプラットフォームづくりが必要である。
- (4) 担当研究者の退職やプロジェクトからの離脱後に、価値のあるデータを管理・理解できる後継者がおらず放置・削除される現状の改善につながるような、データの集約的保存基盤等の構築が必要である。
- (5) オンラインジャーナル購読料の高騰に伴い、大学等における研究成果（論文）に対するアクセスが困難になりつつある現状の改善（シリアルズ・クライシスの打破）に繋がる仕

組みを構築する必要がある。

- (6) イノベーションを誘発するためには、最新の研究成果や他領域での新たな知見、データなどを総合的に扱いながら課題を解決する能力を持つ人材の育成が不可欠であり、そのための基盤としての成果・データ共有プラットフォームの構築は、これからの我が国における人材育成という観点からも必要である。

## 2. 研究成果の質と透明性の確保

- (1) 科学技術の複雑化・専門化に伴い、社会からの科学技術に対する理解と支持を得ることが難しくなりつつある状況を認識する必要がある、研究成果の公開により社会からの信頼を確保できる体制を構築する必要がある。
- (2) 将来、科学技術が更に進歩した時点で過去の真偽確認ができない等の問題を回避することが可能となることを認識し、論文、研究データの長期保存ができる基盤を構築する必要がある。
- (3) 論文、研究データの公開は、研究不正を回避する意味でも重要であり、科学技術の進展と研究活動の透明性、公正性を確保することが必要である。

## 3. オープンサイエンスの推進と人文・社会科学の役割の増大

第4期科学技術基本計画においても、科学技術政策と科学技術に関連するイノベーションのための政策を総合的に推進する「科学技術イノベーション政策の一体的展開」を掲げている。また、科学技術と社会との関係が深化する中、科学技術イノベーション政策を「社会及び公共のための政策」の一環として、国民の幅広い参画を得つつ、理解と信頼を得ながら進めていくことが必要となる。このためには、科学技術イノベーションが社会にもたらす効果や影響を可視化するなど、客観的な根拠に基づき政策の企画立案及び推進等を行い、政策形成プロセスをより合理的なものにするとともに、国民に対してより一層の説明責任を果たしていくことが必要となる。

こうした観点からも、オープンサイエンスを推進することは、人文・社会科学の専門家が科学技術と社会の関係を探る課題に対し、社会のための科学技術、社会の中の科学技術という観点に立った人文・社会科学的研究を推進し、その成果を踏まえ媒介的活動が活発に行われることが期待される。

## 4. 国際的潮流に乗り遅れることのデメリット

我が国においては、論文に係るオープンアクセスについて一部取組が進んでいるが、研究データに関しては、これまで組織的に十分な議論がなされておらず、以下のようなデメリットが指摘されていることも踏まえ、我が国としての特徴を生かしつつ、欧米諸国、アジア諸国など諸外国と積極的に連携し、国際的な発信力を確保する必要がある。また、日本の科学コミュニティがオープンサイエンスの流れに乗り遅れることで、日本の研究者が世界から取り残され、さらには日本全体の科学の質の低下を招く可能性があることを認識しておく必要

がある。

## (1) オープンサイエンスが進まないことのデメリット

### ① 地球規模研究におけるデメリット

研究データのオープン化への対応が十分ではない場合、共有すべき研究データの共有が進まないことで、地球規模で推進される研究に参加できず、国際競争力の低下を招くなど、日本の研究者からの成果が反映されないことで日本が世界から取り残され、日本の科学技術が国際的に存在感を低下させる懸念がある。

### ② 研究活動における効率化、国際発信におけるデメリット

我が国としての明確な意思表示（基本方針等）が示すことができなければ、オープン化のデファクト・スタンダードが海外主導によって形成され、日本における研究活動及び国際発信に不利益が生じる可能性がある。

### ③ 日本の「見えない化」

世界でオープン化が進み、研究論文、データの相互交流が活発になると、国内に閉じた研究が相対的に存在感を低下させていく可能性がある。その結果、日本の研究者の業績が正当に評価されず、海外での研究や国際共同研究の機会を失う可能性がある。

## (2) RDA(Research Data Alliance)の議論の特徴に見る日本にとっての示唆

研究データのオープン化に係る議論は、RDA が国際的にリードしており、我が国として議論の動向を把握するとともに、我が国としての基本姿勢を明らかにしつつ、議論の輪に加わり、世界各国と対等な関係を構築していく必要がある。

### ① RDA における議論の特徴

#### (a) スピードが早い

RDA に設置される WG は、12 ヶ月～18 ヶ月を目途に、推奨テクニカルペーパーを出すなどの成果が求められている。

#### (b) 波及効果が大きい

米国、欧州委員会、豪州、英国、カナダ、ICSU（国際科学会議）等の政府系機関、国際的組織や大学等が参加している。（G8 科学大臣会合でも報告）

#### (c) 参画しないことの日本のデメリット、リスクの可能性

グローバルなデファクト・スタンダード形成に向けた議論が加速的に行われているが、現状では日本からの意見を十分に盛り込めていない。

### ② 日本にとっての示唆

RDA 以外に検討のない分野、技術については、RDA での検討結果がルール決定の際の有力なたたき台、あるいは、事実上のルールとなる可能性が高い。RDA 以外に検討する舞台がなければ、RDA が当該分野の議論を支配することになり、事実上のルールメーカーとなっていく可能性もある。



### Ⅲ. オープンサイエンスに関する国際動向への対応

#### 1. 国としての基本姿勢・基本方針の明確化

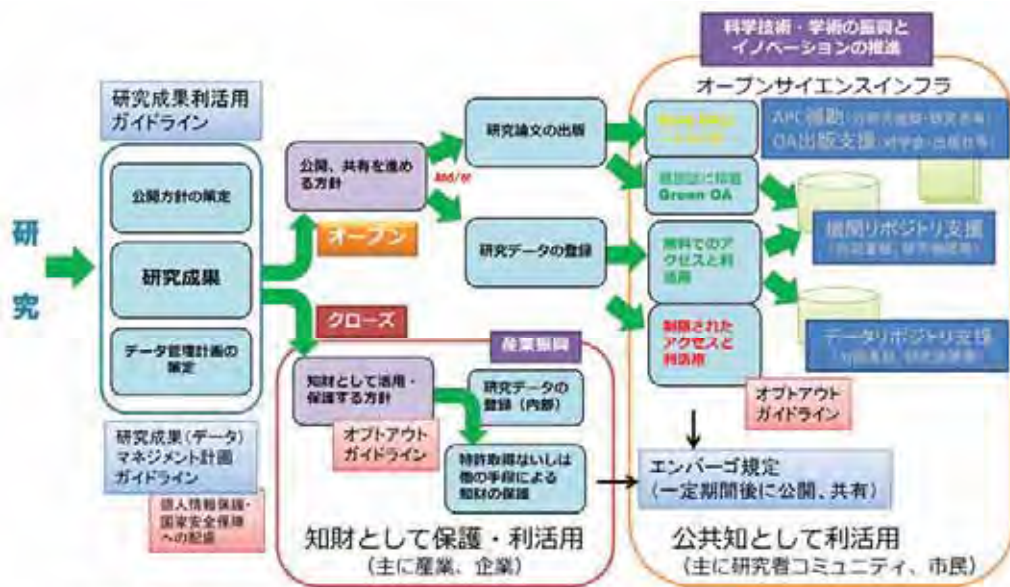
公的研究資金による研究成果（論文、研究データ等）の利活用促進を拡大することを我が国のオープンサイエンス推進の基本姿勢とする。

その上で、各省庁、資金配分機関、大学・研究機関等のステークホルダーがオープンサイエンスの実施の責任を果たし、オープンサイエンスの実施方針及び実施計画を策定するものとし、それに当たって、参照すべき共通事項や留意点等をオープンサイエンスに関する基本方針として明示する。

各省庁等のステークホルダーは、オープンサイエンスを推進すべき領域、プロジェクトを選定し、科学研究活動上の便益・損失や研究途上の取扱い及び機微の判断など各分野の専門家・研究者、技術者の意見を十分に取り入れ、その分野の活動・研究成果が最大化されることを旨として、オープンサイエンス実施方針等を定める。

オープンサイエンスの推進に当たっては、内閣府及び総合科学技術・イノベーション会議が政府全体を通じた旗振り役として中核的な役割を担い、各ステークホルダーにおける進捗状況をフォローする。

図 研究成果の利活用、オープンサイエンスの推進に係る概念図



下記図表を参考に和訳、改定  
 Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 Version 1.0 11 December 2013 p.4  
[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)

## 2. オープンサイエンス推進の基本的考え方

### (1) オープンサイエンス推進の目的・意義

公的研究資金による研究は、新たな発見や洞察を提示し、それにより、健康、エネルギー、環境、農業といったあらゆる分野の進歩を強力に牽引してきた。このため、そうした公的研究資金による研究成果として得られた論文や研究データへのアクセスを可能とすることで、研究成果の理解促進と同時に、成果の再利用による新たな発見や、新たな研究概念の創出とイノベーションを加速し、新たな産業の創出、競争力の強化、地球規模での研究の促進、経済成長等に貢献するものである。

### (2) オープンサイエンス推進に係る公開の範囲

公的研究資金による研究成果のうち、論文及び論文のエビデンスとしての研究データについては、原則公開とし、その他研究開発成果としての研究データについても可能な範囲で公開することが望ましい。これにより、国内外の研究者、大学・研究機関等の中でデータアクセスとデータ共有を促進し、あらゆるユーザーが研究成果を自由に活用・再利用することを可能とし、新たな知見や価値を生み出すことを可能とするものである。

なお、ここでいう研究成果のうち、機密保持、企業秘密、国益及び国家安全保障に関わるもの、研究成果の商用化・産業化を目的として収集されたデータ、又は民間企業が保有するデータ並びに共同研究契約などで研究成果の公開に制限がある場合などは、公開適用対象外として考えるべきである。

そのほか、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必要がある。

### (3) 「公的研究資金を用いた研究」の定義及び研究データの範囲

#### ① 「公的研究資金を用いた研究」の定義

競争的研究資金及び公募型の研究資金に該当するものとする。また、国費が投入されている独立行政法人及び国立大学法人等の運営費交付金等を100%活用した研究活動等も対象とすべきである。

#### ② オープン化の対象となる研究データの範囲

研究データは原著論文と同等又はそれ以上の研究開発成果として位置づけられ、公開適用の対象となる研究データは、メタデータ、数値データ、テキストレコード、イメージ、ビジュアルデータなど多様なデータが含まれる。これには原則として以下のものは含まれない。研究ノート、予備的分析、論文の草稿、将来の研究計画、同僚との私信、物理的対象物（研究室の試料、バクテリアの菌株、実験動物など）等及び所属機関との契約により公開に条件が付されているもの。

### (4) 公的研究資金を用いた研究を実施する機関の責務

公的研究資金を用いた研究を行うための設備、ルール、人材等を具備している機関にお

いては、論文、研究データ等の研究成果の管理に係る規則を定め、特に、研究成果の散逸、消滅、損壊を防止するための具体的施策を講ずる必要がある。

具体的には、保存すべき研究成果及び研究資源の全てに永続性のあるデジタル識別子（Persistent Object Identifier）を付与し、管理する仕組みを確立する必要がある。

（参考） データ共有の利点と問題点

データ共有の利点	データ共有の問題点
<ul style="list-style-type: none"> <li>✓ 誤りと不正の減少</li> <li>✓ 研究投資への費用対効果を増大</li> <li>✓ 助成機関及びジャーナル規定の遵守</li> <li>✓ 重複とバイアスの減少</li> <li>✓ 研究の再現／検証を促進</li> <li>✓ さらなる仮説の検証</li> <li>✓ 教育目的での利用</li> <li>✓ 他のデータセットとの統合</li> <li>✓ 引用の増加</li> </ul>	<ul style="list-style-type: none"> <li>✓ 不適切な再利用に関する懸念</li> <li>✓ 限られた時間と人的資源</li> <li>✓ データの賞味期限やデータ共有に伴う費用</li> <li>✓ 医療関連データなど、プライバシーに関する懸念</li> <li>✓ データの所有権／データ公開権限の所在が不明確</li> <li>✓ 学問上のインセンティブ／認識の不足</li> <li>✓ リポジトリ不足またはリポジトリの認知不足</li> <li>✓ 企業の秘密情報の保護</li> </ul>

### 3. 関係機関におけるオープンサイエンス実施方針等のあり方

#### (1) 総論

関係省庁、研究資金配分機関及び大学・研究機関は、あらゆるユーザーからのアクセス拡大を促進するため、必要に応じ、機関同士が協同して計画を策定するなど実効性を高めることが望ましい。

関係機関が定めるオープンサイエンスに関する実施方針及び実施計画には、次の要素が含まれていることが必要であると考えられる。

- イノベーションや競争力を強化する取組
- ステークホルダーにとって透明なプロセス
- 公的研究資金による研究成果に係るデジタル研究データの所在を特定しアクセスできる方法
- 研究成果の長期的管理を保証した上で、アクセスのし易さや相互運用を奨励するような検索、保管及び普及機能を最適化するためのアプローチ方法
- 公的研究資金を受けている研究者に対して義務を周知するためのガイドラインの策定、関連する規則の改正等
- 機関の計画等の遵守に関する評価とともに、必要に応じて、計画等の実行徹底を図るための戦略
- 計画等を実行するための運用経費等の確保
- 計画等を実行するための優先度を考慮したロードマップ策定
- データインフラ（リポジトリ等）の整備計画

## (2) 論文へのアクセスのためのポリシー

論文のオープンアクセス化については、2002年4月のブダペスト・オープンアクセス・イニシアチブに基づいて、オープンアクセスを推進する。

したがって、公的研究資金から発生した論文（出版物等）については、あらゆるユーザーがアクセス、検索、読み出し、分析できるよう、長期間にわたって保存しなければならない。

- 各関係機関の計画等は、以下の点を明確にしたものでなければならない。
  - 実行可能な範囲において、学術出版物が不正に大量配布されることを防ぐための手順を示すこと。
  - デジタル形式での公開、検索、ダウンロード、分析、アクセス性の向上の促進。
  - 論文を利用可能とするためのガイドラインとして、即時のオープンアクセスが難しい場合は、適切な期間のエンバゴを設けることが望ましい。その際、各機関においては、各分野及び固有の研究課題や国益等に配慮するため、独自の取組とすることができる。また、特定分野においては、計画と整合がとれない場合においては、利害関係者がエンバゴの変更を求めることができる仕組みを用意するものとする。
  - エンバゴ終了後は、論文全体を提供すること。論旨を示すデータとともに、関連する補足資料へのアクセスも提供することが望ましい。
  - 著者、ジャーナル及び出版元に対する権利の帰属が維持されるようにすること。
- 出版物及びメタデータが以下の観点から確実に保存される必要がある。
  - コンテンツの長期保存と無料アクセスを提供すること。
  - テキスト及び付随コンテンツ（画像、映像、補足データ等）について、広く普及している標準的な保存フォーマットを使用すること。
- オープンアクセスを可能とするためのリポジトリによるグリーンオープンアクセス（グリーン OA）、ゴールドオープンアクセス（ゴールド OA）への対応を示しておく必要がある。
- 特に研究論文のグリーン OA 化を前提とするオープンアクセス義務化に関しては、グローバルな関係者間のコンセンサスが得られつつあることを念頭に置いた方針の策定が必要である。

## (3) デジタル化された研究データへのアクセス

研究分野によって研究データの保存と共有の作法に違いがあることを認識し、特性に応じた計画等を策定する必要がある。

その上で、公的研究資金による研究の結果、生成されたデジタル形式の研究データについては、公開するデータとなった場合、あらゆるユーザーがアクセス、検索、読み出し、分析等再利用できるよう保存しなければならない。

- デジタル形式での研究データへのあらゆるユーザーからの無料アクセスを最大化する

必要がある。ただし、

- 機密性及び個人のプライバシーを保護すること。
- 財産的価値のある成果物を保護すること。
- 長期間保存が可能となるようデータインフラのロードマップを作成すること。
- 提供するデータを自由に利用できることを示す利用ルールを付す必要がある。
  - 研究データは著作物ではないため、CC0 を採用することが望ましい。また、データの集積・整理や品質管理などの点で専門的研究者の労力や高度なノウハウが入ったデータベースなど、知的生産物として著作権が発生する場合には CC-BY を採用することが望ましい。
- 公的研究資金による支援を受ける研究者が、必要に応じ、データ管理計画を作成するようにすること。同計画は、データの長期保存及びアクセスが困難な場合において理由を説明することにも活用される。
- 研究者から提出されたデータ管理計画が適切に評価されるようにすること。
- 研究者自ら提出したデータ管理計画に従っていることを確実にするための仕組みを盛り込むこと。
- 適当な場合には、ユーザーがアクセス可能なデータベースにデータを預けることを奨励すること。
- 計画の下で利用可能となる研究データセットに対し、適切な属性を特定し、提供するための基盤を開発すること。
- 研究データ管理、分析、保管、保存及び共有技術等に関する訓練、教育及び人材養成を支援すること。

#### (4) オープンサイエンス計画の実施とフォローアップ

- 各機関は計画等が確定した場合は、ホームページ等に掲載するなど広く周知できるよう努力しなければならない。
- 内閣府及び総合科学技術・イノベーション会議は、各機関における計画等の実施状況を定期的に確認する。
  - データマネジメントの構築、進捗状況の確認等を行う。
  - データの利活用の促進に向けて、プラットフォームの構築、保存・共有技術、人材育成等の進捗状況等についてフォローアップを行う。

(参考) 政策立案及び実施における相関図 (イメージ)



#### 4. オープンサイエンスを推進する際の留意点

オープンサイエンスを推進していくためには、指摘されている課題についても十分に留意しておく必要がある。このため、各関係機関においては、以下の留意点を踏まえた実施方針及び実施計画が策定されることを期待する。

##### (1) 「無料」と「自由」の定義問題

オープンアクセスに関して、「無料」と「自由」の定義について、関係者間で混乱があることが指摘されており、定義の明確化による共通認識の醸成に努める必要がある。

##### (2) 著作権ポリシー

研究成果の利活用促進の上で避けては通れない著作権の課題について取り組む必要がある。

例えば、学協会著作権ポリシーデータベース (SCPJ: Society Copyright Policies in Japan) によると、登録学協会 2,609 件のうち、著作権ポリシーを有していない (非公開、無回答を含む) 学協会が 1,355 件となっている (2015 年 3 月 20 現在)。

##### (3) 大規模データセットの所有権の定義

コンピュータ又はソフトウェアプロバイダーが収集する可能性がある大規模データセットの所有権の定義、とりわけ医療記録や生体記録など特定クラスのデータに関連するプライバシー、機密性又はセキュリティの問題に対応しておく必要がある。

(4) 学術誌出版界との共存関係の構築

オープンアクセスのための新たなビジネスモデルの開発や新しいデータキュレーションと保存サービスの提供に取り組んでいる学術誌出版界への対応（出版社や第三者機関に預けるのか、自前（日本）で管理するのかという公的な知的財産の保有ルール）についても整理しておく必要がある。

(5) 研究者及び科学コミュニティに対するインセンティブ

研究者自身の研究成果への世界からのアクセスは、個々の成果の価値の敷衍、深化、発展を促進し、科学自体の進歩にととまらず、社会的意義も高めることが期待される。一方、特定の研究者が全ての情報を優先的に利用するような方針とならないよう計画を策定する必要がある。研究界の秘密主義や一方的な公表に対する競争環境が研究者の協力を阻み、研究者が研究データやアイデアの共有やアクセス拡大を行う際の妨げになる可能性がある。これまでデータを作成し、他の研究者に提供・利用できるようにする活動は、論文投稿に比して、十分に評価されてこなかった分野も多い。このため、政策誘導として、研究者及び科学コミュニティに対するインセンティブを高め、オープン化に対する努力を評価することが重要である。

(6) 研究分野の特性に対する配慮

研究分野（物理学、化学、材料科学、地球科学、バイオサイエンス、人文・社会科学等）によって研究データの保存と共有の作法に違いがあることを認識し、特性に応じたルールづくりが必要である。研究分野でのデータ共有やオープン化を理解する関係者の多寡、データ生成や整理に対する研究当事者の労力の大小や特殊なスキル・研究能力の必要性の関係、個々のデータセットの取り扱いに対する技術的容易さ、困難さなど、様々な固有の事情が各分野にありうることを認識しておく必要がある。また、高度な研究データであっても、人工衛星観測、地震・気象分野など、大学・研究機関等が新規性のある実験や装置で取得する研究開発目的のデータと、行政データとして整理・公開が行われる可能性のあるデータが混在する分野もあることも認識が必要である。

(7) オープンサイエンスのためのスキル

研究者が論文又はデータをオンラインでオープンに共有するために必要なスキルが偏在している。一部の研究分野においては、大規模な研究データを蓄積する慣例を有する場合もあるが、大部分の研究者にはオープンサイエンスを実現するために必要なスキルを開発するための訓練、教育が必要である。

(8) 技術的インフラ・人材育成

オープンサイエンスには、研究データを記述するメタデータ、データを処理するアルゴリズム、オンラインインフラ等の整備が不可欠である。異なるインフラ間の相互運用性も

データ共有の障壁になり得る。さらにはデータを有効に再利用できるようにするために、データの品質管理とともにメタデータとも関連づける必要がある。大学・研究機関等においては、技術職員、URA (University Research Administrator) 、大学図書館職員等を中心としたデータ管理体制を整備できるように、データサイエンティストやデータキュレーターなどを研究支援人材として位置づけられるよう、包括的な育成システムを検討し、推進することが必要である。

#### (9) 適切かつ持続可能な資金提供モデル

オープンアクセスを着実に実施するためには、大学・研究機関等の研究者がオープン化の義務を果たし、マネジメント及びモニタリングが行えるよう資金を確保する必要があり、公的研究資金の中に含めるのか、オープンアクセス維持費として別途の支援スキームを用意するのか、制度設計が必要である。

### 5. 今後の検討課題等

オープンサイエンスの推進に当たっては、長期的視点から取り組むべき課題が多く存在しており、特に下記事項については、継続的に検討を行う必要がある。

内閣府及び総合科学技術・イノベーション会議においては、関係機関の取組状況等に応じ、戦略的かつ継続的に検討することが求められる。

#### (1) 論文、研究データの公開・共有化に係る検討

- 科学技術の進展を妨げない工夫が不可欠
- 企業のイノベーションを促進する観点も重要
- 共有すべきデータは、知的財産戦略を含め、総合的に検討を行う必要
- 学協会等における著作権ポリシーの公開に向けた取組
- 運営・促進するのに必要な資金の制度化

#### (2) 研究データの保存に係る検討

- 研究終了後も研究データの保存・整備が必要
- データ保存・整備のしくみ（データリポジトリ等）が必要
- 論文や研究データの永続的、長期的保存を担保するためには、国立国会図書館、大学図書館、情報基盤センターをはじめ、科学技術・学術文書・書籍を扱える図書館や情報基盤センター等の協力を得ることが有効
- 公開されたデータや情報の権利関係（ライセンス）を明確にし、また永続的保存や提供サービスに関しての制度づくり・責任体制の議論が必要
- 保存に関わる経費は、コンピュータなどの経費だけではなく、データを管理・維持するための人的コスト等が重要であることを考慮すべき。計算機資源としてはクラウドを



活用するなど、コストの圧縮を考慮しておくことが必要

- 将来に向けて持続的なオープンサイエンス制度として自立できるよう、組織的なデータ利用をライセンス化するなどのビジネスモデルの創出（著作権制度などのイメージ）

(3) 保存すべきデータ及び保存期間等

- すべてのデータを保存することは現実的ではないことに留意
- 世界のルールづくりに参画していく姿勢が重要
- 「データ出版」の考え方についても配慮が必要

(4) 研究データの技術的な品質の評価等

- 研究データの技術的な品質の評価、またそれぞれの研究分野の固有の基準に沿った評価について合意しておく必要がある。論文のピアレビューに相当するような評価の仕組みが必要

(5) 研究者に対するインセンティブ等

- 研究者に対するインセンティブが必要。例えば、高品質なデータを提供した研究者に適切な報酬（助成金や昇進機会）を与えるなど、成果に見合う処遇の仕組みを設けるための検討

(6) データ駆動型の研究をサポートするサービスを企画、開発、運用する人材の確保

- 現状の ICT の進化による爆発的な研究データの増大と保存技術の変化
- データサイエンティスト等の育成と活用
- 科学技術分野ごとに異なるデータの属性、管理手法、利用者（特定のスキルの有無）、利用局面等を理解できる人材の確保（データサイエンティスト、データキュレーター）

## おわりに

本報告書は、世界的な潮流となっているオープンサイエンスにおいて、我が国のステークホルダーがオープンサイエンスのメリットを最大限に享受することを可能にする一方、オープン化に伴う不安や課題を着実に解消するための基本方針を示したものである。

今後、この基本方針に基づき、オープンサイエンスを着実に実行し、我が国としての科学技術イノベーションにおけるパフォーマンスを高め、諸外国と積極的に連携し、国際競争力を維持していくためにも、ステークホルダーとなる関係者が共通認識の下、オープンサイエンス推進のための実施方針及び実施計画を策定するなど、実行可能な体制を形成していくことを期待する。

このため、内閣府及び総合科学技術・イノベーション会議に対しては、オープンサイエンスは一朝一夕で実現し得るものではないことを十分に理解した上で、第5期科学技術基本計画に確実に反映するとともに、各ステークホルダーがオープンサイエンスを推進するに当たり、今後、検討すべき長期的課題等について継続的に議論を行い、各機関における計画の実施状況等の定期的なフォローアップを行うことを求めたい。

最後に、この報告書が契機となり、新たなサイエンスの手法の普及・定着により、分野融合や国際協働が進展し、従来の枠を超えた価値が生み出されるサイエンスの新たな飛躍の時代の幕開けが訪れることを期待するとともに、このオープンサイエンスの推進が我が国に明るい未来をもたらすことを願うものである。