

# 国立研究開発法人理化学研究所 革新知能統合研究センター 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（AIP 研究データガイドライン）

令和3年3月5日

革新知能統合研究センター

## 1. 目的

国立研究開発法人理化学研究所（以下、研究所）革新知能統合研究センター（以下、AIP）は、革新的な人工知能基盤技術を開発し、それらを応用することにより、科学研究の進歩や実社会における課題解決に貢献することを目指すとともに、人工知能技術の普及に伴って生じる倫理的・法的・社会的問題に関しても研究している。これらの研究開発においては、内部で生成したデータだけでなく、様々な企業・大学・研究所・プロジェクト等の連携先から取得したデータも活用している。

AIP 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。AIP 研究データガイドラインは、AIP 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、AIP における研究データの取り扱いに必要な事項を定義することを目的とする。

## 2. 研究データおよび利活用データの定義

[研究データ] 基本指針において、研究データとは研究所の研究活動を通じて取得、生成されたあらゆるデータをいう。AIP においては、原則として研究データを電磁的な方法により管理する。電磁的な方法で管理できない紙媒体などのデータは可能な範囲で電磁的な方法により管理できるよう適切に変換する。研究データはメタデータを含むが、原則として研究ノート、論文、およ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は含まない。研究データには、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。

[利活用データ] 基本指針の定義に従い、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）および利活用によって

科学技術の発展に貢献できると AIP が研究所を代表して判断したデータを「利活用データ」という。

[ローカル研究データ] AIP の研究者あるいは研究プロジェクトが専ら使用し、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、AIP の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、種々のディスクや非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。

[開示フラグ] 利活用データの開示先を指定する情報をいう。基本指針に従い、開示フラグには、「非開示」、「所内限定共有」、「所内外限定共有」、「公開」がある。

[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よいし、別の者であってもよい。

[データ利用者] 利活用データを受け取り利用する者をいう。

### 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、研究所リポジトリ)を利用する。研究データの本体を他のストレージに保管する場合は、そのデータへのリンクを研究所リポジトリに登録することが望ましい。

研究所リポジトリの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。AIP は研究所リポジトリの実効的な運用を確保するため、AIP 内で生成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. 個人情報の取り扱い

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やガイドラインに従う。特に、人を対象とする医学系研究等の実施に当たっては要配慮個人情報を含む研究データを厳重に取り扱う。

さらに研究所外から受領したデータについては、上記規定等を含む情報管理に関するガイドラインに沿って処理を施したのち、研究データとして扱う。

## 5. 研究データの研究所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、データの取得・生成に始まり、共有や保管のために研究所リポジトリへの登録、共有、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。ここで、データ生成者とデータ管理者は異なっている場合でも同一の人物であってもよい。

### a. データの取得・生成

取得・生成した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. 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。

### c. 研究所リポジトリへの登録

データ管理者は、研究所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研究所リポジトリにアップロードして登録する。登録の際、不用意にデータ開示が行われないよう、開示フラグは「非開示」を設定しておくことが望ましい。

### d. 開示フラグの設定

後述の手続きにより開示フラグを設定する。

### e. データの削除

後述の手続きによりデータを削除する。

研究データから利活用データを選択する際に、データ生成者あるいは管理者が、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、利活用データは下記を含む。

- ・ラボノートに記載があり、保管が必要となっているデータ
- ・共同研究で共有が予定されているデータ
- ・研究の次の段階で必要となるデータ
- ・将来の論文掲載、根拠に必要なデータ

ここで、研究データを研究所リポジトリに保管する際には研究所リポジトリの管理、維持のため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する、圧縮する、プログラムで生成可能なデータは保管しないなど、費用削減や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。それらのデータを研究所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. 開示フラグの変更

開示フラグ「非開示」を付して研究所リポジトリに登録された利活用データは、データ生成者あるいはデータ管理者が、研究動向やデータ利用者を見極め、開示フラグの変更を所属長の承認を得て行う。開示フラグは、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。

[非開示] データ生成者あるいはデータ管理者のみがアクセス可能なデータに付与される。

[所内限定開示] 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。研究所内の者すべて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。

[所内外限定開示] 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。

[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。

データ提供の手段については、データの特性等に応じて個別に選択する。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、例えば、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。

開示フラグは、研究動向、利活用データの価値、共同研究の進展、共有の有効性等を常に総合的に勘案し、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、原則として公開する。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

データの著作権を守るため、開示フラグ(開示の範囲)とともに適切な利用許諾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を求める。

受取先を特定しないで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 アドレス等を記録するとともに、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成果を論文等で発表する場合には、論文の本文や謝辞にて提供元情報の掲載を義務付ける利用許諾を明記する。

利用許諾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む自由な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を利用する。

## 9. 利活用データの保管期間と研究所リポジトリでの管理方法

基本指針が定める研究所リポジトリに登録される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年間以上である。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ため、データ管理者は、定期的に研究所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するように開示フラグと利用許諾を修正するよう努める。

## 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、研究分野やデータ種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公開期間や利用許諾等の公開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研究所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開示を始めており、メタデータがつながり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。AIP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。研究データを生成した研究者は、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の活用、技術支援を得ながら世界最高品質のメタデータの生成に努める。

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、利活用データの選定時やデータの開示時にメタデータが付与され、メタデータの生成に要する労力が大きくなる場合が多い。このため、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しながら研究データの取り扱い計画を立案するよう努める。

メタデータが実験条件や観察された現象等の科学的知見の詳細を記述する場合、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、適切な開示フラグや利用許諾を付するよう努める。データのカatalogとして生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、所属長の了承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。

データとその注釈であるメタデータとは密接な関係にあるが、開示の設定は異なってもよい場合がある。例えば、データの知財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など、メタデータは公開するがデータは公開しないこともあり得る。さらに、データ開示を緊急に行う必要があるが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てデータを先行的に開示することも許容される。

## 12. データ管理者の転出対応

研究所リポジトリに登録された利活用データの管理は、原則として管理者あるいは管理者の所属長が責任を持って行う。管理者が AIP から転出する場合には、管理者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った利活用データについては、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. データの削除

基本指針に従い、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許可を得たうえで研究所リポジトリから削除するとともに、その理由を研究所リポジトリ上に削除前と同一の開示フラグを設定して登録する。

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

#### 14. 研究者等の評価及び報奨

利活用データを生成、取得し、さらに他者と積極的に共有した者については、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