

国立研究開発法人理化学研究所 創発物性科学研究センター 研究データの管理、公開、共有及び利用に関する実施ガイドライン（創発物性科学研究センター研究データガイドライン）

令和3年3月3日

創発物性科学研究センター

1. 目的

国立研究開発法人理化学研究所（以下、研究所）創発物性科学研究センター（以下、本センター）は、物理・化学・エレクトロニクス of 3 分野にまたがって創発現象の学理の解明及びその応用についての研究を行っている。本センターは、研究活動を通して、以下に定義する研究データを産出する組織である。

本センターは、研究所が定める「研究データの管理、公開、共有及び利用に関する基本方針」（以下、基本方針）を継承し、情報技術を通じた研究所の研究推進及び研究成果最大化に貢献する。創発物性科学研究センター研究データガイドライン（以下、本ガイドライン）は、本センターに所属する研究者等が適切かつ確実に研究データの管理が行えるよう、本センターにおける研究データの取り扱いに必要な事項を定めることを目的とする。

2. 研究データおよび利活用データに関連する用語の定義

[研究データ]基本方針において定義する通り、研究データとは研究所の研究活動を通じて取得、作成されたあらゆるデータをいう。研究データには、実験機器を通して測定、産出されたデータ、計算機を通して生成されたデータが含まれる。ただし、研究の途上において得られるデータでデータ生成者による十分な検証を経ていないデータ、明らかに学術の発展に資しないと判断されるデータは研究データに含めない。また、研究ノート、論文、及びデータベースシステムやソフトウェアコンポーネント等の計算機上で実行されるプログラム類の著作物は、原則として研究データに含めない。本センターにおいては、原則的にすべての研究データを電磁的な方法により管理する。紙媒体のような電磁的な方法で管理できないデータは電磁的な方法により管理できるよう適切に変換するよう努める。研究データは、その利用、収集する状況、由来等に応じて、以下のように分類される。

[一次データ] 研究の過程で新規に取得または生成される研究データを一次データという。一次データには、測定機器、撮像装置等の実験機器から生成され収集されたメタ

データを含む研究データ、研究者自身が計測・考案して作成した研究データ、計算機による数値計算等を通して得られた研究データが含まれる。

[二次データ] 一次データや二次データ、あるいはその双方を加工して生成されたデータを二次データという。

[利活用データ] 基本方針の定義に従い、メタデータを含む研究データのうち、論文発表等に伴い公開が義務付けられたデータ（以下、論文データ）、利活用によって科学技術の発展に貢献できるとデータ管理者及びその所属長が判断したデータを「利活用データ」という。

[ローカル研究データ] 研究所の研究者あるいはその研究者の参画する研究プロジェクト内のメンバーが専ら使用し、第三者からアクセスできないデータ管理領域に置かれる研究データをローカル研究データという。ローカル研究データの管理媒体には、研究者が専ら個人で利用する計算機に接続された固定記憶装置、種々のディスクや不揮発性メモリを備えた記憶メディア、ネットワーク経由でのアクセスが可能で専ら承認された利用者のみがアクセスできる記憶領域等が挙げられる。

[ローカル利活用データ] ローカル研究データのうち利活用データであるものをローカル利活用データという。

[公開共有フラグ] 研究データリポジトリに登録された利活用データについて、その開示先を指定するための情報を公開共有フラグという。基本方針に従い、公開共有フラグには、「非公開」、「所内限定公開」、「所内外限定公開」、「公開」が含まれる。各公開共有フラグについては、以下の「6. 公開共有フラグの変更」において定義する。

[データ生成者] 研究データを生成した者をいう。

[データ管理者] データ生成者により生成された研究データを、情報基盤や情報媒体を用いて管理する者をいう。データ管理者は上記データ生成者と同一の者であっても、別の者であってもよい。

[データ利用者] 利活用データを受領し利用する者をいう。

3. 研究データリポジトリの選定及び利用

研究データの管理には、原則として、研究所が整備する研究データリポジトリ（以下、理研リポジトリ）を利用する。研究データを他の管理領域に保管する場合は、そのデータへのリンクを理研リポジトリに登録することが望ましい。

理研リポジトリ機能の増強を研究所に要請するには、客観性、定量性を持った根拠が必要となる。本センターは常に理研リポジトリの実効的な運用が行えるよう、本セン

ター内で産出される研究データの質や量を正確に把握するよう努める。

4. 個人情報等の取り扱い

本センター内で生成された、あるいは研究所外から受領した個人情報を含む研究データの取り扱いについては、関係法令・ガイドライン・倫理指針や、研究所が定める個人情報等の取り扱い規程等に従う。

5. 研究データの理研リポジトリへの登録

研究の過程で取得、生成、管理される研究データは、一次データの取得・生成に始まり、共有や保管のための理研リポジトリへの登録、公開、削除に至る一連の流れが典型的なライフサイクルである。このライフサイクルをより具体的に記述すると以下の通りとなる。

a. 一次データの生成と管理

データ生成者により生成された一次データを、データ管理者がローカル研究データとして管理する。

b. 二次データの生成と管理

データ生成者が一次データあるいは二次データに解析・加工等の処理を施し、二次データを生成する。データ管理者は、生成された二次データをローカル研究データとして管理する。

c. 利活用データの選択

データ管理者が、ローカル研究データとして管理されている一次データあるいは二次データから、保管・共有・公開の対象となるデータを選択し、ローカル利活用データとする。

d. 理研リポジトリへの登録

データ管理者は、理研リポジトリが提供するアップロード機能を用いて、ローカル利活用データを理研リポジトリにアップロードし登録する。登録の際、不用意にデータ公開が行われないよう、公開共有フラグは「非公開」を設定しておくことが望ましい。

e. 公開共有フラグの設定及び変更

データ管理者は、後述「6. 公開共有フラグの変更」に記した手続きにより公開共有フラグを設定、変更する。

f. 利活用データの削除

データ管理者は、後述「13. 利活用データの削除」に記した手続きにより利活用データを削除する。

研究データから利活用データを選択する際に、データ管理者及びその所属長は、データ保管あるいは他者とのデータ共有の必要性を、研究の状況、研究データの質、研究知財の管理、論文根拠としての必要性等を含む多様な側面から、研究者としての良心に従って総合的に評価して判断する。具体的には、利活用データとして選択されるデータとして、

- ・論文あるいは実験ノートに記載があり、保管及び公開が必要なデータ
- ・共同研究プロジェクト内での共有が予定されているデータ
- ・公開することにより学術の発展に資する可能性のあるデータ
- ・公開することによりデータ生成者以外の研究者による重複したデータの生成を避けられる可能性のあるデータ

などが挙げられる。

ここで、理研リポジトリの管理、維持のためには相当の費用がかかることも考慮する。データ流通や利活用に支障のない範囲でデータを割愛・圧縮する、プログラムで生成可能なデータはその元データのみに限る等の検討を積極的に行い、費用削減による実効性向上に努める。

ローカル研究データ、ローカル利活用データについては、当該のデータ管理者が責任をもって管理する。それらのデータを理研リポジトリ以外の情報基盤や媒体で管理する場合には、情報媒体の盗難、紛失、第三者からの不正アクセスを防止する措置、廃棄時のデータ漏洩対策を適切に施す。

6. 公開共有フラグの変更

理研リポジトリに登録された利活用データの公開共有フラグの変更は、データ管理者が研究動向やデータ利用者を見極め、所属長の承認を得て行う。公開共有フラグは、対象となるデータ利用者の範囲が狭い順に列挙すると以下の通りとなる。

[非公開] データ生成者、データ管理者及びその所属長のみがアクセス可能なデータに付与される。

[所内限定公開] 研究所内のデータ利用者が読み込むことができるデータに付与される。研究所内の全ての者を対象とすることも、パスワードや IP アドレスによるアクセス制限などの手法を適用してその中の一部の者のみを対象とすることもできる。

[所内外限定公開] 研究所内外を問わず、単数あるいは複数のデータ利用者を対象にしたデータに付与される。パスワードや IP アドレスによるアクセス制限などの手法を適用することもできる。

[公開] データ利用者を限定せず、何人も読み込み可能なデータに付与される。

ここで、データ提供の手段について、データ管理者は、データ特性等に応じて個別に選択する。例えば、常にネットワーク等を通じて読み込み可能な状態にしておく必要はなく、書面を通じて読み込み依頼を受け、情報媒体を用いて提供することも可能とする。

公開共有フラグは設定後、常に研究動向、利活用データの価値、共同研究の進展、公開の有効性等を総合的に勘案し、適切な設定となるよう努め、必要に応じた設定変更を所属長の承認の下でデータ管理者が行う。

7. 利活用データの公開猶予期間

利活用データは、当該データが生成された時点から 2 年以内あるいは利活用データとして選択された時点から 1 年以内のいずれか早い時点で、原則公開する。しかしながら、当該利活用データに関係する研究実施時に契約や取り決めがある場合には個別に対応する。さらに論文データについては、論文掲載条件に論文データ公開猶予期間が規定されている場合には、当該規定に従う。

8. 利活用データの利用許諾

データ管理者は、データの知財権を守るため、公開共有の範囲（公開共有フラグ）に応じた適切な利用許諾条件をデータの受取先に示すよう努める。共同研究契約や MTA(Material Transfer Agreement)が締結されている場合には、それに従った利用許諾条件を設定・提示する。データの利活用状況を把握したい場合には、受取先がデータをダウンロードする前に IP アドレス等の記録とともに、個人情報であることに留意しながら受取先の氏名や所属などの情報の提供を受取先に求める。さらには、受取先でデータを利用した結果を論文等で発表する場合には、論文の本文や謝辞、あるいはレファレンスにおいて提供元情報の掲載を義務付けることを利用許諾条件中に明記する。利用許諾条件の作成に当たっては、データ生成者やデータを参照する識別子を論

文に掲載するように求める条文を含める等、論文の読者が引用元データを参照できるよう配慮する。一方で、プログラムによるデータ利用なども含め、障壁なくデータの読み込みを可能としてより多くの受取先にデータを提供したい場合には、オープンデータに適したクリエイティブコモンズライセンス等の利用許諾条件を利用する。

9. 利活用データの保管期間と理研リポジトリでの管理方法

基本方針に従い、理研リポジトリに登録されるメタデータを含む利活用データの保管期間は、登録されてから原則として10年間とする。利活用データ保管期間中に科学技術が進歩あるいは潮流が変化する等の事情により、当該利活用データの利用価値が変動することが想定される。このため、データ管理者は、定期的に理研リポジトリに登録された利活用データを確認し、当該データに最も適合した公開共有フラグと利用許諾条件の設定に努める。

10. 公共リポジトリの利用

公開データの取り扱いにおいては、研究分野やデータの種類によっては公共リポジトリの利用が好ましい場合がある。公共リポジトリが定めるデータ保管・公開期間や利用許諾等の条件を総合的に勘案し、公共リポジトリの利用が適切な場合には公共リポジトリに公開データを登録する。この時、メタデータについては理研リポジトリに登録し公開する。

11. メタデータの付与

既に大小を問わず国内外の研究機関がウェブ上でメタデータの公開を開始しており、メタデータがつなぐ研究データのグローバルネットワークが形成されつつある。本センターもこの潮流に沿って、オープンサイエンスや研究データ利活用推進のため、利用価値の高い研究データを選んで高品質なメタデータを付与するよう努める。メタデータの記述には、グローバルな相互運用性を担保する点でも標準化されたオントロジーやデータ項目を活用することが望まれる。データ生成者あるいはデータ管理者は、情報システム本部の担当部署が提供、推奨するオントロジーやツール類を活用し、その技術支援を得ながら世界最高品質のメタデータの作成に努める。

メタデータの生成は、対象となるデータが生成された時点で行うことが、データ管理のコストや便宜を考慮すれば最も好適である。しかしながら、メタデータは利活用データの選定時やデータ共有・公開時に付与されることも多く、メタデータの生成に大きな労力を要する場合も多い。このため、本センターでは、研究計画の段階でメタデータの詳細度、付与する時期等も考慮するよう努める。

メタデータが実験条件、計算条件、解析方法等の科学的知見の詳細に及ぶ場合、メタデータ自体が価値の高い研究データとなりうる。この場合には、メタデータではない利活用データと同様に、適切な公開共有フラグや利用許諾条件を付すよう努める。データのカatalogとして作成されたメタデータは、プログラム等から機械的に検索が行えるようなクエリの標準技術の普及を考慮して、所属長の承認の下でクリエイティブコモンズライセンスを付し、アクセス障壁を低く抑えながら公開する。

データとそれを説明するメタデータとは密接な関係にあるが、公開共有の設定は異なってもよい。例えば、データの知財権を保護しながらもメタデータを活用して研究内容の公開を積極的に行う場合を考慮して、メタデータは公開であるが、データは非公開であることも許容される。さらに、データ公開を緊急に行う場合で、詳細なメタデータの付与に時間を要する場合には、データ生成者あるいはデータの所在等の最低限のメタデータを付与した上でデータを先行して公開することも許容される。

12. データ管理者の転出対応

理研リポジトリに登録された利活用データの管理は、原則としてデータ管理者が責任を持って行う。データ管理者が本センターから転出する場合には、その所属長が管理責任を請け負う。ここで請け負った所属長は、別の担当者に管理を依頼することができる。誰も管理責任を負えない状況となる利活用データについては、事前に情報システム本部の担当部署に相談し、その管理方法について決定する。

13. 利活用データの削除

基本方針に従い、メタデータを含む利活用データは10年間を保管期間とし、保管期間終了後はデータ管理者が所属長の承認の下で削除する。しかしながら、学術上の重要性に鑑みて保管期間を延長することが望まれる場合には、所属長の承認を得て、期限をつけて保管期間を延長する。

他方、保管期間前に削除することが妥当と判断される利活用データは、妥当性を結論づける理由を添えて所属長の承認を得たうえで理研リポジトリから削除するとともに、その理由を理研リポジトリ上に削除前と同一の公開共有フラグを設定して登録する。

データとそのメタデータとは同時に削除する必要はなく、それぞれ適切な削除計画を立てるよう努める。

14. 免責事項

利活用データの公開・利用について、本センターの免責事項を以下の通り定める。

- a. 本センターは、利活用データの完全性・正確性・有用性・安全性等の品質の担保に努めるが、いかなる保証もしない。
- b. 利活用データを利用したことや利用できなかったこと、あるいは利活用データを用いてデータ利用者が下した判断や起こした行動により生じたいかなる結果・損害について、本センターは責任を負わない。
- c. 事前の予告なく利活用データの改編、修正、削除、公開の中止を行う場合がある。

15. 研究者等の評価及び報奨

本センターは、研究データを取得・生成し、さらに利活用データとして他者と積極的に共有した者について、当該活動を業績として考慮し、適切な評価及び報奨を行う。

以上