

## 2.5 専門人材の不足、育成

データシェアリングがその意義を発揮するためには、研究データを収集、蓄積するだけでなく、研究データへのアクセシビリティの確保と、容易に再利用可能な環境を構築することが必要不可欠である。しかし、データシェアリングの各プロセスにおいて必要とされるインフォマティクスを主軸とした多様な専門人材が研究の現場に存在せず、わが国の研究情報基盤は脆弱であるといえる。

また、大学等高等教育機関では、そうした人材を育成する教育環境は用意されておらず、キャリアパスが不透明で雇用も不安定であることから必要な専門人材の確保、育成が困難である。

データシェアリングに必要とされる高度な技術を持つ多様な専門人材を持続的に確保し、研究者が適材適所で活躍できる社会システムの構築が求められている。

## 2.6 データ研究、基盤構築に係る研究業績評価

現在、研究プログラムにおいて貴重かつ重要なデータを作成した研究者であっても、その成果が業績評価にしっかりと反映される仕組みがない。また、地道なデータ基盤構築、データ整備を必要とする分野の研究活動については、持続的な研究資金が獲得しにくくなっており、その結果、わが国の研究の現場においてデータ整備が進められていない状況となっている。

そうした中、公的研究資金配分機関や研究コミュニティは、従来の論文発表による業績評価に加え、研究データの生成やデータ共有に触発されるデータ引用についても、研究者の貢献と認知し、研究成果として評価の対象とすることが求められる。

## 2.7 研究者、研究コミュニティのデータ共有意識

研究者の多くは、データの保存、交換の取組みに関し、自身の研究内容の先進性、独自性などに対する排他的利用の観点から、自身のデータの公開に抵抗感を持つケースや、データ共有に係る保存・管理などの作業負担を感じている。また研究者個人にデータ管理を依存している状態は、科学的に貴重なデータの散逸、紛失、滅失、廃棄等による再現不可能性に陥ることが考えられ、社会に対する科学研究の信頼性に多大な影響を及ぼす。

このような研究者に根強く残るカルチャーを脱却し、わが国でデータシェアリングを推進するために、従来の論文や特許による評価に加え、データシェアリングによって得られるアウトカムを把握し、評価する技術・方法を開発し、研究者の意識高揚、データ共有意識の醸成にかかわる様々な議論、施策を行っていくことが必要である。

### 3. データシェアリングのあり方に対する提言

#### 提言1：データシェアリングポリシーを早急に策定すべき

##### 【データシェアリングポリシー策定】

- わが国の研究環境を世界最高水準の環境にしていくためには、これまで研究者が個別に行っていたデータ共有の仕組みを、組織の仕組みとして稼働させ、グローバルかつ分野横断的な観点からできる限り幅広くデータが共有されるためのデータシェアリングポリシーを策定する必要がある
- 日本の総合的・基本的な科学技術・イノベーション政策の企画立案及び総合調整を目的としている内閣府総合科学技術・イノベーション会議(CSTI)が策定するオープンサイエンスにおけるデータシェアリング方針に基づき、実質的推進機関となる大学等研究機関、公的研究資金配分機関、研究コミュニティ及び各業界の経済団体等が関与し、必要に応じてデータシェアリングポリシーを早期に策定することが望ましい。
- データシェアリングポリシーには、日本の国際的な研究開発力を維持、向上するための情報戦略を意識し、「研究データ収集基準」、「研究データシェアリング範囲の設定」、「データ品質保証方法」、「データ格納手法」、「研究データ基盤の整備ルール」などを規定することが必要である。
- 研究データを活用した新たな研究手法の取り組みを推進する世界情勢に対し、日本が取り残され、遅れていくという危機感または切迫感をステークホルダー皆で共有することが重要である。

##### 【公的研究資金配分機関等におけるデータシェアリングのあり方】

- 公的資金による研究成果の論拠となる研究データは、データシェアリングポリシーに基づき、各研究プロジェクト等におけるデータの取り扱いを定めるデータ管理計画書作成などの方策を導入することが考えられる。
- データ管理計画書の提出により、日本の研究データの所在を把握することが可能となり、研究開発における研究データの再利用が加速し研究開発が効率化する。また、結果的に研究者のデータシェアリングに対する意識高揚を促し、その枠組みを強化できる可能性がある。
- 分野による違いはあるが、国内におけるデータシェアリングの需要は存在している。研究プロジェクトにおいて、研究者自らが研究データを提供・管理するという意識を持つことが必要である。一方で、国として研究者に過度な負担を負わせない方策を検討することが必要である。
- データ管理計画の実行には、個々の研究プロジェクト、研究組織のそれぞれのレベルで対象データを格納する基盤整備等が必要であり、その構築、運営にかかるリソースとして、必要とされる費用を直接、間接の研究費から充当する仕組みと、人材確保の枠組みを導入することが望ましい。
- データ管理計画書の項目は欧米の例を参考にして、日本においても同様の項目を設定するとともに、研究データをシェアリングする基盤(リポジトリ)の構築計画についても記載されることが望ましい。参考までに米国NSFの計画書項目を下記に示す。

**【参考】NSF Data Management Plan 記載事項<sup>1</sup>**

- ①データの種類、サンプル、物理的収集物、ソフトウェア、カリキュラム資料等、それらの管理者、セキュリティ事項
- ②データ・メタデータフォーマット、コンテンツの基準(データフォーマット(現行の基準がなかったり適切でない場合は改善案等を記載する))
- ③アクセスおよび共有のための方針(アクセス方法、個人情報、機密事項、知的所有権、法的要求事項に関する方針)
- ④データの再利用、再分配、派生データに関する方針
- ⑤データ、サンプル、他の研究成果の保管およびアクセス保持に関する計画(長期保存、アクセス保持、プロジェクト終了後保存するデータの選択と保管方法)

**【データシェアリングを行うべき分野、領域の選定】**

- わが国として情報戦略に基づいたデータシェアリングの導入により、わが国が世界と比較して強み持つ分野、領域を更に強化し、研究開発力を向上させることが、当該分野、領域の国際的研究開発をリードしていくことにつながる。こうした点から、内閣府総合科学技術・イノベーション会議(CSTI)が主導し、実現可能な重点的分野から順次、データシェアリングを強力に推進し、リソースを集中的に投下することが望ましい。
- 特に、新たな公的研究資金配分機関等が設置される場合や、「革新的研究開発推進プログラム(ImPACT)」、「戦略的イノベーション創造プログラム(SIP)」のように新領域・新分野を対象とする研究助成制度の立ち上げにあたっては、その当初よりデータシェアリングの仕組みを導入していくべきである。
- 選定された各分野の主要学協会、研究コミュニティは、分野毎の専門的、特徴的な事由に応じたデータシェアリング方法を検討し、共有する対象データ、データの公開制限、非公開データの判断基準、データ利用者、シェアリング目的等を明確にした情報戦略を考慮し、あるべき手法、共有ルールを検討していくことが重要である。一方、分野横断した領域や新規領域などは、公的機関が中核機関となっていくことが考えられる。

<sup>1</sup> NSF Grant Proposal Guide, Chapter II.C.2.j  
[http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg\\_2.jsp#dmp](http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp#dmp)

---

**提言2 :わが国としての持続的な研究データ基盤を構築すべき**

---

- 日本の公的研究資金で産出された研究成果データをできる限り幅広く共有し、持続的に情報を蓄積・提供するため、日本におけるデータ基盤を構築し、共有データをマネジメントできる仕組みを持つことが必要となる。データシェアリング手法の導入には、データ基盤の対象データを産出する個々の研究プロジェクトよりも長期的な持続性を持っていなければならない。
- 蓄積された情報が散逸することのないように、持続的にデータ基盤を構築、維持するため、日本におけるデータ基盤構築は、内閣府総合科学技術・イノベーション会議(CSTI)の政策として位置づけることが望ましい。
- データ基盤構築にあたっては、共有されたデータを第三者が安全かつ容易に再利用可能となる相互運用性及び共有対象となるデータの価値を評価、保証するデータ品質管理の仕組みとともに、データのライフサイクルを管理する体制を構築することが重要である。データベース構築に多くの経験を有する公的機関が中核となり、学協会及び研究コミュニティと連携し、データ基盤構築を促進することが望ましい。
- データ構造は多様であり、かつデータ提供者と利用者のニーズをマッチさせることが重要である。利用者ニーズを考慮し、データを有効活用するためのサービスを開発するために、オントロジーや、データ構造化方策についての先行事例を踏まえた RDF(Resource Description Framework)等のセマンティックウェブ技術の活用など、更に研究を進めていく必要がある。

**提言3 : 研究データに係る人材の確保と育成を政策的に推進すべき**

- 研究データシェアリング、データ駆動型研究手法による科学研究活動にあたっては、研究開発プログラムを実施する研究者とは別にインフォマティクスを主軸にした多様な専門人材が必要であり、科学技術イノベーション創出活動を高度化する上で必要不可欠である。こうした専門人材は、インフォマティクスと専門研究分野の知識を併せ持つ研究人材であることが望ましい。
- わが国における研究と教育の中核機関である大学において、研究データの整備からデータ基盤の構築・情報発信までの各プロセスにおける専門性を持つ優れた研究人材の育成が急務であり、大学における育成コース、カリキュラムを制定することが望ましい。
- 優れた研究人材を持続的に育成・確保するには、育成カリキュラム等の高等専門教育と同時に、研究人材が個々の研究プロジェクトや研究組織において、適材適所で活躍可能な将来のテニユアに至る道筋となる明確なキャリアパスを形成する必要がある。大学等研究機関では、実際の研究開発を通じ、研究人材を育成すると同時に、公的研究資金配分機関等は、研究開発プログラムや研究組織内における研究人材をしっかりと位置づける競争的経費研究費改革が求められる。この点において、文部科学省や内閣府総合科学技術・イノベーション会議(CSTI)が主導し、研究開発における制度検討を進めることが望ましい。
- ただし、インフォマティクスと専門研究分野の知識を併せ持つ研究人材の育成には時間を要するため、短期的にはデータ提供者である研究者とインフォマティクス人材の意思疎通を図る仲介役として、コーディネータ的人材を配置することが得策である。公的研究資金配分機関は、研究プログラム等の設計において、こうしたコーディネータ的人材をしっかりと位置づけていくことが重要である。

#### 提言4 : 研究データに係る研究業績評価の仕組みを見直すべき

---

- 研究開発プログラムの研究業績評価において、従来の論文発表による業績評価に加え、研究データの生成・整備・管理・保管・提供又はデータ共有に触発されるデータ引用についても研究者の貢献と認知し研究成果としての評価の対象とすることが、研究者の研究データシェアリングに対する意識を高揚させることになる。また、データ研究、データ基盤構築に係る研究業績が機関業績として評価されることは、研究組織そのもののデータ共有意識の高揚にもつながる。
- データ研究、データ基盤構築に対する業績評価の仕組みの見直しを行うにあたっては内閣府及び各省の研究開発評価指針の改定をすることが考えられる。
- 研究開発評価指針の改定にあたっては、データ研究、データ基盤構築に係る研究業績の新たな評価軸の検討、研究費の一定額を研究データ生成・整備・管理・保管・提供等へバランスよく投入できる枠組みの検討など、当該研究に携わる研究組織、プログラム及び研究者個人についてそれぞれ業績面でしっかり評価できる仕組みを策定することが必要である。

---

**提言5 : 国はデータシェアリングを推進する研究コミュニティに対し、充実した支援サービスを行うべき**

---

- 公的研究資金の配分を受ける研究コミュニティ、研究者は、これまでのような学術論文等による研究発表だけではなく、その論拠となる研究データのシェアリングが科学技術の発展に重要な施策であることを認識することが必要となる。
- 国は、データシェアリングを推進する研究コミュニティ、研究者に対して、相互に連携しつつ、公的機関を通して支援サービスを提供していくことが必要と考えられる。
- 国全体の支援サービスとしては、共通的な研究データ基盤の構築とその持続的な運用、データを利用しやすい仕組みづくり、信頼性を担保するためのデータ品質検証、従来の論文や特許による評価に加え、データシェアリングによって得られるアウトカムを把握し、評価する技術・方法の開発などが考えられる。
- 公的機関が共通的なデータ基盤としての支援サービスを持続的に実施できるよう内閣府総合科学技術・イノベーション会議(CSTI)の政策として位置づけることが望ましい。
- データシェアリングを機会に、わが国の各学協会についても海外の学協会と戦略的に協力・連携し、論文が海外のジャーナルに投稿される場合にあっても論文に直接かかわるデータ、サプリメントデータ等をわが国の学協会に保存するなど、論文と異なる手段による科学技術情報の確保が重要である。わが国の学協会に持続的にデータを格納できる仕組みを確保することは、国際的評価指標において苦戦しているわが国のジャーナルに比して、高い水準にある我が国の科学技術に比肩し得る科学技術情報戦略と戦術を策定する好機であると言える。公的機関は科学技術外交の観点において、それらの活動を費用、人材の側面から支援する方策が考えられる。
- 新領域や新分野におけるデータシェアリングの手法研究を活性化するため、人材の確保と育成の観点を含めて必要な支援策を講じるべき。
- データ処理可能な形式は分野により様々であるが、それらを突合し、つなぐための研究開発も必要である。新領域や新分野におけるデータシェアリングでは、新しいデータ形式、構造など、シェアリングのやり方そのものに研究要素があり、これら研究に対する支援が必要である

#### 4. おわりに

科学技術創造立国を目指すわが国として、研究データシェアリング、データ駆動型科学研究手法による目的実現に向けて議論を加速しつつある世界情勢に乗り遅れることなく、わが国が目指すべき目的の達成にデータシェアリングをどう有効活用するかの方角性について提言をとりまとめた。

データシェアリングを効率的、効果的に推進するためには、まず国や公的研究機関が各分野を主導する主要学協会や研究コミュニティと相互に連携し、わが国における議論を活性化すると同時に研究者への業績評価や研究効率の最大化によるインセンティブによるデータ共有文化の醸成に取り組むことが重要である。

こうした取組みを通し、データシェアリングが将来の科学技術の発展に貢献し、イノベーション創出がより促進されることを期待する。