

大学における研究データ管理実施要領ガイドライン

目 次

1.	大学における研究データ管理の実現に向けて	1
(1)	大学において管理する研究データ	1
(2)	大学における研究データ機関管理の形態	2
(3)	大学における研究データ管理実現のための制度設計	3
(4)	大学における研究データ機関管理のための学内体制	4
(5)	大学における研究データ管理のための情報基盤	5
(6)	大学における研究データの機関管理コスト	6
(7)	その他の留意点—大学が十分なリソースを有さない場合	7
	(参考資料) AXIES-RDM 部会研究データポリシーWG メンバー一覧	8

※ 本文書の内容は、当初、「大学における研究データポリシー策定のためのガイドライン」の第 5 章として作成されたものである。

大学の、特に執行部が、研究データポリシーを策定したところで満足してしまうのではなく、学内の研究データ管理体制の整備まで検討することを願って、作成したものであるが、策定した研究データポリシーに基づいた研究データ管理を実施するための具体的な事項について言及することになり、いささか長くなってしまった。

このため、元の第 5 章を切り離し、「大学における研究データ管理実施要領ガイドライン」として別途取りまとめることが、検討チーム内で急遽提案された。

本文書はこのような経緯のため、前文もなく、やや乱雑であるが、今後、AXIES-RDM 部会等の議論を経て、充実させていく予定である。

ぜひ、多様な意見を送っていただきたい。

(AXIES-RDM 部会 研究データポリシー検討 WG)

大学における研究データ管理実施要領ガイドライン

1. 大学における研究データ管理の実現に向けて

大学において策定される研究データポリシーは抽象度が高いため、ポリシーを策定しただけでは学内に研究データ管理体制を敷くことができない。学内の研究データ管理に向けた方法論の概要は、AXIES-RDM 提言のⅣ－Ⅶ章を参照されたい。

以下には、大学における研究データ管理体制の実現に向けた、いくつかの具体的な事項について述べる。

(1) 大学において管理する研究データ

研究データポリシーは、基本的には、研究過程において生じるあらゆるデータを対象とする(3(6)b節参照)。研究者が、研究活動期間を通じて研究データを管理するという意味ではそれは正しい。しかし機関としては、自機関のもてるリソースの範囲と、大学の戦略性をバランスさせて、管理する研究データの範囲を決めていく必要がある。

表 5.1 に、大学において機関管理される研究データの種類と段階を挙げた。AXIES-RDM 提言「II. 学術機関における研究データ管理の導入目的の実際」に記したように、大学は、それぞれの特質や考えに基づき、機関管理する研究データの対象と付随する学内サービスを定め、展開していく必要がある。

表 5.1：大学において機関管理される研究データの種類

① 機関管理が必須のデータ (文書管理、研究データ 10 年保存ルール等への最低限の対応) ② 研究者の要望に応じて、大学が保存・公開するデータ - 機関リポジトリを通じて公開されるデータ - 論文投稿時に必要とされる、根拠データ¹ ③ 堅実な学術継承のために保全するデータ (データアーカイブ。特に公的資金を得て取得されたデータ) ④ その他、機関が定めたデータ - 機関のヴィジビリティ向上に資するデータ - 外部機関との連携（共同研究、産学連携）に資するデータ - 研究公正を担保するためのデータ 等
⑤ 機関が提供する保存領域に保存されるデータ (研究者の研究活動時に利用される保存領域。 機関は保存領域を提供するのみで、データは研究者が管理 研究不正発生時に強制介入できる権限が機関にあると望ましい)

(2) 大学における研究データ機関管理の形態

もう一つ踏まえるべきは、研究データの種類や管理する場面によって、機関管理の形態が異なるということである。公開の度合い、管理する対象、研究データに関連して、機関が提供する機能などが異なる。

公開の度合いについては、「公開」か「非公開」があるが、研究データが機微な情報を含んでいる可能性を踏まえると、その中間段階の公開方法の方が、現実的な可能性がある。中間段階としては、機関が認めた特定の人や層のみに研究データを公開する「制限公開」や、コンテンツの外形的情報のみを公開する「メタデータ公開」などがある。

管理する対象としては、研究データの容量や管理体制の問題から、研究データそのもの（「コンテンツ」）の管理は研究者に委ね、大学は「メタデータ」のみを管理する場合がある。

研究データに関連して、機関が提供する機能としては、「メタデータ管理」、「保存」、「公

¹ 論文投稿時に必要とされる、根拠データの保存・公開：多くの国際的学術雑誌が近年、研究の再現性確保の観点から、要求する。このように提出される研究データ（data supplement）は、論文査読前で、他者が再利用するレベルにまで至っていないこともあることから、機関リポジトリとは分けて、管理されることが望ましい。また、機関リポジトリで公開するデータに比べて、データの外形的情報（メタデータ）に厳密性を要求されないため、研究者自らがデータのアップロード等の作業（self-archive）を行うのでも良い。

開」、「保全/長期保存」などがある。「メタデータ管理」のみであれば、コンテンツの保存領域は最小で済む。ただし、必要な時に研究データに即座にアクセスできる体制や、研究データを管理する研究者が異動/離職/退職する場合の手続きを定めておく必要がある。なお、「保全/長期保存」は、10年以上の半永久的保存を指す。

表 5.2：大学における研究データ機関管理の形態

大学が機関管理する研究データの対象と形態										
大学が機関管理する研究データ	公開の有無				機関が管理するデータ			機関の有すべき機能		
	非公開	メタデータ公開	制限公開	公開	メタデータ	コンテンツ	管理メタデータ	保存	公開	長期保存 / 保全
① 機関管理が必須のデータ	●	(○)	(○)	(○)	●	●	●	●	(○)	(○)
② 研究者の要望に応じて、大学が保存・公開するデータ										
機関リポジトリを通じた公開	○	●	○	●	●	●	●	●	●	(●)
論文投稿時に必要とされる、Supplemental Dataの公開	○	○	○	○	●	●	●	●	○	(○)
③ 堅実な学術継承のために保全するデータ	○	●	○	●	●	●	●	●	●	●
④ その他、機関が定めたデータ										
機関のヴィジビリティ向上に資するデータ	(○)	●	○	●	●	○	●	●	●	(●)
外部機関との連携（共同研究、産学連携）に資するデータ	(○)	●	○	●	●	○	●	●	●	(●)
研究公正を担保するためのデータ 等	(○)	(○)	(○)	(○)	(○)	(○)	(○)	(○)	(○)	(○)
⑤ 研究活動中に、機関提供の保存領域に保存されるデータ										
研究者に提供される、研究活動中の保存領域内のデータ	●	-	-	-	-	●	-	●	-	(○)
研究者に提供される解析ツールの仮保存領域内のデータ	●	-	-	-	-	●	-	●	-	-

注) ●：主要機能、○：付随してもよい機能、-：不要な機能。括弧付きは、一定条件下においてありえる機能。

(3) 大学における研究データ管理実現のための制度設計

大学における研究データ管理のための枠組みは、策定する研究データポリシーにより与えられる。しかし学内に研究データ管理を実現するためには、表 5.3 に挙げるような課題が複数立ち上がる。これらについては現段階では、広く通用する解決方法は存在しないが、今後各種コミュニティにおいて議論されていく必要がある。

表 5.3：大学における研究データ管理実現のために必要な検討項目

① 研究者の離職・退職時の研究データの取扱い

研究データ（特に公的資金を得たもの）が、人類に永続的に利活用できるようにするためには、研究者の離職・退職時に、研究データが機関の管理下に置かれ、保全されることが望ましい。この場合の研究データの権利関係や管理コストの取扱いも含めた制度設計が必要である。

② 一定期間を経た研究データの廃棄の考え方

研究データは、保管コストの問題もあるため、いずれかの時点で破棄も検討されなければならない。ただし貴重・希少なデータについては、可能な限り、一律の破棄が回避されることが望ましいが、「誰が」「いつの時点で」「どのような基準のもとに」その判断をするかが悩ましい。

③ 研究データ管理におけるコスト負担の考え方

研究データが大学の情報基盤に保存・管理される場合の保管コストは、大学が全額負担することは難しい場合が多いこと、研究分野によって必要とされる容量が大きく異なること、また、研究活動のための外部資金は研究者が獲得していることから、一定の受益者負担のルールが考えられて良い。ただし、研究プロジェクト終了後は、受益者負担が難しいことは踏まえる必要がある。

④ 大学が研究データを提供する場合の、研究データの権利関係の考え方

大学が、研究者に代わって研究データを外部提供する場合、研究データの権利関係が整理される必要がある。研究データのクレジットの付し方、問い合わせへの対応、利益が出た場合の利益分配の方法などが課題となる。研究者が他機関に異動する場合の扱いや、逝去後の扱いも検討される必要がある。

⑤ 大学が外部提供（公開）する研究データの質保証の考え方

機関リポジトリ等を通じて公開される研究データについて、大学は外形的なチェックはできるが、その中身の正確性や利用可能性については、当該データを生成した研究者に委ねるほかない。一方、データ公開後に不備が発覚した場合のバージョン管理の方法、外部クレームへの対応、不備が故意だった場合の研究者への対応などが課題となる。

(4) 大学における研究データ機関管理のための学内体制

大学が研究データを機関管理する場面は、AXIES-RDM 提言の「V. 学術機関における研究データ管理のための仕組み」に挙げた 10 程度のシナリオが考えられる。これらの多く

の場合、1) 研究者からコンテンツとメタデータを受け取り、2) 事務局において管理し、3) 必要に応じて外部利用者に公開・提供するための窓口が必要となる。

大学が機関管理する研究データは、外部にその存在が知られ、利活用されてはじめて、管理・保全した意味が生じるため、3) は忘れがちであるが、確実に整備する必要がある。加えて、外部の者が研究データを発見し再利用できるように、「データキュレーション」を行う体制が必要となる。データの生成者や生成日などの外形的情報は、機関リポジトリを通じて公開するのであれば一般に、大学図書館にて付される。他方、研究データの説明情報等、専門分野の知識を必要とするキュレーションは、研究者に近い研究支援者により行われる必要がある。

AXIES-RDM 提言のV章には、研究データ管理の場面別に、関係する可能性のある部署を記している。学内に特定の研究データ管理を導入する場合は、関係する可能性のある部署で十分に議論し、学内手続き（ワークフロー）と必要となる情報基盤をデザインする必要がある。

(5) 大学における研究データ管理のための情報基盤

大学における研究データ管理に必要なデジタルプラットフォームの機能要件は、AXIES-RDM 提言のVI章に記した。それら機能は、表 5.4 に挙げた情報基盤の組み合わせにより実現する。

表 5.4 : 大学における研究データ管理のための情報基盤

① 記憶媒体（ストレージ）

研究データを管理・保存・公開するための記憶媒体。記憶媒体は、機関にて調達する場合（オンプレミス）と、クラウド等外部記憶媒体を利用する場合とある。また、保存される研究データの利用頻度に応じて、ホット/コールドストレージを使い分けることができる。

② 研究者向け研究データ基盤

研究者が研究活動中に利用する保存領域の管理を可能とする基盤。ユーザの認証・認可を通じたアクセスコントロール、研究プロジェクトごとのファイル管理、ファイル編集に関わる証跡管理などの管理を可能とする。研究データ管理計画と連動した機能を有しても良い。

③ 機関向け研究データ管理基盤

研究不正防止や研究データの長期保存を目的として、機関が研究データを管理するための基盤。研究データが保存される保存媒体や研究者向け管理基盤と連動し、メタデータレベルの管理を可能とする。必要に応じて、研究データへの直接アクセスや、管理する研究データの評価・分析も可能とする。

④ 研究データ公開基盤（機関リポジトリ、データアーカイブ、データカタログ等）

機関や研究者が研究データを公開することを可能とする基盤。研究データへのメタデータや永久識別子（DOI）の付与、ワークフロー管理を含む、研究データの管理を管理者に可能とし、メタデータレベルの検索を外部利用者に提供する。研究データの公開、限定公開、メタデータのみ公開、非公開を可能とし、限定公開およびメタデータのみ公開については、研究データの利用申請も可能とする。管理する研究データおよび利用状況に関する評価・分析も可能とする。

⑤ 研究データ解析基盤

研究者が研究データを解析することを可能とする基盤。標準的な解析ツールや機関が提供を望む解析ツールを装備する。研究者向け研究データ管理基盤を介して、研究データの入出力を行う。実験装置や大型計算機から出力されたデータを、解析することも可能である。

（注）ここに挙げた情報基盤の多くは、国立情報学研究所が提供する NII Research Data Cloud（NII-RDC）で実現可能である。但し「①ストレージ」は、各大学にて整備要。

(6) 大学における研究データの機関管理コスト

大学において研究データを機関管理するためには、表 5.5 に示すようなコストが発生する。

表 5.5：大学における研究データの機関管理コスト

- | |
|-------------------------------------|
| ① 記憶媒体（ストレージ）の導入・維持経費 |
| ② 研究データ管理および公開のためのシステム導入・維持経費 |
| ③ 研究データの管理コスト（事務経費、研究者による管理費） |
| ④ 研究データの公開・外部提供コスト（事務経費、広報・外部連携費等） |
| ⑤ 研究データの解析環境の提供コスト（ソフトウェア費用、導入・運営費） |
| ⑥ 研究データ管理の規定や手続き整備等、全体統括費（事務管理費） |
| ⑦ 研究データ管理の学内周知と手続きコスト（事務管理費） |
| ⑧ その他、都度判断を必要とする事案への対応コスト（事務管理費） |

コストの多くは、機関において発生する事務管理費、研究データ管理のためのシステムやソフトウェアなどの導入・維持管理費、外部利用者との広報・外部連携経費、そして、研究データを生成した研究者において研究データを成形し説明情報等を付加するのにかかるデータキュレーションのためのコストである。

事務管理費は一般に見逃されがちであるが、全学的な負担となるため、十分に視野に入れる必要がある。システムやソフトウェアなどの導入・維持管理費は一般に情報基盤センターにて手配されるが、費用は全学にて負担される。広報・外部連携費は、節約の観点か

ら縮小されることが多いが、研究データは外部に利用されて初めて、それまでに至る管理コストをかけた意味が生じることを認識する必要がある。併せて、研究データが外部に利用されるためには、研究データが十分に成形され、説明情報も付加される必要があり、これは研究者自らが作業をするか、データキュレータ等を雇用して行う必要がある。

このように、機関内で生成された全ての研究データを維持・管理するのは相当のコストが発生すると想定されることから、各機関において、機関管理対象とする研究データを一定程度制限するなどの方策が考えられる。

なお、記憶媒体（ストレージ）や、研究データ管理のためのシステムやソフトウェアの利用の範囲は、学問分野や研究者の研究テーマ、研究手法により大きく異なるため、機関がすべて負担するのではなく、一定容量を超える分については研究者が負担する従量課金制とすることも考えられる。ただし、研究プロジェクト終了後、または研究者の異動や退職の後に、研究データの長期保存や公開にかかる費用を当該研究者が直接負担することは原理的に難しいことにも留意しなければならない。

このように、研究実施期間中および研究終了後の研究データの保存・管理、そしてその共有・公開と再利用を成り立たせるための、持続可能な費用負担モデルを、機関ごとに確立する必要がある。

(7) その他の留意点—大学が十分なリソースを有さない場合

大学が研究データの保存・管理や、一般への利活用の供与のためのリソースを十分に有さない場合がある。そのような場合においても、大学がコンプライアンスの観点から最低限管理しなくてはならない研究データは、機関において保存・管理の義務がある。文書管理や研究データ 10 年保存に基づく、研究データの管理である。

一方、この最低限の範囲を超えて、大学が研究者の要望に応じて保存・公開する研究データについては、大学外にある分野別あるいは一般的なデータリポジトリを用いた公開を、機関が研究者に勧めることもありえる。ただし、クラウド上のフリーの保存領域を研究者が利用し、研究データが散逸することは望ましい状況ではないため、他大学と連携した研究データ保全の仕組みなども検討されて良い。

なお、国立情報学研究所の運営する機関リポジトリクラウドサービス JAIRO Cloud は、大学の規模に応じた価格設定で、比較的安価かつ容易に大学に導入可能なため、その利用が検討されても良い。

(参考資料) AXIES-RDM 部会メンバー一覧