「研究データ管理サービスの設計と実践」

全体スクリプト

第1章　序論

1.1 第1章の概要

研究データには、どのように生成され、保存され最終的に再利用されていくのかといったライフサイクルがあります。この講座では、研究データのライフサイクルに沿った形で、サービス設計や研究前の支援、研究中の支援、研究後の支援、そして日常的な支援について学びます。学習中は、研究データのライフサイクルのどのプロセスにあたる部分を学んでいるのかをこの図を見ながら常に意識してください。

第1章では、研究データ管理の基礎的な知識として、研究データ管理の背景、研究データ、研究データ管理、研究データ管理サービスについて学びます。

まず、「研究データ管理の背景」では、大きく３つのことを学びます。

一つ目は、オープンサイエンスの潮流です。論文や研究データへのオープンアクセスによる新たなイノベーションの可能性に鑑み、国は、資金配分機関、大学等（など）の研究機関、研究者等と連携し、推進体制の構築を図るとしています。

二つ目は、研究不正防止への期待です。国は、研究不正防止における研究データ保存の必要性を明確に認め、日本学術会議ではデータの保存期間を「原則、論文発表後10年」としています。

三つ目は、研究助成機関からの要求です。近年、説明責任や助成の投資効率向上を目的として、多くの助成機関がそのポリシーにおいて、適切な研究データ管理を求めています。助成申請時に「データ管理計画」の提出が義務とされることもあります。

次に、「研究データとは」では、研究データの定義や具体例を踏まえ、研究データの「生成」、「加工」、「分析」、「保存」、「公開」、「再利用」という一連の「研究データのライフサイクル」を学びます。そして、各段階で行われるデータ管理の具体的な取り組みを概観します。

そして、「研究データ管理とは」では、研究データ管理の定義やメリットを踏まえ、研究データの再利用を推進するための指針である「FAIR（フェア）原則」をとりあげます。「Findable（ファインダブル）=発見可能性」、「Accessible（アクセシブル）=アクセス可能性」、「Interoperable（インターオペラブル）=相互運用可能性」、「Re-usable（リユーザブル）=再利用可能性」の4つの原則は、研究データ管理に取り組む上で重要な原則です。

また、「研究データ管理サービスとは」では、研究データ管理を支援するための具体的なサービスを学びます。研究データ管理サービスには、ポリシーの策定、研究前、研究中、研究後、そして日常的な支援まで、研究データのライフサイクルに沿ったさまざまな支援があります。これらのサービスは、ストレージなどの管理基盤の支援とスタッフが直接提供する人的支援に分類することができます。

それでは、第1章を詳しく学んでいきましょう。

研究データのライフサイクルと本教材との関係

研究データには、データが生成され、保存され、最終的に再利用されていくまでのライフサイクルがあります。本教材では、研究データのライフサイクルに沿った形で、サービス設計や研究前の支援、研究中の支援、研究後の支援、そして日常的な支援について学びます。

学習中、研究データのライフサイクルのどのプロセスにあたる部分を学んでいるのかを、この図を見ながら常に意識してください。

1.2 背景

初めに、近年、研究データ管理の重要性が増している背景として、オープンサイエンスの潮流、研究不正対策、研究資金の助成機関のポリシーについて学びます。

1.2.1 オープンサイエンスの潮流

2016年1月に内閣府が策定した第5期科学技術基本計画には、オープンサイエンスの推進と、国の基本姿勢が述べられています。

基本計画によれば、オープンサイエンスとは、主に論文へのオープンアクセスと研究データのオープン化、すなわちオープンデータを含む概念です。オープンアクセスが進むことにより、研究成果の幅広い活用が可能となり、所属や専門を越えた新たな協働による知の創出を加速し、新たな価値を生み出していくことが可能となるとしています。また、オープンデータが進むことで、研究プロセスの透明化や研究成果の幅広い活用が図られるとしています。

近年、こうしたオープンサイエンスの潮流が世界的に急速な広がりを見せており、オープンイノベーションの重要な基盤としても注目されています。

さらに基本計画では、「国は、資金配分機関、大学等の研究機関、研究者等（など）の関係者と連携し、オープンサイエンスの推進体制を構築する」と述べられており、論文のみならず研究データもオープンにしていくことが、国の基本姿勢となることが明記されています。

1.2.2 不正対策

研究データ管理が注目される背景には、研究不正防止への期待もあります。

文部科学省の『研究活動における不正行為への対応等に関するガイドライン』は「研究機関が研究者に対して一定期間研究データを保存し、必要な場合に開示することを義務付ける必要がある」と明記しています。

文部科学省の依頼を受けた日本学術会議は、『科学研究における健全性の向上について』において、研究データを含む資料の保存期間を原則、論文発表後10年間とすると定めています。

1.2.3 助成機関のポリシー

研究データ管理が求められるもうひとつの背景として、研究資金の助成機関のポリシーを挙げることができます。データ管理計画、英語でData Management Plan(データマネジメントプラン）、略してDMP（ディーエムピー）と呼ばれる、研究データ管理に関する計画書の提出を、資金提供の条件にしている助成機関もあります。

英国では、表に示すように多くの研究助成機関がDMPの提出を求めています。欧州でも、欧州委員会による助成プログラム「Horizon2020（ホライズン・トゥエンティトゥエンティ）」が、パイロットプログラムに参加する研究について、DMPの提出を求めています。米国の一部の助成機関でも、DMPが導入されつつあります。日本では、科学技術振興機構＝JST（ジェーエスティー）が2016年から一部助成プロジェクトで義務化し、日本医療研究開発機構＝AMED（エーメド）も2018年度からデータ管理計画の提出を要求しています。

このような動きは、納税者への説明責任や、研究助成の投資効率の向上といった点からも、研究データ管理が求められていることを表しています。

1.3 研究データとは

以上のような背景の下、研究データ管理が注目を集めていますが、そもそも研究データとは何でしょうか。

1.3.1 研究データとは

研究データとは、研究の過程、あるいは研究の結果として収集・生成される情報です。

研究データは仮説の検証や、結論を導くために用いられ、また研究結果を確認するために使用されることもあります。

電子的なデータのみ研究データと定義する場合もあれば、紙ベースの情報も研究データに含める場合もあります。

研究データの定義は、研究機関や助成機関によって異なる場合があるので、必要に応じて確認をして下さい。

1.3.2 研究データを構成するもの

研究データには、ノートに記録されたものからデータベースに格納されたものなど、さまざまなものが含まれます。ここでは、その一例を挙げています。

研究データの内容は、試験や検査の結果からモデル、また方法論や操作手順などさまざまです。さらにデータの形式も文書、ノート、音声、写真、標本データファイルなどさまざまな形があります。

1.3.3 研究データの具体例

研究データは、目的に応じて多様な手段で取得されます。ここでは、研究データの具体例を学びます。

観測データは、天文観測や気象観測といった自然現象等の観測により取得されたデータです。

実験データは、DNA（ディーエヌエー）シーケンスやクロマトグラムなど、実験によって得られるデータです。

モデルやコードを用いてシミュレーションを行った結果得られるデータもあります。この場合、得られたデータに加え、シミュレーションに使用したモデルやコード、パラメータセットも管理すべきデータとして重要です。具体的には気象モデルや経済モデルなどがこれにあたります。

取得した生データを編集・加工したデータもあります。データマイニング、テキストマイニング、3D（スリーディー）データなどにより処理されたデータがこれにあたります。

研究の過程で参照するデータセットもあります。具体的には、遺伝子配列データバンク、結晶データベース、歴史画像アーカイブなどがあげられます。

1.3.4 研究データのライフサイクル

研究データとは何かを学んだところで、改めて本章の始めに示した、研究データのライフサイクルの図を見てみましょう。ここでは、各段階に必要なデータの取扱いの概要を学びます。

データの「生成」では、研究そのものの計画とともに、データ管理計画の策定やデータの共有方針が定められます。また、既存のデータの検索や、新たなデータの収集が行われます。

データの「加工」では、収集されたデータの、デジタル化、フォーマット変換、確認、検証、クリーニングなどが行われます。データに含まれる個人情報の匿名化作業が必要となる場合もあります。

データの「分析」では、収集・加工したデータから必要なデータを抽出したり、情報を読み取ります。ここで、論文など、研究のアウトプットの作成や、データの保存に向けた準備を行います。

データの「保存」では、データを適切なフォーマットやサイズに変更し、バックアップを作成し、メタデータや証拠文書の作成も行います。

データの「公開」では、データの配布、共有、アクセス制御、著作権処理等を行います。

公開されたデータは、追跡研究や新たな研究で「再利用」されます。

重要な点は、組織化、文書化、保存・共有が適切になされたデータは、プロジェクト終了後も長く再利用され続け、研究を推進し、イノベーションの機会を提供してくれるということです。

1.4 研究データ管理とは

研究データとそのライフサイクルの概要を理解したところで、研究データ管理について学びます。なお、本教材では電子的なデータの管理のみを対象とします。

1.4.1 研究データ管理とは

研究データ管理は英語でResearch Data Management（リサーチデータマネジメント）、略してRDM（アール・ディー・エム）と呼ばれます。

「研究データ管理」とは、一般的に、ある研究プロジェクトにおいて使用された、あるいは生成された情報を、どのように組織化、構造化、保管、管理していくのか、を指す言葉です。

例えば、研究データの取り扱い計画の策定や、研究中の日々の情報の取り扱い、研究後の長期的なデータの取り扱いなどが研究データ管理に含まれます。

1.4.2 研究データ管理のメリット

研究データの適切な管理にはさまざまなメリットがあります。

メリットの一つ目は、研究の効率化です。研究データを管理することで、自らのデータを発見し、利用することが容易になります。さらに、他の研究者により適切に管理されたデータを再利用することで、自身でデータを生成する労力を省くことができるだけでなく、さまざまな比較対象のデータを入手してより容易に分析を行うことができます。全体として研究にかかる時間や資源の節約につながります。加えて、データを安全に管理することでデータ消失のリスクが減ることも、研究効率につながります。

メリットの二つ目は、資金提供元の要求を充たせることです。適切なデータ管理計画を作成し、実施することで、助成を受けることができると同時に、資金提供元が資金の使いみちや助成の成果などの透明性を維持することにも寄与します。また、産業界や出版者等を含む業界の慣行に従ったデータ管理を行うことができます。

メリットの三つ目は、研究倫理の確保につながる点です。研究結果の根拠となるデータを適切に管理することは、研究の再現性の確立、研究データや研究記録の正確性、完全性、信頼性の保証などにつながります。

1.4.3 研究データ管理と大学

海外の大学の中には、研究データ管理に関する支援を図書館や他の関連組織が共同で行っているところもあります。

米国のバージニア大学、英国のエディンバラ大学、ケンブリッジ大学、オランダのアムステルダム大学の事例は、その一例です。

これらの大学では、図書館やIT（アイティー）部門、その他関連する組織が連携して、「データ管理計画の作成支援」や、「データの保存」支援、相談への対応や、ワークショップの開催など、多くのサービスを提供しています。

1.4.4 FAIR原則

研究データ管理を考える上で、大切な考え方として「FAIR（フェア）原則」があります。これは、研究データの再利用を推進する上で重要な事項を4つの柱に整理したもので、2016年に、研究者、図書館員、出版者、助成団体等から成るコミュニティであるFORCE11（フォースイレブン）によってまとめられました。

４つの柱は、Findable（ファインダブル）＝発見可能性、Accessible（アクセシブル）＝アクセス可能性、Interoperable（インターオペラブル）＝相互運用可能性、Re-usable（リユーザブル）＝再利用可能性であり、英語の頭文字をとってFAIR（フェア）原則と呼ばれています。

それぞれの柱には、具体的な行動指針が示されており、例えば「アクセス可能性」を実現するには、「データは、標準的な通信プロトコルを使用し、識別子によって検索できる」という指針が示されています。

FIAR原則は、機関によるデータ管理のポリシー策定や、研究データ管理サービスの設計・実施をする上で重要な原則です。

1.5 研究データ管理サービスとは

研究データ管理が求められる背景と、そのメリットを踏まえると、日本の研究機関においても「研究データ管理サービス」を整備していくことが喫緊の課題であることがよくわかります。

1.5.1 研究データ管理サービスとは

研究データ管理サービスとは、研究者によって生み出される研究データを適切に管理できるように、研究を支援するスタッフが協力して、必要なサービスを提供することです。

研究データ管理サービスにはさまざまなサービスが含まれます。ポリシーの策定から、研究前、研究中、研究後、そして日常的な支援まで、研究プロセスに沿ったさまざまな支援があります。また、これらのサービスは、スタッフが直接提供する人的支援とストレージなどの管理基盤の支援に分類することもできます。

研究データ管理サービスを検討する場合には、所属機関の状況に合わせて、これらを組み合わせ、担当部門が連携して設計するのが理想的です。別々の部門が連携することで、無駄や重複を防ぎ、研究者が混乱することなくサービスを利用することができます。

1.5.2 研究データ管理サービス：人的支援

はじめに、研究データ管理サービスを構成するサービスのうち、人的支援について学びます。

事前に策定された機関のポリシーを踏まえ、研究の開始前には、助成申請時に必要なデータ管理計画作成の支援サービスが求められます。データ管理計画作成ツールなども活用しながら、各助成機関の要求に応じた計画の作成を支援します。

研究中には、多くの情報源から研究に有用なデータを探す手助けを行ったり、データ分析の講習会なども開催しながら、収集したデータの分析を支援する必要があります。

研究後は、研究データの公開や他の研究者による「再利用」を促すための支援が求められます。メタデータの作成支援、公開したデータの利用条件を明確にするために、データにライセンスを付与することなどを支援します。欧米では、こうしたサービスを行うために、データライブラリアンと呼ばれる専門的な知識を持ったスタッフを配置しているところもあります。

研究段階に応じて提供するサービスに加えて、日常的に提供するサービスもあります。

研究支援者、研究者、大学院生向けの研修では、それぞれが研究データ管理やその支援にどのように取り組むべきかを学ぶためのセミナーや、オンライン学習、ワークショップの提供が考えられます。

研究データに関する相談をメールや電話で受けることができる窓口サービスも大切です。研究者から具体的な相談を直接聞き、必要に応じて他の部門と連携しながら解決していくサービスです。

研究データ管理の必要性や重要性は、現状では、まだ広く認知されていないことから、リーフレットやポスター、ウェブサイトなどで継続的に広報を行うことも重要です。研究者コミュニティにアプローチし、研究データ管理の重要性についての理解を促し、誤解や疑問を解消するアドボカシー活動も時間をかけて行う必要があります。

1.5.3 研究データ管理サービス：管理基盤

研究データ管理サービスには、ストレージなどの管理基盤の提供も含まれます。

オンラインストレージの提供は、研究中に「生成」されたデータのバージョン管理や、チーム内でのファイル共有に役立ちます。最近では、クラウドストレージのようなデータの保存領域を研究機関が提供することも多くなってきました。情報セキュリティの観点からも、機関がオンラインストレージを提供することはとても重要です。

研究後のデータを「保存」する段階では、長期保存を目的とした、データを頻繁に出し入れしないストレージの提供が考えられます。研究成果の根拠であるデータを長期保存することは、研究公正の側面からも非常に重要です。しかし、すべてのデータを長期間保存し続けるということは現実的ではないため、一定の条件でデータを選別して残す必要があります。

さらに、データを「公開」する際には、データリポジトリの提供による支援が考えられます。データの公開にあたっては、「人的支援」で述べたメタデータやライセンスの付与に加え、データなどのデジタル・オブジェクトの識別子であるDOI（ディーオーアイ）の付与や、研究者個人を識別できるIDであるORCID（オーキッド）の登録も大切です。こうした識別子の付与により、データの引用のモニタリングやカウントが容易になり、研究者のインセンティブにつながります。DOI（ディーオーアイ）やORCID（オーキッド）は機関単位で管理して付与することも可能ですので、他のサービスと合わせて検討するとよいでしょう。

第2章　研究データ管理サービスの設計

2.1 第2章の概要

第2章では、研究データ管理サービスの設計について、サービスの全体像、サービスの設計、研究データ管理に取り組む組織体制づくり、サービス設計のための事前調査、サービスの試行・評価、海外大学の事例について学んでいきます。

まず、「研究データ管理サービスの全体像」では、研究前・研究中・研究後そして日常的な研究データ管理支援サービスの概要を学びます。それぞれの段階で提供できるサービスの詳細は3章以降で学んでいきますが、ここでは各章で示される具体的な支援項目を概観します。

次に、「研究データ管理サービスの設計」では、英国のデジタル・キュレーション・センターのガイドブックを参考に、サービス設計の3つのプロセスを概観します。サービス設計では、始めに研究データ管理にとりくむ組織体制づくりをおこないます。そして、サービスを設計するための事前調査をおこないます。その結果を踏まえて、設計したサービスの試行と評価を繰り返します。サービス設計では、これらのプロセスすべてにおいて全学レベルで部門横断的に取り組むことが重要です。

「研究データ管理に取り組む組織体制づくり」では、ステークホルダーをもれなく想定し、全学レベルで研究データ管理に取り組む体制を構築する方法を学びます。

「サービス設計のための事前調査」では、必要なサービスを特定するための調査方法について学びます。研究者の取り扱うデータの実状を定量的に把握する手法としてはアンケート調査が有効です。また、研究者の研究データ管理に対する予備知識や意識を把握するためには、インタビューによる調査が効果的です。

「サービスの試行、評価」では、設計したサービスの試行と評価、改善をくりかえしながら、自機関に合ったサービスを構築するプロセスについて学びます。試行、評価、実施のプロセスは、予算や人材確保、サービスの定着にかかる時間を踏まえて、長期的計画で実施していくことが大切です。

最後に、「海外大学の事例」では、エディンバラ大学とワーゲニンゲン大学の研究データ管理サービスの概要を学びます。両大学では、研究データ管理やデータ管理計画作成に関するポリシーが制定され、また国レベルでも方針が定められています。二つの大学はともに図書館、IT部門、その他複数の部門が連携した組織体制を構築し、部門が連携して研究データのライフサイクルに沿った多様なサービスを提供しています。

それでは、次の動画からそれぞれについて詳しく学んでいきましょう。

2.2 研究データ管理サービスの全体像

始めに、研究データ管理サービスの全体像を学びます。

2.2.1 研究データ管理サービスの全体像

1章でも説明したとおり、研究データ管理サービスとは、研究者によって生み出される研究データを適切に管理できるように、研究を支援するスタッフが協力して、必要なサービスを提供することです。

研究データ管理サービスにはさまざまなサービスが含まれます。ポリシーの策定から、研究前、研究中、研究後、そして日常的な支援まで、研究データのライフサイクルに沿ったさまざまな支援があります。また、これらのサービスは、ストレージなどの管理基盤の支援とスタッフが直接提供する人的支援に分類することもできます。

研究データ管理サービスを設計する場合には、所属機関の状況に合わせて、これらを組み合わせ、担当部門が連携して設計するのが理想的です。別々の部門が連携することで、無駄や重複を防ぎ、研究者は混乱することなくサービスを利用できます。

これから、研究前、研究中、研究後、日常的な支援サービスについて、それぞれの概要を学んでいきます。

2.2.2 研究前の支援

まず、研究前の研究データ管理支援サービスの概要を学びます。

研究データ管理支援の前提となるのが、ポリシーです。研究者は所属する研究機関、研究助成機関、学会や出版社のポリシーをそれぞれ守る必要があります。研究データ管理支援者は、これらのポリシーの要求内容等も踏まえながら研究データ管理に関する支援を提供していくことが求められます。

ポリシーで求められるデータの適切な管理や保存、必要に応じた開示に対応するためには、研究を開始する前の段階で、ポリシーに基づいた研究データ管理を実施するための「計画」を策定します。これが、データ管理計画、Data Management Plan(データマネジメントプラン）、略してDMP（ディーエムピー）と呼ばれるものです。研究助成機関が、助成申請時にデータ管理計画の提出を義務化する動きも見られます。各機関が求めるデータ管理計画の項目はさまざまですので、その要求内容に沿った計画を策定できるよう、研究データ管理支援者は適切な情報提供や作成支援を行うことが求められます。

データ管理計画の作成支援サービスの１つとして、データ管理計画作成ツールの提供があげられます。このツールは、テンプレート等を用いて計画書を容易に作成したり、既に作成された計画書を参考にできるオンラインツールです。データ管理計画作成ツールを利用すると、作成の負担を抑えることができますので、研究データ管理支援者は、このようなツールについての情報提供をおこない、研究者が必要に応じて研究データの特性に合ったものを選べるように、サポートをするとよいでしょう。

研究前の支援のより具体的な内容は、3章で学びます。

2.2.3 研究中の支援

次に、研究中の研究データ管理支援サービスについて学びます。

論文の根拠となるデータは論文発表後10年間の保存が原則とされています。研究機関は、組織として研究データを安全に管理・保存するためのストレージ領域を提供し、研究データ管理支援者はその活用を促進、支援するとよいでしょう。

研究中に、自分が作成したデータ以外の既存のデータを探して、再利用するための検索支援も大切です。研究データ管理支援者は、リポジトリやデータジャーナル、論文から探す方法や、利用申請をして非公開のデータを探す方法、関連する研究プロジェクトや研究者を探しデータを見つける方法などを理解しておくことが重要です。

また、研究中の主たる研究活動であり、研究成果に大きく関わる「データ分析・可視化」を支援するため、研究者を対象に汎用的な統計、解析ツールの利用方法に加え、分野特化型のツールについてもコースを提供するとよいでしょう。

さらに、論文発表に向けて、事前にデータの共有・公開について、関係者との合意や、関連機関のポリシーの確認を促すことも大切です。また、研究で参照したデータを論文で引用することを勧め、具体的な引用の記載方法についても情報提供を行います。

研究を開始する前に作成したデータ管理計画は、研究実施中も必要に応じて振り返り、更新していくことが重要です。データ管理計画作成ツールでは、一度作った計画を編集することもできます。

研究中の支援のより具体的な内容は、4章で学びます。

2.2.4　研究後の支援

続いて、研究後の研究データ管理支援サービスについて学びます。

研究中に作成されたデータを公開する場合は、データリポジトリを利用します。データリポジトリには、特定分野のデータを集録する分野別リポジトリと、研究機関が所属する研究者に対し提供する機関リポジトリ、複数分野を取り扱う汎用リポジトリがあります。関連機関のポリシーやガイドライン、研究コミュニティへのインパクトなどを踏まえて、研究者がデータを公開・保管するリポジトリを適切に選択できるように支援しましょう。

研究データをデータリポジトリに登録して保管・公開する際には、メタデータを付与します。メタデータを付与することによって、登録されたデータセットの検索や引用が可能になります。分野別のリポジトリにおいては、分野特有の詳細なメタデータが求められることも多いため、研究データ管理支援者は、メタデータ作成のマニュアルやエディターの等のツールの情報提供を行います。機関リポジトリへのデータの登録の際は、大学図書館員などが直接データの入力をしたり、専門的なアドバイスを行うとよいでしょう。

データの公開に伴い、コンテンツの利用許諾の条件＝ライセンスをメタデータ中に明記することで、データの再利用を促すことが可能となります。関連するポリシーや、対象となるデータの性質、データ保有者の意向等を踏まえながら、適切なライセンスについて情報提供を行うことも大切な支援活動です。

研究後の支援の詳細は、5章で学びます。

2.2.5 日常的な支援

ここでは、日常的な研究データ管理支援サービスについて学びます。

まずは、研究データ管理に関する研修の実施です。研究データ管理支援者には、オンライン講座や「シナリオ」を用いた実践的な研修、研究者にはポリシーや研究活動を踏まえた研修、大学院生には情報リテラシー教育、研究倫理教育の観点での研修、という風に、対象者に合わせた研修を企画するとよいでしょう。

また、研究者が研究データ管理を行う上で最初に参照する場所として、研究データ管理＝RDM（アールディーエム）のポータルサイトを提供することも大切です。ポータルサイト上では、研究データ管理の基本情報や提供されるサービス内容、担当者の連絡先といった情報をまとめ、研究データ管理と支援サービスへの理解を図ります。

研究データに関する相談窓口は、研究データ管理に関する問合せをメールや対面などで、ワンストップで受け付ける窓口です。研究者や大学院生がどこに相談してよいか迷うことのないように、関連部署が連携して窓口を運営しましょう。

さらに、研究データ管理の簡潔な広報資料の作成や、ミニワークショップの開催、学部・学科の教授会などでのブリーフィングを通して、大学の関係部署のスタッフを含む、ステークホルダーの理解を促進する、アドボカシー活動も継続的に行っていくとよいでしょう。

日常的な支援の詳細は、6章で学びます。

2.3 研究データ管理サービスの設計

ここからは、研究データ管理サービスの設計プロセスを学びます。サービス自体の設計プロセスに加えて、サービスの前提となる戦略やポリシーの設計プロセスも学びます。

2.3.1 研究データ管理戦略の策定

効果的な研究データ管理を支援するには、包括的な戦略を立て、一貫したサービスの開発を行うことが大切です。

戦略は、現状把握、将来像の明確化、そして将来像に移行するための作業計画の作成、という3つのステップで策定します。

現状と将来像を明確にした後は、その間にどのようなギャップがあるのかを分析します。例えば、研究データ管理に影響を及ぼす内外の要因は何か、守るべきポリシーや、組織のあるべき姿と研究データ管理活動のメリット等を踏まえて、ギャップを分析し、把握するとよいでしょう。

次に、将来像に移行していくための作業計画を作成します。作業計画の内容は、具体的な例を参考にするとよいでしょう。エディンバラ大学のロードマップでは、データ管理計画、アクティブデータ基盤、データ・スチュワードシップ、およびデータ管理支援の4つの主要エリアの下に作業計画がまとめられています。作業エリアごとにいくつかの目標とそれに伴う具体的な作業、成果物および目標期日がリストされています 。

2.3.2 研究データ管理ポリシーの策定

次に、研究データ管理ポリシーの重要性と策定のプロセスについて学びます。

ポリシーは、研究データ管理について一貫した展望を示すことで組織内の利害を調整する役割があります。また、より具体的に研究データ管理サービス導入のあり方を示し、サービスの構築を後押しします。

ポリシーを策定する際には、さまざまなステークホルダーの役割や課題、ニーズを踏まえながら草案を策定します。その後、小規模なグループで草案を検討します。グループの構成は、例えば、研究者や、研究機関の執行部、主なサービススタッフなどが考えられます。

小規模なグループで検討し、草案を修正するプロセスを何度か経た後、研究機関の運営組織による批准を受ける必要があります。批准には時間がかかる場合もあるため、ポリシーは簡潔で高いレベルの原則に焦点を当てるとよいでしょう。

ポリシー策定では、広範なコンサルテーションを通して、さまざまなステークホルダーからのフィードバックと、研究データ管理への関与を引き出すことが大切です。つまり、ポリシー作成のプロセス自体が、効果的なアウトリーチになっている、ということを踏まえておくとよいでしょう。

ポリシーの内容の詳細や事例は3章で説明します。

2.3.3 サービスの設計

次のステップでは、戦略やポリシーに基づいて作成された計画に従って研究データ管理サービスを設計します。

英国のデジタル・キュレーション・センターは、研究データ管理サービスの設計をすすめるにあたり、主に3つのプロセスを推奨しています。

まずは、研究データ管理に取り組む組織体制づくりです。研究データ管理を主幹する委員会等を設置したり、実際に業務を実施するメンバーで研究データ管理チームを組成し、全学的な体制を作ります。

次に、自機関に見合ったサービス設計を行うために、研究者や学部単位で研究データ管理に対してどのようなニーズがあるのかを調査します。

求められるサービスを把握できたところで、実施主体を決定し、実際のサービスをまずは試行するところから始めます。そして、試行結果を検証し、サービスの本実施につなげます。

ここで説明したプロセスは手順の一つの例ですが、研究機関によって、サービス設計を進める主体、順序はさまざまです。海外の研究機関の事例なども参考にしながら、全学レベルかつ部門横断的に取り組むことが、研究データ管理サービス設計の成功のカギといえます。

ここからは、それぞれのプロセスについて詳しく学んでいきます。

2.4 研究データ管理に取り組む組織体制づくり

はじめに、研究データ管理支援サービスの設計プロセスの一つ目である、組織体制づくりについて学びます。

2.4.1 ステークホルダーの抽出と組織体制づくり

研究データ管理サービスを実際に運用する主体は、機関の中のさまざまな既存の組織に所属する研究支援スタッフです。研究者以外のステークホルダーをもれなく想定し、それぞれが既存の役割に応じて全学レベルの研究データ管理において何ができるかを考え、責任をもって取り組めるような体制を整備しなければなりません。

こういった組織横断的な体制づくりは長期的な取り組みとなることが予想されます。なお、最初からすべてのステークホルダーに参加を呼びかけるのではなく、研究データ管理に関心のある研究者コミュニティに参加してもらうなど、さまざまな進め方があるでしょう。

その後、研究担当理事や副学長のリーダーシップの下、研究データ管理を主幹する委員会等を設定します。また、業務を実施する責任・技能を備えたメンバーで構成された研究データ管理チームを組成します。体制づくりの際は、ボトムアップとトップダウンの両方のやり方が考えられますが、組織横断的な体制を組むことがのぞましいため、最終的には執行部レベルのサポートがあることが望まれます。

2.5 サービス設計のための事前調査

次に、研究データ管理支援サービスの設計プロセスの二つ目である、サービス設計のための調査について学びます。

2.5.1 事前調査の必要性 Data Asset Frameworkとは

サービスの設計にあたっては、まずは必要なサービスを特定するために、調査を実施して、研究者の研究データ管理の実態やニーズを把握します。

調査の具体的な方法については、英国で使われたData Asset Framework（データアセットフレームワーク）、DAF（ディーエーエフ）と呼ばれる、研究データを資産と捉えた一連の調査の枠組みを参考にするとよいでしょう。

DAF（ディーエーエフ）の目的の一つ目は、研究者の取り扱うデータの実状を定量的に把握し、研究データ管理基盤の整備の参考にすることです。このためには、アンケートを用いた調査が役立ちます。

目的の二つ目は、研究者の研究データ管理に対する予備知識や意識を把握し、必要な研修や、アドボカシー活動の計画に活かすことです。このためには、インタビューによる調査が効果的です。

次に、具体的な調査内容について紹介します。

2.5.2 アンケート調査の例

これは、アンケート調査の設問の一例です。

研究者に対して、研究プロジェクトで生み出されるデータの種類や容量、保存方法、研究データ共有に関する意識や、サービスへの関心を尋ねます。

研究者には、研究の合間の貴重な時間を費やしてアンケートに答えてもらうことになります。このため、研究機関として、アンケートへの協力を求める際には、調査の意義や目的等についての充分な説明や、回答することへの動機付けも工夫してください。

2.5.3 インタビュー調査の例

これはインタビュー調査の設問の一例です。

質問する項目のうち、基本的な項目はこのようにリスト化してからインタビューに臨むとよいでしょう。アンケート調査と同じような項目もありますが、選択回答などでは答えにくい、長期的な管理への不安や、機関への期待、研究データ管理への関心や意識を柔軟に尋ねることを心がけます。

インタビュー調査は、さまざまな分野の研究者の生の声を聞くと同時に、研究データ管理への関心を高めるよい機会ともなりますので、調査対象の人数は少なくてもよいので、実施することが望まれます。

2.6 サービスの試行、評価

さらに、研究データ管理支援サービスの三つ目のプロセスである、サービスの試行と評価について学びます。

2.6.1 試行、評価、実施のポイント

組織体制を構築し、研究者のニーズや現状を事前調査で把握したら、自機関の状況に合わせてサービスを設計しましょう。

そして、できるところから試行を始めましょう。試行にあたり、新たな資金が必要になる場合は、機関の内外から資金を獲得することも必要になるでしょう。また、ハードウェアに関するサービスを試行する場合には、機関のシステム変更などのタイミングを確認することも必要となります。最初はスモールスタートで小規模に始めてみることがポイントです。

試行後は、提供するサービスが機関の現状やニーズに合致しているか、評価をおこなうことが大切です。評価を基に、さらなる改善を図り、再び設計、試行、評価、改善を行う場合もあるかもしれません。このPDCA（ピーディーシーエー）サイクルをくりかえしながら、自分の機関に合ったサービスを構築していきます。

平行して、研究者に対する広報やアドボカシー活動を通じて関心・理解度を高め、サービス内容の評価、改善を繰り返します。

これらの流れを経て、サービス提供の本実施となります。

このように、サービス設計においては、試行、評価、改善をくりかえしながら、自分の機関に合ったサービスを構築し、組織内の意識を高めていくことが大切です。また、サービスの定着には時間がかかります。組織内で理解が得られ、予算や人材の確保など準備が整ったところから順に実施し、長期的計画で進めていく必要があるという点も重要です。

長期的計画の例として、先進的な研究データ管理サービスを提供するエディンバラ大学では、2008年に研究データに関する調査を行い、試行等を経て包括的な研究データ管理プログラムが導入されたのは2012年です。現在も試行を重ね、システムの改善など長期的に取り組んでいます。

2.7 海外大学の事例

2章の最後に、参考として海外の大学における研究データ管理支援サービスの事例について、特にポリシーと組織体制、提供するサービスに着目して学んでいきます。

2.7.1 事例：エディンバラ大学（英国）

最初の事例は、英国のエディンバラ大学で提供されている研究データ管理支援サービスです。ここでは特に大学のポリシー制定と組織体制、提供されるサービスの概要を紹介します。

エディンバラ大学は、2011年に大学全体に向けたデータ管理ポリシーを作成しました。ポリシーでは、「研究データは、そのライフサイクルを通して最高の水準で管理される」ことを明言しています。

大学に加えて、国の方針においても、研究データ管理は重視されています。国の研究助成機関は、同じく2011年に、助成を受けた研究については研究データを公開する旨の原則を定めています。加えて、英国内の大学の研究を評価する国の枠組みにおいても、研究データが研究成果の対象として含まれることが新たに定められています。

このような国の方針と、大学のポリシーに従って、エディンバラ大学では、機関全体で包括的な研究データ管理支援サービスを提供する体制が構築されました。研究データ管理支援サービスを所管するInformation Service Group（インフォメーションサービスグループ）は、図書館、IT部門を含むさまざまな部署が含まれています。

部署間で連携して提供されるサービスは多岐にわたります。研修の実施や、データ管理基盤の提供、データ管理計画の作成、個別の相談対応など、研究データのライフサイクルに合致した包括的なサービス提供が行われています。

2.7.2 事例：ワーゲニンゲン大学（オランダ）

二つ目の事例として、オランダのワーゲニンゲン大学で提供される研究データ管理支援サービスについて、特に大学や国の方針、組織体制、提供されるサービスの概要を学びます。

ワーゲニンゲン大学では、2017年から博士後期課程の大学院生及び研究チームに対し、データ管理計画の作成が義務化されています。

また、オランダにおける研究公正の要求として、最低10年のデータセットの保存義務などが定められています。

さらに、ワーゲニンゲン大学が参加する国内大学の連携組織などでは、研究データ管理ポリシーやサービスインフラの作成計画を大学同士で連携するイニシアチブが示されています。

このような背景のもと、ワーゲニンゲン大学では図書館、IT（アイティー）部門や法務部門などが連携してデータ管理サポートグループを組織しています。

グループが提供するサービスは、研修やデータ管理計画作成といったものから、分野に特化した専門的なものまで、多岐にわたります。

第3章　研究前の支援

3.1 第3章の概要

第3章では、研究前の支援として、ポリシーや、データ管理計画、データ管理計画作成支援について学びます。

まず、「ポリシー」については、研究者が守るべきポリシーや、ポリシー策定が求められる背景、またさまざまな関連機関のポリシーの実例などを学びます。

研究者が守るべきポリシーは主に3種類あります。一つ目は研究者が所属する研究機関のポリシー、二つ目は研究や研究機関に資金を助成する政府や助成団体といった研究助成機関のポリシー、三つ目は研究成果を発表する媒体を管理する学会や出版者のポリシーです。

これらの関係機関がポリシーを定める背景には主に二つの流れがあります。一つ目は、データの利活用を促進するためのオープサイエンスの流れ、二つ目は研究活動の正当性を検証するためにデータの保存等を求める研究公正の流れです。日本では、特に研究公正の流れを受けたポリシー策定が先行しています。

次に、「データ管理計画とは」では、データ管理計画の概要やその意義、義務化の動向や、データ管理計画の作成支援について学びます。

データ管理計画を作成することにより、研究の初期段階からデータ管理の検討をおこない、準備ができます。計画と入念な準備に基づき、研究中の適切なデータ管理が可能となり、研究の質と効率性の向上につながります。このため、ポリシー等で要求されているか否かを問わず、データ管理計画の策定にとりくむ必要があるでしょう。

そして、「データ管理計画作成支援」では、計画に含まれる具体的な内容について学びます。

データ管理計画で記載が求められる事項は、機関等によってさまざまです。英国のデジタル・キュレーション・センターは、データ管理計画を作成する際に求められる可能性がある情報を「管理上のデータ」、「データ収集」、「文書化とメタデータの付与」、「倫理・法律のコンプライアンス」、「保管とバックアップ」、「選定と保存」、「データ共有」、「責任とリソース」の8つの項目で整理しています。

データ管理計画は、研究の初期段階で作成した後、計画を実行しながら必要に応じて修正をおこないます。そして研究終了後は計画の達成状況について検証をおこないます。「研究の初期段階で作成して終わり」ではなく、このようなライフサイクルに沿って更新・活用していくことで、データ管理計画に基づく適切な管理が可能となります。

それでは、次の動画からそれぞれについて詳しく学んでいきましょう。

3.2 ポリシー

始めに、研究データ管理を行う際の指針となる「ポリシー」について学びます。研究データ管理支援者は、研究者が守らなければならないポリシーの種類や内容を理解し、研究前の段階から適切な情報提供を行うとよいでしょう。

3.2.1 ポリシーとは

ポリシーとは、政策、方針、規定、などの意味を持つ用語です。

「セキュリティポリシー」「プライバシーポリシー」のように、何らかの物事について、組織としての方針を文書にまとめて公開したものを「ポリシー」と言います。

ネットワーク管理の分野では、管理者が運用方針に基づいて記述した、設定ファイルや運用ルールなどのことをポリシーということもあります。

3.2.2 研究者が遵守を求められる3つのポリシー

研究者は、関連する複数の機関のポリシーについて、それぞれ遵守することが求められます。

守るべきポリシーの一つ目は、大学や研究所などの「研究機関」のポリシーです。

二つ目は、研究機関に対して資金を配分する、政府や助成団体などの「研究助成機関」のポリシーです。

三つ目は、研究成果を発表する学術雑誌などの媒体を管理する学会や出版者のポリシーです。

3.2.3 研究データに関するポリシーの背景

こういったポリシーが作成され、遵守が求められる背景には、大きく2つの流れがあります。

一つ目がオープンサイエンスの流れです。

2015年3月に、報告書『我が国におけるオープンサイエンス推進のあり方について」が内閣府から示されました。2016年7月には日本学術会議からも提言がなされており、異分野での利活用を促進するため、積極的に研究データをオープン化することが求められています。

二つ目に、研究公正の流れがあります。

2014年8月に文部科学省により示されたガイドラインでは、研究活動における不正行為が疑われた際に検証できるようにするため、研究データを原則として10年間保存・管理することが求められています。

オープンサイエンスの流れは積極的なオープン化を促進するものである一方、研究公正の流れは、検証が目的であり利活用は想定されていません。この二つの異なる背景を受け、関連機関でポリシーが作成されていますが、日本国内では特に研究公正の流れを受けたポリシー策定が先行しています。

3.2.4 研究公正に関するポリシーの指針

日本学術会議は、2015年3月に、特に研究公正の観点から守るべき原則を示しています。

研究公正に関するポリシーの対象には、論文の根拠資料となる研究データで、実験ノート、数値データ、画像などが含まれます。

また、一定期間の保存が義務付けられており、日本学術会議では保存期間を「原則として論文発表から10年間」としています。

必要に応じて研究データを開示することや公開することも推奨されています。

また、データを作成し、管理する教職員が異動したり退職したりした後も、機関として研究データを保存し、データへのアクセスを確保することも求められています。

3.2.5 研究機関のポリシーの例（研究公正）

ここでは主に研究公正に対応するポリシーの例として、２つの研究機関のポリシーについて学びます。

京都大学は、2015年7月に「京都大学における公正な研究活動の推進等に関する規程第７条第２項の研究データの保存、 開示等について定める件」を発表しました。その中で、「教職員等は、当該論文等に疑義が呈された場合、その他必要に応じて保存する研究データを開示しなければならない」としています。

高エネルギー加速器研究機構は、2015年3月に「公正な研究活動の推進に関する規程」を発表しました。第4条で、「研究者は、自らの研究が公正に実施されたことを示し、必要に応じて第三者による研究成果の検証を可能とするため、研究活動によって得られた研究データ等を一定期間保存し、必要に応じて開示できるようにしなければならない」と明記しています。

3.2.6 研究機関のポリシーの例（オープンサイエンス）

次に、オープンサイエンスに対応する研究機関のポリシーの例について学びます。

海洋研究開発機構では、2007年に策定したデータポリシーの中で、「科学的・教育的利用のため、国内外の研究機関及び研究者などがデータ・サンプルを利用できるように、適切に管理・保管するとともに迅速で円滑に提供するよう努める」としています。また、「国民に利益が還元されるよう産業利用等を促進する」ことも明言しています。

国立環境研究所も同様に、2017年にデータポリシーを策定し、「国立環境研究所が研究活動を通じて取得・作成したデータのうち、研究成果として公開したデータの他、公益性や社会的ニーズが高く、公開することが適当であると判断したデータを公開対象とする」としています。

3.2.7 研究助成機関のポリシーの例

研究助成機関のポリシー例として、科学技術振興機構＝JST（ジェーエスティー）の基本方針につい学びます。

JSTは、2017年4月に「オープンサイエンス促進に向けた研究成果の取扱いに関するJSTの基本方針」を発表しました。その中で、JSTが研究資金を配分し実施する研究プロジェクト等において、研究データの取り扱いについて重要な点を3つ挙げています。

一つ目は、データマネジメントプラン、データ管理計画の作成です。

JSTが研究資金を配分し実施する研究プロジェクト等においては、研究データの取扱いを定めたデータマネジメントプランを作成し、JSTに提出する、としています。なお、データマネジメントプラン、すなわちデータ管理計画の概要は、次のパート以降で学びます。

二つ目は、データの保管、管理についてです。

研究代表者等は、データマネジメントプランに基づいて、研究データを適切に保管・管理する、としています。

三つ目は、研究データの公開についてです。

研究データのうち、エビデンスデータは公開することを推奨しています。また、同時に、それ以外の研究データについても公開することを期待する、としています。

3.2.8 学会・出版者のポリシーの例

近年、学会や出版者も研究データについてのポリシーを定めつつあります。

例えば、Nature（ネーチャー）とNature関連誌12誌は、2016年9月から、受理された論文の基礎データについて、著者以外の者による利用可能性を明示した「データ利用可能性ステートメント」を論文中に記載することを義務化しました。

また、著者は、デジタル・オブジェクト識別子（DOI（ディーオーアイ））が付与されたデータセットを引用することが推奨されています。

3.2.9 ポリシーを策定するために

ここまで見たように、多くの研究機関や関連機関ではデータ管理に係るポリシーが定められています。

2018年2月に、内閣府の「国際的動向を踏まえたオープンサイエンスの推進に関する検討会」に、「研究機関の公的資金による研究データの管理・利活用ポリシー策定に関するコアガイドライン（案）」が提示されました。

ガイドライン案では、ポリシーに含むべき要素として、ポリシーの目的や、研究データの定義の他、データ管理を実施するにあたり、踏まえるべき重要なポイントが挙げられています。

データに関するポリシーを策定する際には、このガイドライン案を参考にするとよいでしょう。

3.2.10 ポリシーを遵守するために

これまで見てきたように、研究者はさまざまなデータ管理に関するポリシーを守る必要があります。

そのためには、研究を開始する前の段階で、ポリシーに基づいた研究データ管理を実施するための「計画」を策定することが重要です。

データ管理の計画には、責任者は誰か、保存先はどこが適しているか、公開が必要か、必要でないか、公開する場合は適切な公開時期はいつか、どのデータを保存・公開するか、といった内容が織り込まれている必要があります。詳しくは次のパートで説明します。

3.3 データ管理計画とは

ここでは、「データ管理計画」とは何か、そして計画を策定することにどのような意義があるのか、ということについて学びます。

3.3.1 データ管理計画（DMP）とは

文部科学省が2016年に発表した「学術情報のオープン化の推進について」では、データ管理計画とは「研究プロジェクト等における研究データの取り扱いを定めるものであり、具体的にはデータの種類、フォーマット、アクセス及び共有のための方針、研究成果の保管に関する計画などについて記載されるもの」と定義されています。

つまり、データ管理計画は、研究の実施段階から終了後に至るまでの期間において、研究データがどのように生成、管理、共有、保存されるのかを文書化したものといえます。この計画に従い、研究中はデータの適切な保管・管理が行われ、研究後はデータの共有や公開が進められます。

3.3.2 データ管理計画（DMP）策定の意義

データ管理計画を策定することにはさまざまな意義があります。

研究の初期段階から、データ管理に関する事柄を検討することで、入念な準備を行うことができます。

計画と準備にもとづいて、研究全体を通して、適切なデータ管理を行うことができます。適切なデータ管理は、研究の公正性や再現性を保証し、また自分の作成した研究データの信頼性も保証します。そして、データのセキュリティ向上にもつながります。

適切かつ計画的なデータ管理を行うことで、データ収集の重複などを回避でき、結果的に、研究の質と効率性を高めることにつながります。

要求されているか否かに関わらず、データ管理計画はより良い研究を行うために策定することが望ましいといえます。

3.3.3 データ管理計画（DMP）策定の義務化

研究データ管理が求められる背景には、先ほど触れたように、研究資金の助成機関のポリシーによる要求もあります。近年、研究助成機関が助成の要件としてデータ管理計画＝DMP（ディーエムピー）の策定を求める動きが国際的に広がっています。

英国では、表に示すように多くの研究助成機関がデータ管理計画の提出を求めています。

欧州でも、欧州委員会による助成プログラム「Horizon2020（ホライズン・トゥエンティトゥエンティ）」に参加する研究については、データ管理計画の提出が求められています。

米国の一部の助成機関でも、提出が義務化されつつあります。

日本では、科学技術振興機構＝JST（ジェーエスティー）が2016年から一部助成プロジェクトで計画を義務化し、日本医療研究開発機構＝AMED（エーメド）も2018年度からデータ管理計画の提出を義務化しています。

3.3.4 データ管理計画（DMP）策定の支援

助成機関等による義務化の方針を受け、海外の大学では、データ管理計画策定の支援に取り組む例が見られます。

イリノイ大学、エディンバラ大学、ケンブリッジ大学、アムステルダム大学などの事例は、その一例です。

それぞれの大学では、データ管理計画の作成ツールの紹介や、実際の計画のサンプルなども提供しながら、研究者のデータ管理計画策定を支援しています。

3.4 データ管理計画作成支援

ここからは、計画に含まれる具体的な内容について学びます。データ管理計画の形式や求められる内容はさまざまですが、主な内容を理解しておくことは、データ管理計画作成を支援するにあたり重要なことです。

3.4.1 データ管理計画の例

初めに、記載すべき具体的な項目の例を学びます。

JST（ジェーエスティー）は2017年4月に発表したガイドラインにおいて、データ管理計画 に明記すべき必須項目を挙げています。

項目は、

・管理対象となる研究データの保存・管理方針、

・研究データの公開・非公開に係る方針、

・公開可能な研究データの提供方法・体制、

・公開研究データの想定利用用途、

・公開研究データの利活用促進に向けた取組み、

の5つとなっています。

3.4.2 データ管理計画に含める内容

実際にデータ管理計画を作成する際には、データ管理について具体的に何をどのように計画する必要があるのでしょうか。

英国のDigital Curation Centre （デジタル・キュレーション・センター）では、計画を策定する上で求められる可能性がある情報を、管理上のデータ、データ収集、文書化とメタデータの付与、倫理・法律のコンプライアンス、保管とバックアップ、選定と保存、データ共有、責任とリソースの8つの項目で整理しています。

各項目について、具体的にどのような情報を含む必要があるのか、学んでいきましょう。

3.4.3 ①管理上のデータ

まず管理上のデータとして把握しておく必要のある情報は、

・助成機関や所属機関が定めた関連ID（アイディー）

・助成機関・助成金番号

・プロジェクト名称・内容・連絡先

・データ収集・生成の目的

・個人識別子、もしくは主要研究者名とそのPI（ピーアイ）

・データ管理計画書の作成日・最終更新日

などです。

また、関連するポリシーとして、

・何か既存の手続きに則っているか

・所属部局等はデータ管理ガイドラインを有しているか

・所属機関はデータの保護やセキュリティに関するポリシーを有しているか

・所属機関や助成機関は研究データ管理ポリシーを有しているか

・何らかの公式規格を採用したか

といった問題についても考えておく必要があるでしょう。

3.4.4 ②データ収集

データ収集については、

どのようなデータを収集・生成するのか、どのようにデータを収集・生成するか、という2つの点から計画を策定する必要があります。

「どのようなデータを収集・生成するのか」について考慮すべきポイントとしては、

・データの種別やフォーマット、容量など

・選択したフォーマットやソフトウェアはデータの共有・長期アクセスに適しているか

・再利用できる既存のデータはあるか

といったことが挙げられます。

また、「どのようにデータを収集・生成するのか」について考慮すべきポイントは、

・どの標準や方法論を用いるのか

・フォルダやファイルをどのように構造化するのか

・バージョン管理をどのように行うのか

といったことです。

3.4.5 ③文書化とメタデータの付与

文書化とメタデータについては

・データが将来にわたって可読性を有し、解釈されうるためには、データにどのような情報が必要とされるか

・どのように文書化し、メタデータを作成するか

・どのメタデータ標準を用いるか。また、その理由は何か

という点について考えておく必要があります。

3.4.6 ④倫理・法律のコンプライアンス

倫理的問題に関しては、

・データの保存や共有についての同意取得の有無や、

・もし必要な場合、参加者の個人識別性をどのように保護するのか（この場合、匿名化などの方法が考えられます）、

・センシティブデータが確実に安全に保管・移動されるようにするため、どのような対処をするか

といった点を踏まえておくべきです。

また、著作権や知的財産権に関する問題については、

・データの所有者は誰か

・再利用のためどのようなライセンスを付与するか

・第三者による再利用について何らかの制限を定めるか

といったことを明確にしておく必要があります。

3.4.7 ⑤保管とバックアップ

研究期間中は、データをどのように保管・バックアップするのか、ということも研究を進める上で重要となります。このため、計画段階から

・十分なストレージを有しているか

・付加サービス利用に必要な料金を含める必要があるか

・データをどのようにバックアップするか

・障害発生時、データはどのように復元されるか

などのポイントについて考えておく必要があります。

また、アクセス制限やセキュリティ確保をどのように管理するのか、といったこともあらかじめ考えておくべきです。

3.4.8 ⑥選定と保存

どのデータを保持・共有・保存すべきか、という点を検討する場合、

・契約上、もしくは法的、もしくは規制上の理由で、保持あるいは破棄しなければならないデータは何か

・保存するその他のデータをどのように決定するか

・データはどの程度の期間、保持・保存するか

といったポイントについて踏まえておく必要があります。

また、データセットの長期保存計画を検討する際には、

どのリポジトリやアーカイブなどにデータを保存するかを考慮しておかなければなりません。

3.4.9 ⑦データ共有

データを共有するか否か、また、限定共有するかなど、共有の方針は各研究プロジェクトによって異なります。

このため、計画を作成する際に、データをどのように共有するか、また、データ共有に何らかの制限が必要か、というポイントについて事前に検討しておく必要があります。

3.4.10 ⑧責任とリソース

責任とリソースについては、データ管理計画の作成時に、データ管理の責任者は誰か、そして計画を実行するために、どのようなリソース、資源が必要かというポイントについて確認および検討する必要があります。

3.4.11 データ管理計画作成ツール

データ管理計画を作成する際には、無料で使えるツールを利用すると、作成の負担を抑えることができます。

テンプレート等を用いて、各助成機関等の要件に沿った計画書を容易に作成することができます。また、共有・公開機能を用いて、他の研究者の作成した計画書を参考にすることもできます。研究データ管理の支援者は、このようなツールについての情報提供をおこない、その使い方についてサポートをするとよいでしょう。

データ管理計画作成ツールの例を紹介します。

英国のデジタル・キュレーション・センターのDMPOnline（ディーエムピーオンライン）は、誰でも無料でアカウントの作成ができます。

また、カリフォルニア大学のDMPTool（ディーエムピーツール）は、2018年3月に最新版がリリースされており、こちらも無料でアカウントを作成することができます。

3.4.12 データ管理計画のライフサイクル

データ管理計画は、研究の初期段階で作成すれば終わり、という訳ではありません。

研究をしている最中（さなか）にも、必要に応じて計画の修正を行うことが必要です。絶えず計画の更新を行うことで、その計画はより実践的で有効性のあるものとなるでしょう。

また、研究終了後は、計画の達成状況について検証を行うことが大切です。

研究データ管理の支援者には、研究のプロセスに沿って、研究者が計画の更新をスムーズに行えるよう支援をしていくことが求められます。

第4章　研究中の支援

4.1 第4章の概要

第4章では、研究中の支援として、研究データの種類とセキュリティポリシーの関係、データの保存と機関の役割、データの発見支援、データ分析トレーニングコースの実施、論文発表に向けた研究データの取扱いに関する支援、データ管理計画の更新支援、について学んでいきます。

まず、「研究データの種類とセキュリティポリシーの関係」では、研究データを含むさまざまな情報を「機密性・完全性・可用性」の観点で格付けする方法を学びます。ほぼ全ての研究データは、格付けの結果「要保護情報」に分類される可能性が高く、各機関で定められた「取扱い制限」に従った管理が必要となります。

次に、「データの保存と機関の役割」では、研究データの適切な保存方法と、それを実施するための研究者・研究主催者・研究機関それぞれの役割を学びます。論文の根拠となるデータは論文発表後10年間の保存が原則とされています。研究機関は、組織として研究データを安全に保存するためのストレージ領域を提供することが望まれます。

実際に研究データを保存する際は、ネットワークドライブや、組織として契約したオンラインストレージなどが適切な保存場所といえます。他にも、研究データのバックアップやセキュリティ対策、保存時に推奨されるファイルフォーマットなど、研究データ保存のポイントについて具体的に学びます。

そして、「データの発見支援」では、研究中に自分が作成したデータ以外の既存のデータを探して活用するための検索方法を学びます。リポジトリやデータジャーナル、論文から探す方法や、利用申請をして非公開のデータを探す方法、関連する研究プロジェクトや研究者を探しデータを見つける方法などを具体的に学びます。

「データ分析トレーニングコースの実施」では、研究成果に大きく関わる「データ分析・可視化」の支援方法を学びます。データ分析トレーニングコースを実施する際は、機関としてその重要性を認識し、組織内で関連部署が連携し、また研究データ管理の流れの中にトレーニングを位置づけ、さらに適切な情報基盤の利用も促しながら行うことが重要です。

さらに、「論文発表に向けた研究データの取扱いに関する支援」では、論文発表の前に行う、関係者との契約・合意内容や、所属機関・助成機関・学会や出版者等のポリシーを確認することの重要性を学びます。また、研究で参照したデータを論文で引用することの意義と、具体的な引用の記載方法についても学びます。

「DMPの更新支援」では、研究の最初に作成したデータ管理計画＝DMPを研究実施中も必要に応じて振り返り、更新していく重要性を確認し、DMP作成ツールを用いた実際の更新方法を学びます。

それでは、今回も研究データ管理について学んでいきましょう。

4.2 研究データの種類とセキュリティポリシーの関係

ここでは、研究データの情報セキュリティを向上させるための、情報の格付け方法と、取扱制限の実施方法について学びます。

4.2.1 情報セキュリティと情報の格付け

情報セキュリティとは、取り扱う情報の、機密性、完全性および可用性を維持することです。研究データは、情報セキュリティを考慮すべき情報の一つです。

情報セキュリティ対策を実施する際は、始めに、取り扱う情報に対して、機密性、完全性、可用性に対するレベルの格付けを行う必要があります。研究データも種類に応じて格付けを行います。

続いて、格付けに応じた取扱い制限に従い、研究データの保存や保管を行います。これにより、情報セキュリティの向上を図ることができます。

ほとんどの大学や研究機関で規定されているセキュリティポリシーの要求を満たすためには、情報セキュリティ対策の実施が必要です。

4.2.2 機密性・完全性・可用性とは

情報セキュリティ対策の第一段階である、情報の格付けを行う際の３つの指標：機密性、完全性、可用性の意味を学びます。

機密性とは、認可された者だけが情報にアクセスできるようにすること、を意味します。機密性を維持するためのセキュリティ対策は、情報漏えい防止や不正アクセス対策です。

完全性とは、情報や情報の処理方法が正確で完全であること、です。完全性を維持するためのセキュリティ対策は、改ざん防止やデータの差異の検出などです。

可用性とは、許可された者が、必要な時に、情報や情報資産に確実にアクセスできるようにすること、です。可用性を維持するための対策は、バックアップ、システムの二重化などです。

4.2.3 機密性についての格付けの定義例

次に、機密性、完全性、可用性に対する格付けの定義の例について学びます。定義は、格付けをする際の指標となるため、十分に理解しておくことが重要です。

機密性には、3つのレベルが定義されています。

機密性3の情報は、秘密文書に相当する機密性を要する情報です。行政文書の管理に関するガイドラインでは、秘密文書は「極秘文書」と「秘文書」に分けられています。極秘文書は、国の安全・利益に関わる保全の必要が高い行政文書です。秘文書は極秘文書に次ぐレベルの秘密で、関係者以外に知らせてはならない情報を含む、極秘文書以外の行政文書です。

機密性2の情報は、秘密文書ほどの機密性は必要としないが、漏えいが、利用者の権利や大学・研究機関の活動に影響し得る情報です。

機密性1の情報は、機密性2及び3の情報以外の情報となります。

4.2.4 完全性・可用性についての格付けの定義例

完全性と可用性にはそれぞれ2つのレベルが定義されています。

完全性2の情報とは、改ざん、誤びゅう又は破損により、利用者の権利や大学・研究機関の活動に影響し得る情報です。

完全性1の情報は、完全性2の情報以外の情報となります。

可用性2の情報とは、滅失、紛失又は当該情報が利用不可能であることにより、利用者の権利が侵害され、または大学・研究機関の活動の安定的な遂行に支障を及ぼすおそれがある情報を指します。

可用性1の情報は、可用性2の情報以外の情報となります。

4.2.5 要保護情報とは

機密性、完全性、可用性が一定のレベル以上にある情報は、定められた方法で管理をしなければなりません。

機密性3または機密性2の情報は要機密情報、完全性2の情報は要保全情報、可用性2の情報は要安定情報と呼ばれ、まとめて要保護情報と呼ばれます。

要保護情報については、各大学や機関で定義された取扱制限に基づいて、情報を管理する必要があります。

なお、機密性、完全性、可用性の全てが1の情報は、非保護情報と呼ばれることがあります。

4.2.6 取扱制限の定義例

次に、要保護情報を管理する際の指針となる「取扱制限」の定義について学びます。

大学や研究機関の情報セキュリティ責任者は、情報について、機密性、完全性、可用性の3つの観点を区別し、それぞれにつき取扱制限を定義する必要があります。

例えば、機密性が高い場合には複製を禁止したり、配布する際には許可を求めたり、暗号化を必須化することが挙げられます。完全性や可用性が高い場合についても同様に、取扱い制限の定義を定めていきます。情報セキュリティ責任者は、ここに示されるような取扱制限を3つの観点それぞれについて、事前に定めなければなりません。

教職員等は、扱う情報を格付けし、事前に定められた取扱制限の定義から実施する取扱い制限を指定して、規定に沿った管理を実施します。

なお、「どのレベルでどのような取扱制限を定めるのか」の決まりはありません。各機関がそれぞれレベルに応じた取扱制限を事前に定め、実施する必要があります。

4.2.7 セキュリティポリシーと研究データの管理

実際に、研究資料やデータの格付けについて考えてみましょう。

個人情報を含んでいたり、漏洩によるインパクトの大きい研究データは、機密性3、完全性2、可用性2に格付けされると考えられます。

それ以外の非公開情報は、機密性は2が該当し、完全性、可用性は2と考えらます。

リポジトリ等で公開しているデータの機密性は1と見なすことができますが、その内容が改ざんされてはならないので完全性は2となります。また、資料やデータがなくなってしまうと研究の遂行に影響を及ぼしますので、可用性も2と考えられます。

オリジナルのデータが別のリポジトリなどでも公開されている場合は、機密性は1、完全性は2となり、可用性は1と考えられます。

機密性、完全性、可用性が全て1の場合は取扱制限を受けない非保護情報となります。機密性はないが公開を前提としない参考資料で、消失しても再作成又は再入手が容易で、研究遂行上支障のないような情報です。

このように見ていくと、ほぼ全ての研究データは要保護情報であり取扱制限の対象となっています。このため、事前に定められた取扱制限に従って研究データを管理する必要があるのです。

4.2.8 研究データの具体的な格付け例

格付け作業のイメージをつかむために、より具体的な例を示します。

研究活動で生じる情報は、その種類や状態に応じて、格付けが異なることがわかります。ここに表示されていないものも含めて、日々の研究で使われる情報は格付けをおこなうことが必要です。

格付けをした後は、それぞれの要保護情報に対する取扱制限を規定する必要があります。研究データを管理するシステムは、取扱制限をサポートする機能を持つものを選定する必要があります。

4.3 データの保存と機関の役割

ここからは、研究中の研究データ管理に適したデータの保存や、機関の役割について学びます。ここで学ぶ内容は、研究後の長期的なデータ保存においても同様に適用される、重要な内容です。

4.3.1 保存の必要性と機関の責任

研究データを保存する必要性については、2015年に文部科学省によるガイドラインをもとに日本学術会議で審議がなされました。

審議の結果、研究者及び研究機関に対し、原則として10年間研究データを保存することが義務づけられています。

保存義務の対象となる研究データは、実験ノート・数値データ・画像などを含む論文の根拠資料です。

保存の義務に加え、必要に応じて研究データの開示や公開が推奨されています。

研究データを作成・管理する教職員が異動したり、退職した後も、機関として研究データを保存したり、データへのアクセスを確保することが求められています。

日本学術会議は、研究データを保存する際の具体的な責任者とその役割についても言及しています。

研究者には、研究記録やメタデータの整理により、検索・抽出可能な形での研究データの整理・保管を行い、また適正なバックアップを作成することが求められています。

研究主宰者には、研究者の教育や指導に加え、メタデータ管理や統一フォーマットの作成が求められます。

研究機関の長には、データ・バックアップ用サーバーなどのインフラを整備をすることが望まれています。

また、セキュリティポリシーへの準拠の観点からも、各ステークホルダーに加えて「組織として」研究データのストレージ領域を提供することが望まれます。

4.3.2 各機関が定めるポリシーの例

文部科学省及び日本学術会議での議論を受け、各大学で研究データの保存に関するガイドラインが整備されています。ここでは、3つの大学の事例を学びます。

まず、3つの大学のポリシーでは共通して、論文発表後10年間の保存期間を原則として定めています。

京都大学では、「公正な研究活動の推進等に関する規定」に関連して、必要に応じて保存する研究データの開示を義務付けています。

大阪大学では、「研究データ保存等に関するガイドライン」において、調査委員会に求められた時の研究データの開示を義務付けています。

九州大学でも、「研究データの保存等に関するガイドライン」において、求めに応じ研究活動の適正性を説明することや研究データの開示を義務付けています。

4.3.3 研究データ保存時の留意点

研究データを適切に保存するには、これまで述べてきた国や機関のポリシーの要求を踏まえて、保存場所を選択し、バックアップ体制を整備し、セキュリティ対策を実施し、最適なフォーマットを選択することが重要です。また前提として、「データ管理計画書」に沿ったデータの保存を行うことも大切です。

ここからは、これらの留意点について、より具体的に説明をしていきます。

4.3.4 研究データ管理に適切な保存場所

はじめに、研究データを管理するために適切な保存場所について学びます。

デスクトップPC（ピーシー）・モバイルPC（ピーシー）は、手軽な保存先ですが、ハードウェアが冗長化されておらず、障害によりファイルを失う可能性が高くなります。このため、研究データを管理する保存場所としては適切ではないと言えます。

外部記憶装置の場合、CD（シー・ディー）やDVD（ディー・ヴイ・ディー）は容量が不十分であり、書き込みエラーへの対処や確認も必要です。USB（ユーエスビー）ハードディスクやUSBメモリ、SD（エスディー）カードなどは大容量化してきていますが、デスクトップやモバイルPCと同じく障害によりファイルを失う可能性が高くなります。外部記憶装置は、持ち運び易い反面、紛失や盗難などのセキュリティリスクも大きくなります。このため、研究データを管理する保存場所としては適切ではありません。

ネットワークドライブとは、LAN（ラン）などのネットワーク経由で接続できる他のコンピュータのディスクスペースを指します。所属機関のシステム管理部門が提供し、安定的に運用されるネットワークドライブであれば、研究データを管理するのに適切な保存場所といえます。

オンラインストレージは、クラウドストレージとも呼ばれ、インターネット上のディスクスペースを利用できるサービスです。一定容量までは無料で利用できるものもありますが、機関のセキュリティポリシーやガイドラインに合ったオンラインストレージを選ぶことが大切です。

4.3.5 オンプレミスとパブリッククラウド

研究データ管理において適切な保存場所であるネットワークドライブとオンラインストレージの利用方法について、詳しく学びます。

大学等の機関のシステム管理部門が提供するネットワークドライブは、多くの場合、機関内でハードウェアを購入し、オンプレミス型のサービスとして提供されています。

オンラインストレージには、大きく分けて二つの利用形態があります。ひとつは、組織が契約するクラウドストレージを利用する場合で、もうひとつは、個人がフリーミアムのクラウドストレージを利用する場合です。

組織が契約するクラウドストレージを利用する場合には、まず付随するアプリケーションから利用する場合があります。教育機関向けのBOX（ボックス）、OneDrive for Business（ワンドライブフォービジネス）のサービスはこれに該当します。

また、組織が別途運用するアプリケーションからクラウド上のストレージを利用する場合もあります。例えばownCloud（オウンクラウド）と呼ばれるアプリケーションを通じて、組織が契約したクラウド上のディスク領域を利用する場合がこれにあたります。

これとは別に、個人がアカウントを登録して気軽に利用できるフリーミアムのオンラインストレージがあります。フリーミアムとは、サービスが無料で提供され、高度な機能などを使用する際に課金する形態のサービスです。

次のパートでは、情報セキュリティの観点から、どのようなオンラインストレージを利用するべきかを説明していきます。

4.3.6 オンラインストレージの情報セキュリティ

一般的に、商用クラウドサービスは「外部委託」しているサービスの一部と見なすことができます。「外部委託」とは、大学の情報処理業務の一部または全部について、契約を通じて学外の業者などに実施させることをいいます。

商用クラウドサービスは、利用者の「情報セキュリティに関する十分な条件設定の余地」があるかないかによって、外部委託における位置づけが異なります。

充分な条件設定の余地のあるものは「クラウドサービス」に分類されます。クラウドサービスは、情報セキュリティ対策を講じるための柔軟な設定ができることを意味します。先ほど紹介した組織が契約するクラウドストレージは、このクラウドサービスに分類されます。

一方で、十分な条件設定の余地のない商用クラウドサービスは「約款による外部サービス」と見なされます。「約款による外部サービス」は、民間事業者等の学外の組織が約款に基づきインターネット上で提供する情報処理サービスのことです。条件設定の余地がないため、情報セキュリティ対策を講じるために必要な設定ができない点に注意が必要です。先ほど紹介したフリーミアムで気軽に利用できるクラウドサービスは「約款による外部サービス」に分類されます。

多くの機関のセキュリティポリシーでは、「約款による外部サービス」を利用する際は、要機密情報を扱うことが禁止されています。こうしたセキュリティポリシーを遵守するためにも、オンラインストレージで研究データを保存する場合は、組織として契約されたサービスを利用するのが、妥当であると考えられます。なお、「クラウドサービス」の利用においても、機関で定められた「外部委託」や「クラウドサービスの利用」に関する規程を遵守する必要があります。

4.3.7 海外事例：機関としてのストレージ提供サービス

先行する海外の大学では、研究データ管理のためのストレージを提供しています。

英国のエディンバラ大学では、RDM DataStore(アールディーエム　データストア）と呼ばれるサービスを提供しており、全研究者に0.5TB（テラバイト）の領域を無償で提供し、0.25TBまでを共同研究プロジェクトに利用可能としています。さらに有償で1PB（ペタバイト）まで利用可能としています。また、DataSync（データシンク）と呼ばれるDropbox（ドロップボックス）のようなユーザインターフェースをもつサービスも提供しており、DataSyncはDataStoreと接続可能となっています。DataSync単体でも、20GB（ギガバイト）の領域を利用できます。

パーデュー大学のストレージサービスでは、全教職員や学生に100GB（ギガバイト）のストレージ領域を提供し、そのうち、1GBまでを公開領域として利用できます。また、外部資金獲得者には1TBの領域を提供し、そのうち10GBまでを公開領域として利用できます。

このように、先行する海外の大学では、無料のストレージ領域を提供するだけではなく、有料で追加領域を提供するなど多彩なサービスの提供が進んでいます。

4.3.8 バックアップ

次に、研究データのバックアップについて学びます。

不慮のトラブルで研究データを失った際にも、バックアップデータがあれば、トラブルの影響を最小限に抑えることができます。また、不用意なファイルの削除や上書き、システム障害にも対応できます。

機関が提供するネットワークドライブが、バックアップサービスを提供している場合もありますが、バックアップの条件を事前に確認し、必要に応じて追加のバックアップを実行することが望まれます。オンラインストレージを利用する際には、機関のセキュリティポリシーに合致しているかどうかの確認が必要です。

バックアップは手動ではなく、バックアップソフトウェアなどを利用して定期的に取得することを勧めます。また、バックアップからリストアできることを事前に確認しておくことも重要です。

ディスク容量やバックアップに必要な時間を勘案し、フルバックアップ、増分・差分バックアップから適切な方法を選択することも大切です。例えば、毎日のバックアップは差分バックアップを実行し、毎月1回フルバックアップを実行するという方法もあります。

また、災害対策として、物理的に離れた複数の場所にバックアップを取っておくことも大切です。

4.3.9 セキュリティ

続いて、セキュリティ対策について学びます。

セキュリティにまつわる問題は、ハードウェアトラブルだけではありません。盗難やウイルスにより研究データを失ったり、データが外部に流出してしまう危険もあります。物理的なセキュリティ対策と、PC上でのセキュリティ対策、個人情報・プライバシー保護を合わせて進めることが大切です。

物理的なセキュリティ対策としては、席を離れる際にPCをロックすること、部屋の施錠を徹底すること、モバイルPCの放置を禁止すること、USBメモリ等の紛失を防ぐこと、などの取り組みが必要です。

PC上でのセキュリティ対策としては、アンチウィルスソフトウェアやファイアウォールの導入、ソフトウェアやOS（オーエス）のアップデート、適切な強度のパスワードや認証方法の利用、ファイルやハードディスク、USBメモリ等の暗号化などが必要です。

個人情報やプライバシー保護の対策として、センシティブなデータを扱う場合は、機密性や完全性の観点で取り扱い方法を明確にした上で、適切な保存や解析の環境を提供することが重要です。

ここまで述べたさまざまな対策は、多くの機関のセキュリティポリシーを遵守するうえで不可欠な内容となっています。セキュリティポリシーに沿った研究環境を整備することで、トラブルを回避し、円滑な研究活動が可能となります。

4.3.10 フォーマット

研究データを保存する際には、適切なフォーマットを選択することも重要です。

研究データの共有、公開、保存では、「論文を執筆するまで」というような短期的な視点ではなく、長期的な視点でデータの利用可能性を確保する必要があります。その際、どのような形式でファイルを保存すべきか、ということが問題となってきます。

一般的に、.pdf（ドットピーディーエフ）や.jpg（ドットジェーペグ）といったファイルの拡張子は、ファイルフォーマットに関する情報を示しています。

長期的な保存のために推奨されるファイルフォーマットは、特定の商用ソフトウェアのみで読み込み可能なフォーマットではなく、仕様が国際標準となっているフォーマットや、広く普及しているフォーマットです。

推奨フォーマットにあわせるため、特定の商用ソフトウェアで作成されたファイルフォーマットから、より汎用的なフォーマットに変換する場合には注意が必要です。フォーマット変換により、ファイル内部に記録されたメタデータや画像の品質、文章構造などの情報が失われる可能性があります。フォーマット変換のリスクを十分考慮したうえで、オリジナルデータをどのようなフォーマットで保存するかを決定する必要があります。

また、ファイルフォーマットは、テキスト形式とバイナリ形式に大別できます。テキストファイルは環境に依存せず開くことができるため、人による可読性が高くなります。バイナリファイルは豊富な情報を含むことができるが、人による可読性が低い、という特徴があります。こうした特徴も考慮したうえで、研究データの共有、公開、保存に適したフォーマットを選択する必要があります。

4.4 データの発見支援

研究中の支援として、ここからは研究に活用するデータの発見支援について学んでいきます。

4.4.1 研究中のデータ利用

研究の最中（さなか）に用いるデータは、必ずしも実験や観測などを通して研究者自身が生成・取得するデータのみとは限りません。他の研究者などにより生成された既存のデータを利用することも考えられます。

既存のデータを利用するには、公開されているデータセットからデータを発見・入手する方法や、存在は公開されているがデータセット自体は非公開となっており申請等によって入手可能なデータを発見・入手する方法、自身の研究に関連性の高い研究プロジェクトや研究者を探し出し、保有されているデータセットについて照会する方法などが考えられます。

研究データ管理の支援者は、研究者が自身の研究に活用できるデータをスムーズに発見できるよう、情報提供や調査支援を行うことが望ましいでしょう。また、利用にあたっては、提供者の利用条件に従う必要がある旨の注意喚起をするとよいでしょう。

ここからは、研究データの発見支援の具体的な方法を見ていきます。

4.4.2 公開されている研究データの検索支援

公開されているデータセットを探し出すには、いくつかの方法が考えられます。

ここでは、一括検索できるツールで複数のデータリポジトリからデータを検索する方法、個別に各データリポジトリにアクセスして検索する方法、データジャーナルから探す方法、論文の引用情報からデータを参照する方法の4つについて見ていきます。

4.4.2 公開されている研究データの検索支援①複数のデータリポジトリ等から一括検索できるツールで探す

研究データを手軽に効率よく検索できるのが、多数のデータリポジトリ等を対象に、キーワードなどで一括検索ができるツールです。ここでは、代表的な検索ツールの例を学びます。

Clarivate Analytics（クラリベイトアナリティクス）のData Citation Index（データサイテーションインデックス）はさまざまな分野の国際的なデータリポジトリから採録された研究データを検索できる有料データベースです。データの被引用数の情報も提供しています。

Elsevier（エルゼビア）のDiscover Mendeley Data（ディスカバー・メンデレイ・データ）では、複数のデータリポジトリ等に収録された研究データを横断して検索することができます。

DataCite（データサイト）では、複数のデータリポジトリ等に収録された研究データで、DataCite（データサイト）を経由してデータの永続的な識別子であるDOI（ディーオーアイ）が付与された研究データを、一括検索することができます。

日本では、国立情報学研究所のJAIRO（ジャイロ）の詳細検索で資料種別「データ・データベース」を指定すると、日本の機関リポジトリに登録されている研究データを一括検索することができます。

これらの一括検索ツールは手軽である一方、検索できる対象や項目が限られていることに注意が必要です。このため、個別にそれぞれのデータリポジトリにアクセスして検索することも必要となるでしょう。

4.4.2 公開されている研究データの検索支援②個別に各データリポジトリにアクセスして探す

次に、個別のデータリポジトリへのアクセスによるデータの検索支援について見ていきます。

リポジトリには、分野特化型のデータリポジトリ、分野を問わない汎用的なデータリポジトリや、機関のリポジトリなど多種多様なリポジトリが存在します。

分野別のデータリポジトリにはさまざまなものがありますが、どのようなデータリポジトリが存在するのかは、データリポジトリのディレクトリre3data（リースリーデータ）で分野を指定して探すことができます。

分野を問わない汎用的なデータリポジトリには、figshare（フィグシェア）やZenodo（ゼノド）、Dryad（ドライアッド）、Mendeley Data（メンデレーデータ）などがあります。

小規模のデータセットなどは、データリポジトリではなく、機関リポジトリに登録されている場合もあります。このため、探しているデータによっては、機関リポジトリもあわせて参照するよう提案するのがよいでしょう。機関リポジトリの検索には、OpenDOAR（オープンドア）などの機関リポジトリのディレクトリを参照してください。

4.4.2 公開されている研究データの検索支援③データジャーナルから探す

データそのものを記述する「データ論文」が掲載されているデータジャーナルから、研究に有用なデータを探すこともできます。

対象のデータジャーナルが一般的な論文データベースに収録されていれば、通常の論文とあわせて検索することができます。

データジャーナルは、論文と同様の出版プロセスを経ているため、査読により一定の品質保証がなされているというメリットもあります。

ここでは、いくつかの分野の代表的なデータジャーナルの例を挙げています。

出版社のデータジャーナルには、Springer Nature（シュプリンガーネイチャー）のScientific Data（サイエンティフィックデータ）や、Elsevier(エルゼビア）のData in Brief（データインブリーフ）があります。

分野別のデータジャーナルには、地球科学のGeoscience Data Journal（ジオサイエンスデータジャーナル）や、考古学のJournal of Open Archaeology Data（ジャーナルオブオープンアーケオロジーデータ）、生物多様性分野のBiodiversity Data Journal（バイオダイバーシティデータジャーナル）、地球システム科学のEarth System Science Data（アースシステムサイエンスデータ）などがあります。

さらに、日本の学術機関による初のデータジャーナルとして2017年1月に、国立極地研究所のPolar Data Journal（ポーラーデータジャーナル）も創刊されています。

4.4.2 公開されている研究データの検索支援④論文の引用情報からデータを参照する

論文の引用文献として、原著論文だけでなく研究データが挙げられる例も次第に増えてきています。このような論文の引用情報からデータを参照する方法もあります。

ここに示すように、論文の参考文献リストに記載されたデータの引用情報をたどると、データリポジトリで公開されるデータセットにたどりつくことができます。

論文の引用情報からの検索は、特に研究者が先行研究調査をする際に、自身の研究に関連のある研究データを発見するのに有用な方法と言えるでしょう。

なお、研究データの引用の記述方法については、この後のパートで説明していきます。

4.4.3 所定の手続等により入手可能な研究データの発見支援

データリポジトリなどでデータが公開されていなくても、データセットの存在や概要についての情報が公開され、申請を行うことで利用できるデータもあります。

研究データ管理の支援者は、研究者に対してこのようなデータセットの存在について、情報提供や入手支援を行うことが望ましいでしょう。また、利用条件は各提供者の方針により異なりますので、必要に応じて研究者に注意喚起をすることも大切です。

利用申請を行うことで利用できるデータの例を見ていきます。

東京大学社会科学研究所附属社会調査・データアーカイブ研究センターの「SSJ（エスエスジェー）データアーカイブ」は、各種調査機関と連携し、社会科学に関する大量の一次データを収集しています。検索をして利用したいデータを発見したら利用申請を行います。

バイオサイエンス・データベースセンターの提供する「NBDC（エヌビーディーシー）ヒトデータベース」では、生命科学分野の、人に関するデータを公開しています。制限公開データの利用には利用申請が必要になります。

国立情報学研究所の「情報学研究データリポジトリ」は、情報学関連分野の研究者向けに各種データセットを提供しています。各データセットの利用条件に従い、必要なものには利用申請を行います。

4.4.4 関連する研究プロジェクトの調査支援

研究データの存在そのものが明らかでなく、検索による調査ができない場合には、自身の研究に関連性の高い研究プロジェクトや研究者を探し、保有しているデータセットについて照会することで、研究に必要なデータを入手できるかもしれません。研究データ管理の支援者は、研究プロジェクトや研究者の調査に活用できるツールについて、情報提供を行ったり、調査の支援を行うとよいでしょう。

具体的なツールの例を見ていきましょう。

国立情報学研究所のデータベース「KAKEN（カケン）」では、文部科学省および日本学術振興会が交付する、科学研究費助成事業により助成を受けた研究の情報を探すことができます。「研究課題を探す」場合と、「研究者を探す」場合の両方について、情報の検索をすることができます。

科学技術振興機構（JST＝ジェーエスティー）の「プロジェクトデータベース」では、JSTの競争的資金制度による研究課題などを探すことができます。

同じくJSTの「Researchmap（リサーチマップ）」は、日本の研究者情報データベースであり、研究者の経歴や研究成果などを参照することができます。

米国国立衛生研究所（NIH=エヌアイエイチ）の「NIH RePORTER（エヌアイエイチリポーター）」ではNIHの研究助成を受けた研究課題を探すことができます。

同じく欧州では、欧州委員会から助成を受けた研究課題を探すことができる「OpenAIRE（オープンエアー）」があります。

4.5 データ分析トレーニングコースの実施

ここでは、実際の研究活動に大きく関わるデータ分析のトレーニングコースについて学びます。トレーニングの意義や、先進大学の事例、トレーニングコースを実際に提供する際のポイントについてそれぞれ学んでいきます。

4.5.1 データ分析トレーニングの必要性

一般的な研究データ管理＝RDM（アールディーエム）のためのトレーニングコースでは、多くの場合、研究データのライフサイクルにしたがって、どのように研究データを扱うかを順番に学んでいきます。

ここに挙げたような、データ管理計画や共有、公開方法、メタデータの記述や適切なデータ保存方法などについては、多くの教材が提供されていますが、データを加工したり、分析することについては、見落とされがちです。

一般に、研究データの加工や分析のサポートは、通常の研究データ管理サービスの枠外とみなされる傾向がありますが、これらの支援活動が、組織として重要になりつつあります。

4.5.2 研究者のニーズ

このグラフは、オーストラリアの研究機関、オーストラリア・バイオインフォマティクス・リソースが国内の生物学分野の研究者を対象におこなった調査結果の一部です。

左のグラフは、バイオインフォマティクスを実践していく上での課題、右のグラフは研究機関に何を望むかについての回答です。どちらも、トレーニングが上位を占めています。

この調査では、研究データの処理や分析手法に対する専門知識や技術の欠如が、研究の推進を制限する主要な要因であることが指摘されています。

生物学分野に限らず、研究活動を促進していくためには、専門的な知識や技術を習得するためのトレーニングコースの実施が重要であるといえます。

4.5.3 データ分析トレーニングの意義

データ分析を学ぶことにはどのような意義があるのでしょうか。

データの分析は、主たる研究活動であり研究成果に関わる重要な活動です。また、データの効果的な可視化は、学会や論文で研究成果を表現するために不可欠です。

データの加工や分析、可視化が必要となるのは、データ中心科学やビッグデータのような大量データを扱ってきた分野だけでなく、幅広い分野で必要とされてきています。

オープンサイエンスの潮流の中で、そうした傾向は更に加速し、さまざまな分野のデータ分析に基づく研究成果が、研究データ利活用のエコシステムを構成していくと考えられます。

研究の現場における、データ分析のニーズの高まりに応えるために、データ分析のトレーニングコースを研究者や学生に提供する取り組みが、注目されつつあります。

4.5.4 Data Carpentryによるトレーニングコース

データカーペントリーは、研究者のデータリテラシー教育を目的として、分野特化型ワークショップを有料で実施する会員制のNPO（エヌピーオー）です。これまでに、世界23か国で合わせて200回以上のワークショップを実施してきた実績があります。

データカーペントリーが実施するトレーニングコースは、コンピュータによるデータ分析について、ほとんど知識がない研究者を対象としています。分野に特化したデータの効率的な管理や組織化、解析、可視化のトレーニングを行っており、具体的にはR（アール）やPython（パイソン）といった、データの加工・分析のためのツール紹介や、SQL（エスキューエル）を使ったデータベース利用、シェルスクリプト入門などのコースがあります。

現在は、エコロジーとゲノムの分野のコースを開催しており、他の分野を対象としたワークショップの準備も進めています。

4.5.5 学内におけるサポート（エディンバラ大学）

大学が、その構成員にトレーニングコースを提供している事例として、エディンバラ大学の取り組みを挙げることができます。

そもそも、エディンバラ大学が実施するトレーニングコースは、研究のIT（アイティー）化やデータサイエンスに対する大学としての戦略設定に基づくものです。組織としてトレーニングコースを実施し、研究者をサポートしていくことが、研究力の向上に繋がる、という認識を持っている点は非常に重要です。

トレーニングコースを計画するにあたっては、まず、研究者や学生が日々の研究活動で利用している解析ツールや、希望する実施形態について調査を行っています。それらを基に、学内の組織が連携してトレーニングコースの実施形態を検討した上で、研究データ管理サービスの一環として、統計解析のコースを実現しました。

さらに、エディンバラ大学では、分野特化型のツールについてのコースも提供されています。

4.5.5 学内におけるサポート（シドニー大学）

次にシドニー大学のデータ分析に関するサービスやトレーニングコースの提供事例について学びます。

シドニー大学では、Sydney Informatics Hub（シドニー・インフォマティクス・ハブ）という組織が、データ分析や可視化に関するコンサルティングやトレーニングコースの提供を行っています。

サービスの内容としては、データ分析のサポート以外にも、そのために必要なHPC（エイチピーシー＝ハイパフォーマンスコンピューティング）や仮想マシンの使い方に関するサポートを一貫して行っています。

R（アール）やPython（パイソン）のコースに加え、解析する前にデータをクリーニングするためのツールの利用方法も、トレーニングコースの内容に含まれています。

4.5.6 無料で利用できる代表的なツール

データ分析のトレーニングコースで利用されているフリーツールの一覧がこちらです。

海外の先進的な大学では、ここで挙げたような、データクリーニング、定性的な分析、統計・定量的な分析、テキストマイニングやデータマイニングなどに利用できるさまざまなツールを対象としたトレーニングコースが実施されています。

こうしたフリーのツールに加えて、大学がライセンス契約をしているツールを使ったコースを提供している事例もあります。

4.5.7 データ分析トレーニングコースの実施のポイント

データ分析のトレーニングコースを実施するためのポイントについてまとめておきます。

まず、データ分析トレーニングコースの実施を、大学として確立していくためには、エディンバラ大学の事例で紹介したように、機関としての位置づけが重要になります。データ解析や可視化の技術の向上が、学力や研究力の向上に繋がる重要な要素であることを、組織として明確に位置付けていることが、先行大学に共通するポイントです。

また、対象とする内容が多岐にわたるとともに、専門性を必要とする場合もあることから、組織内の連携が重要になります。図書館、情報基盤センターだけでなく、データ分析や可視化の技術に関する専門家の多い学部や学会と連携してトレーニングコースを実施するのが効果的です。

さらに、先行大学では、研究データの管理を有効に実施する一連の流れの中で、データ分析や可視化のトレーニングコースを位置付けているという特徴があります。

最後に、組織の提供するHPC（エイチピーシー）や仮想マシン環境、ストレージ環境の利用と関連付けて、計画していくことがカギとなります。

以上が、データ分析のトレーニングコースを実施するためのポイントとなりますが、まずは研究データ管理の支援者自らが、データ分析や可視化を実際に体験することも有効です。海外の大学図書館では、電子ジャーナルやリポジトリの利用統計データや、図書の貸出記録などを対象として、図書館員自らがデータ分析を学んでいるという事例もあります。海外で提供されている教材やサンプルコードを参考にしながら、データ分析を体験することを勧めます。

4.6 論文発表に向けた研究データの取扱に関する支援

ここからは、論文発表に向けた研究データの取扱いに必要な支援について、特に関係機関のポリシーとデータの引用方法について見ていきます。

4.6.1 データの取扱指針の確認

研究成果の論文発表に向けて、データをどのように取り扱うべきかを確認する際は、まず初めに各ステークホルダーのポリシーによる規定の有無や、その内容について確認しておく必要があります。

特に、データ取得時に関係者と何らかの契約や合意を締結した場合や、所属機関や助成機関、学会・出版者によるポリシーが関係する場合は、注意が必要です。

論文の発表にあたっては、用いる研究データにどのような研究倫理上のルールや、公開に関する要件が定められているか、事前に確認しておくようにしましょう。

4.6.2 出版者によるデータ共有ポリシー

最近では、特に出版者がデータ共有を求めるポリシーを策定する例が相次いでいますが、データ共有の要求レベルや、共有手段についての指示は各出版者によってさまざまです。

データ共有については、査読や出版の条件としている場合や、奨励している場合、論文補足資料としてデータを受け付ける場合などがあります。またポリシーで明示されていない場合もあります。

データの共有手段も、リポジトリを限定せず共有を指示する場合、指定したリポジトリで公開を求める場合、データディレクトリから選択する場合、また機関リポジトリや助成機関のリポジトリが認められる場合などさまざまです。

論文公開に先立って、研究者に出版者の最新のポリシーを提供することは重要な支援サービスのひとつです。また、出版者がデータ共有の場として、一定の条件を満たした機関リポジトリを認めている場合もあります。所属機関のリポジトリが、出版者の要件を充たしているかを確認し、充たしている場合は、データ登録先の候補の1つとして案内するのもよいでしょう。

4.6.3 データの引用

論文公開に先立つデータの取扱いについて、ポリシーに加えて注意する点として、データの引用があります。近年、論文を執筆する際、研究に利用した研究データを、引用文献として掲げることが推奨されています。ここからは、データ引用の重要性と、具体的なデータ引用の方法について学びます。

研究データを独立した研究成果として引用することで、次のようなことが期待できます。

・執筆者は、研究成果の根拠となった情報源を読者に対して明確に提示することができます。

・読者は、研究に利用された研究データを容易に特定することができます。

・第三者が、引用された研究データを用いて、研究結果の再現可能性を検証することにより、学術的な透明性を高めることができます。

・引用されることで、引用された研究データの影響度を計測・追跡することができます。

・引用された研究データの作成者の功績が認められ、業績評価につながります。

支援スタッフは、研究者に対してこれらのメリットを説明し、意識を喚起をするとともに、データの引用方法についての情報提供を行うことが望ましいでしょう。

4.6.4 データ引用原則に関する共同宣言 FORCE11（2014年）

データ引用を推進する国際的な動きとして、FORCE11（フォースイレブン）が2014年に発表した、データ引用に関する共同宣言があります。この共同宣言は、データ引用のあり方に関する原則をまとめたもので、多くの学術団体や研究者の支持を得ています。宣言は、「重要性」、「功績と帰属」、「証拠」、「ユニークな識別子」、「アクセス性」、「永続性」、「特殊性と検証可能性」、「相互運用性と柔軟性」という8つの原則から構成されています。

4.6.5 データ引用の方法

ここからは、実際にデータを引用する際の記述方法を学びます。

研究データの共有と活用を促進する活動を行う国際コンソーシアムDataCite（データサイト）は、引用情報に、少なくともデータの作成者、出版年、タイトル、出版者、タイプ、識別子の要素を含めることを推奨しています。これらに加えて、追加的な要素として、バージョン、データの種類を含めるとよい、としています。

なお、タイプとは、データセット、データベース、表、地図、音声ファイル、静止画、原稿、webサイト、動画等のような、引用対象である当該データの種類を指します。

また、識別子とは、対象データを確実に識別するアクセス番号や永続的ID（アイディー）のことで、代表的なものとしてDOI（ディーオーアイ）やハンドルが挙げられます。

このほか、分野やデータの特性に応じて要素を追加する場合もあります。

経済・社会科学系のデータへのアクセスを提供するUK Data Service（ユーケーデータサービス）は、データがカバーする地理的情報と時間的範囲も示すことを推奨しています。

生命科学分野では、時系列データを参照する場合に、シンプルにデータベース名とアクセス番号を記載するという方法も取られています。

また、継続的に更新される動的なデータベースの場合、データのダウンロード日を示すとよいでしょう。

4.6.6 データの引用例

実際の引用例を見てみましょう。データの引用は、このように各分野やジャーナルで定められている引用スタイルに従い記述されています。

文献管理ソフトを活用すると、データを含む引用文献の記載を効率よく進めることができます。

4.6.7 データの引用ツール

データの引用を支援するため、DOI Citation Formatter（ディーオーアイ・サイテーション・フォーマッター）というツールも作成されています。

データのDOI（ディーオーアイ）を入力すると、特定の引用書式に沿った形でデータが出力されるツールです。

DOI登録機関のうち、DataCite(データサイト）、Crossref（クロスレフ）、mEDRA（メドラ）、中文DOI（ディーオーアイ）の4機関が運用をしています。また、日本のDOI登録機関であるジャパン・リンクセンターも、近日中に運用を開始する予定です。

機械的に連携が可能な仕組みであるAPI（エーピーアイ）も提供されており、テキスト出力のほか、RDF-XML（アールディーエフ・エックスエムエル）、Bibtex（ビブテフ）、Citeproc JSON（サイトプロック・ジェイソン）等に対応しています。

4.7 データ管理計画の更新支援

研究中の支援として、最後に、研究前の段階で作成したデータ管理計画の更新支援について学びます。

4.7.1 DMPの実施とその検証

研究前に立てたデータ管理計画＝DMP（ディーエムピー）は、研究の初期段階で作成して終わり、という訳ではありません。計画どおりにデータ管理計画が実行できているか、何らかの問題が生じていないか、研究の進展やさまざまな要因により当初の計画を修正する必要が生じていないかなど、データ管理計画を基に折々で研究者の相談に乗り、機関としてどのような支援ができるかを検討することが望ましいでしょう。

4.7.2 DMPの更新

研究助成機関によっては、研究プロジェクト期間中にデータ管理計画を更新するよう明示的に求めている機関もあります。

例えば、欧州の助成プログラム”Horizon2020”（ホライズン・トゥエンティー・トゥエンティー）は、助成するプロジェクトにおけるデータ管理のガイドラインの中で、研究プロジェクト開始後6か月以内に最初のデータ管理計画を提出すること、プロジェクト期間中に重要な変更が生じた場合には必ずデータ管理計画を更新すること、プロジェクトの定期評価等に合わせて、あるいは遅くとも最終審査に合わせてデータ管理計画を更新することを求めています。

このような規定は、データ管理計画が、作成後も更新され続ける、生きた文書であることを示しています。

次のパートでは、具体的なデータ管理計画の更新方法について、説明していきます。

4.7.3 DMP作成ツールを用いたDMPの更新

助成機関によるデータ管理計画の更新についての要求に対応するため、DMP（ディーエムピー）作成ツールの中には、提出のタイミングに応じて、容易にデータ管理計画の作成・更新ができるような専用テンプレートを提供しているものもあります。

英国のDigital Curation Centre（デジタル・キュレーション・センター）のDMP作成ツール「DMP Online（ディーエムピーオンライン）」の例を見てみましょう。

先ほどDMPの更新を求める機関として例に挙げたHorizon2020(ホライズン・トゥエンティー・トゥエンティー）向けのページを見ると、「Initial DMP(イニシャル・ディーエムピー）」と表示された最初のDMP作成ツールのタブに加えて、「Detailed DMP（ディテールド・ディーエムピー）」「Final review DMP（ファイナル・レビュー・ディエムピー）」というように、段階に応じてDMPを更新するためのタブが設けられています。

研究データ管理支援者は、所属機関の研究者の助成取得状況を把握して、活用できるDMP作成・更新ツールがあれば、積極的に研究者へ情報提供するよう努めることが望ましいでしょう。

また、助成機関からの要求がなくても、機関としてDMPの更新を促したいという場合は、DMP作成ツールに機関専用のテンプレートを設定し、所属する研究者に利用してもらう、といった運用も考えられます。

第5章　研究後の支援

5.1 第5章の概要

第5章では、研究後の支援として、データ公開前の確認作業や、研究データの公開手順、メタデータやライセンスの付与について学びます。

まず、「研究データ公開前の確認作業」では、研究者や研究データ管理支援者が知っておくべき公開前の確認事項を学びます。たとえば、関係者からの許諾や、個人情報の匿名化処理が済んでいるか、データの粒度が再利用や引用に適した単位になっているか、ファイルフォーマットが保存に適しているか、また、データに関する問合せ先の管理責任者や、データ公開までの猶予期間の制約についても確認が必要です。

次に、「研究データの公開」では、データの保管や公開をおこなうデータリポジトリについて学びます。データリポジトリの種類には、特定分野のデータを集録する分野別リポジトリ、研究機関が所属する研究者に提供する機関リポジトリ、複数分野を取扱い、商業ベース又はプロジェクトベースでサービスを提供する汎用リポジトリがあります。データを公開するためのリポジトリを選ぶ際は、出版者や助成機関のポリシー、機関の研究データ保存ガイドライン、研究コミュニティへのインパクトなどを考慮しましょう。

「メタデータの付与」では、メタデータの概要と記載内容を学びます。メタデータとはデータの詳細や文脈を記述した「データに関するデータ」のことで、一定のフォーマットで記述することによって、システム間の相互運用性を向上させ、データセットの検索・引用を可能にします。

メタデータの標準スキーマには、一般的に用いられるDublin Core（ダブリンコア）、Data Cite（データサイト）に加え、日本で策定されたjunii2（ジュニーツー）の後継であるJPCOAR（ジェーピーコア）スキーマがあります。その他、分野特有のメタデータ標準も多数存在します。

メタデータの重要な項目には、DOI（ディー・オー・アイ）とORCID ID（オーキッド・アイディー）があります。

DOIは、電子データに付与される国際的な識別子です。DOIを付与することで引用の利便性が増します。

ORCIDは、研究者や研究貢献者に対し、国際的にユニークな16桁のIDを付与する団体です。ORCID IDを利用することで、研究者を一意に識別でき、また業績の管理を一元化・簡略化することができます。

「ライセンスの付与」では、コンテンツの利用許諾の条件として、ライセンスを明示する方法を学びます。メタデータにライセンスを記載することで、データを利用する際の条件が明確になります。クリエイティブ・コモンズ・ライセンスは、データの作成者が表示・非営利・改変禁止・継承の4条件を組み合わせて、利用条件を表現できます。リポジトリごとに利用できるライセンスは異なるため、確認が必要です。

それでは、次の動画からそれぞれについて詳しく学んでいきましょう。

5.2 研究データ公開前の確認作業

ここでは、研究データの公開に先立ち、研究データ管理支援者が理解しておくべき前提や確認事項について学びます。

5.2.1 研究データを公開することの意義

研究データの公開はさまざまな意義を持っています。

例えば、研究データを公開することによって、研究成果の透明性や公正性を確保したり、研究助成機関からの要求に応えることができます。

また、研究成果への直接的な引用の促進、新たな研究コミュニティの創出や産学連携への発展、またイノベーションの創出への貢献といった効果も考えられます。

他にも、公開や開示に備え、適切な管理・保存をすることで、データの再利用を促進したり、若手研究者や学生に対するデータ収集や解析に関する教育的な効果も期待されます。

さらに、市民や納税者の理解向上にも貢献することが考えられます。

5.2.2 研究データを非公開とする必要性

研究データの公開には多くの意義がある一方で、研究データを非公開にすべき場合もあります。どのような場合に非公開にすべきか、また、その際に確認すべき事項について学びます。

知的財産的な価値がある研究データに対しては、機関として戦略的に非公開を選択するべきです。機関の知財及びデータ保存等に関するポリシーの確認や、知財担当部署への照会などが必要となります。

複数の関係者が知的財産権を保持していたり、秘匿義務がある場合は、データ共有者間での取り決めに沿って、非公開にする必要があります。この場合には、データ共有者の所属機関のポリシーや共有者間での取り決めの確認が必要となります。

個人情報やプライバシー情報などのセンシティブな情報を含む研究データは、法的に公開が禁じられています。この場合にも、データ共有者間での取り決めを確認することが必要となります。

研究データ管理の支援者は、研究データを非公開にすべき場合や、その際に確認すべき事項についても、研究者に適宜情報提供を行うとよいでしょう。

5.2.3 公開前の確認事項

次に、研究データを公開する前に、確認すべき事項について学びます。

関係者からの許諾については、許諾を取得しているか、許諾内容が文書化されているか、またポリシーなどにより法的担保があるか否かを確認する必要があります。

個人情報への対応としては、匿名化処理が適切になされているかどうかも確認が必要です。総務省のガイドラインなどを適宜参照するとよいでしょう。

また、データの粒度が再利用や引用に適切な単位になっているか、ファイルフォーマットが保存に適しているか、また問合せ等に対応するデータの管理責任者が決定しているか、といったことも確認する必要があります。

さらに、データ公開までの猶予期間（エンバーゴ期間）の有無も確認が必要です。「論文出版まで研究データは非公開」と規定されている場合などもあるので、事前に確認をおこないましょう。

データ管理計画=DMP（ディーエムピー）が作成されている場合、これらの事項について記載されていることがあります。計画書を入手し、確認をするとよいでしょう。

この他にも、データを公開するに当たっての留意点はいくつかあります。研究データ管理の支援者は、公開に先立って確認すべき事項を把握し、研究者に対して適切なサポートができるように心がける必要があります。

5.3 研究データの公開

ここでは、研究データの公開を行う場所であるリポジトリについて、その選択方法や公開の手順について学びます。

5.3.1 データリポジトリとは

データリポジトリとは、一般に研究データやそれに付随するプログラム等を収集し、保管することを目的として設置されたリポジトリを指します。

研究データを適切なメタデータと共にリポジトリに保管し、そこで公開することで、オープンサイエンスにおけるデータの利活用や研究公正目的の適切なデータの保存といった多面的なニーズに対応することができます。

5.3.2 データリポジトリの種類

データリポジトリは、大きく3つの種類に分けることができます。

分野別のデータリポジトリは、特定の分野のデータを対象に収録するリポジトリです。データリポジトリの横断検索サイトre3data（リースリーデータ）から、さまざまな分野のリポジトリを探し出すことができます。

機関リポジトリは、大学等の研究機関が所属する研究者に対して提供するリポジトリで、論文などの文献情報の他、研究データも取り扱うことが可能となっています。

日本では、ほぼ全ての大学が機関リポジトリを保有しており、国立情報学研究所のJAIRO（ジャイロ）を使って、国内の機関リポジトリに蓄積されたコンテンツを統合的に検索することができます。

汎用リポジトリは、商業ベースまたはプロジェクトベースでサービスが提供されています。例えば、Zenodo（ゼノド）、Figshare（フィグシェア）、Mendeley Data（メンデレ―データ）などがあります。汎用リポジトリは、DOI（ディーオーアイ）登録機能のほか、リンクやタグの付与、SNS（エスエヌエス）サービスとの連携が充実している点に特徴があります。

このように、データリポジトリは対象分野や設置目的、機能に違いがあるため、データを取り扱う目的によって使い分ける必要があります。

5.3.3 出版者の動向

近年、学術雑誌の出版者は、その論文投稿ポリシーにおいて、論文の根拠データをデータリポジトリへ登録することを義務化するようになってきました。登録されたデータは、参照文献や謝辞などにその保存先を記載するよう求められています。また、出版者によっては、論文の根拠データの利用可能性も論文中に記載することを義務化している例もあります。

Nature（＝ネイチャー）とNature関連誌12誌では、2016年9月から、掲載受理された論文の基礎データについて、著者以外の者による利用可能性と、利用可能な場合はその利用方法を明示した「データ利用可能性ステートメント」を論文中に記載することを必須としました。

また、複数の学術雑誌を発行するAGU（エイジーユー）＝米国地球物理学連合は、2016年に発表した出版データポリシーで、作成者が基礎となるデータを利用可能にすることを望まない場合、または本データポリシーの遵守を拒否する場合、論文の出版を拒否する権利を留保すると規定し、同時に謝辞にデータの保存先を記載するよう求めています。

5.3.4 出版者の推奨するリポジトリ

いくつかの出版者は、推奨するデータリポジトリのリストを提供しています。多くは分野別のリポジトリが対象となっていますが、出版者自身がデータリポジトリを提供し、登録を推奨している例もあります。

Elsevierも（エルゼビア）は、同社の運営するリポジトリであるMendeley Data（メンデレーデータ）をデータの登録先として勧めています。

また、AGU（エージーユー）は自身のリポジトリを持っていませんが、データの公開先として推奨するリポジトリの一覧を公表しています。

オープンアクセス出版者であるPLOS（プロス）も推奨リポジトリを公表しています。

5.3.5 公開場所の選定の際に考慮すべき情報

ここで、データリポジトリを選定する際に考慮すべき点をまとめておきます。

出版者のポリシーによってデータリポジトリへの登録が義務付けられている場合は、それに従う必要があります。中には、編集者や査読者から具体的なリポジトリの指定がある場合もあります。また、助成機関などが、登録するリポジトリの条件を設けている場合もあります。データ管理計画が作成されている場合には、その記述も確認しましょう。

所属機関のポリシーの中で、データの公開先が指定されている場合には、それに従う必要があります。機関が管理する機関リポジトリをデータの公開先として活用することで、長期間の検索可能性や追跡可能性を担保することができます。

データ公開のインパクトを考慮することも大切です。出版者からの指定がない場合でも、できるだけ分野の研究者コミュニティによく知られたリポジトリを選ぶと良いでしょう。

なお、これらの基準から具体的なリポジトリを絞り込めない場合は、第三者による認証を受けている、出来るだけ安全性の高いリポジトリを選ぶべきです。

5.3.6 データリポジトリの認証

第三者によるデータリポジトリの認証基準には、大きく3つのレベルがあります。

コア・レベルは、DSA-WDS（ディーエスエー ダブルディーエス）のワーキンググループにより設定された認証です。エクステンディッド・レベルの認証は、ドイツ工業規格により設定された認証です。フォーマル・レベルの認証は、ISO（アイエスオー）による認証です。

3つの認証のレベルは、上下関係を構成したり、レベル間に優劣があるわけではなく、異なるニーズに応じた条件を有しています。

データリポジトリのディレクトリであるre3data（リースリーデータ）では、この認証の種類を指定してリポジトリを検索することもできます。

データリポジトリを選択する際の判断要素の一つとして覚えておきましょう。

5.3.7 リポジトリへの登録のステップ

実際に研究データをリポジトリに登録する際のステップは次のとおりとなります。

アカウントを作成し、登録するデータの粒度と数を決定します。

データと説明文書も準備します。

権利表示や許諾手続を行い、必要であればエンバーゴ期間も決定します。

これらのステップに加えて、登録先のリポジトリの投稿マニュアルの確認も事前にしておくようにしましょう。

エディンバラ大学では、大学の機関リポジトリの登録の手順や注意点を説明する詳細なチェックリストも公開されています。分野別のデータリポジトリに登録する際にも参考になります。

5.3.8 代表的なデータリポジトリでの公開例

データリポジトリでの研究データの公開事例を紹介します。

左側の汎用のリポジトリサービスであるFigshare（フィグシェア）では、研究データはこのように公開されています。

右側は、地球環境分野の代表的なリポジトリであるPANGAEA（パンゲア）での公開例です。

5.4 メタデータの付与

次に、リポジトリにデータを登録する際に必要となるメタデータの付与について学びます。メタデータは、データセットの検索や引用を可能にする重要な要素です。

5.4.1 メタデータとは

メタデータは、データに関するデータと定義されます。

メタデータとして記載するのは、データの詳細情報や文脈です。一般的に人の利用を想定した「文書」とは異なり、メタデータは、情報の機械的な処理を想定しています。

メタデータを一定のフォーマットに沿って記述することで、システム間の相互運用性も確保することができます。

また、適切なメタデータを付与することによって、データセットの検索や引用といったサービスの展開が可能になります。

データを非公開の状態で管理する場合にも、適切なメタデータを付与し、検索可能性や追跡可能性を担保しておくことで、作成者自身による再利用が簡単にできるようになります。また、研究公正の観点からも、非公開の場合にもメタデータを付与しておくことが重要です。

5.4.2 メタデータのタイプ

メタデータには大きく3つの種類があります。

記述的メタデータは、タイトル、著者、抄録といったデータセットの基礎的な内容を表します。利用者が検索によりデータへアクセスする際に重要な情報です。

管理的メタデータは、保存状態や権利情報、フォーマットなどの技術的な情報を表します。利用者が実際にデータを利用する際に重要な情報です。

構造的メタデータは、データ間の関係性に関する情報を表します。利用者がデータについて詳細に理解する際に重要な情報です。

5.4.3 さまざまなメタデータの標準

メタデータにはさまざまな標準スキーマが存在します。

例えば、分野を問わない一般的なものとしては、Dublin Core（ダブリンコア）やDataCite（データサイト）と呼ばれるメタデータスキーマを挙げることができます。

日本の機関リポジトリでは、Dublin Core（ダブリンコア）を発展させたjunii2（ジュニーツー）と呼ばれるスキーマが一般的に用いられています。2017年には、junii2（ジュニーツー）の後継となるJPCOAR schema（ジェーピーコアスキーマ）が策定されたため、今後はJPCOAR schema（ジェーピーコアスキーマ）が標準になると思われます。

分野特有のスキーマもさまざまなものが存在します。ここでは、社会科学分野と宇宙物理分野のスキーマの例を挙げています。

分野ごとの標準的なメタデータスキーマの種類については、Digital Curation Centre（デジタル・キュレーション・センター）のウェブサイトにまとめられているので、参考にしてください。

5.4.4 メタデータの入力支援

ここでは、研究データ管理の支援者によるメタデータの入力支援について学びます。

多くのリポジトリには、そのリポジトリが採用するメタデータスキーマが標準的に備わっています。

分野別リポジトリでは、多くの場合、分野毎のメタデータスキーマが採用されており、分野特有の詳細なメタデータが求められるため、メタデータ作成者用のマニュアルやチュートリアル、メタデータのエディターといったツールが提供されていることがあります。

先ほど紹介した社会科学分野のData Documentation Initiative（データ・ドキュメンテーション・イニシアチブ）や、宇宙物理学分野のSPASE Group（スペースグループ）も、このようなトレーニングやチュートリアルを提供しています。支援スタッフは、事前にこのような情報を収集し、研究者に案内できるとよいでしょう。

メタデータのうち、管理的メタデータについては、支援者が付与することができます。一方、記述的メタデータと構造的メタデータを付与するには、データセットに対する専門的な知識が必要となります。専門知識を有するデータ保有者による補助のもと、支援者がメタデータの入力支援を行うことも考えられるでしょう。

また、日本の機関リポジトリでは、junii2（ジュニーツー）やJPCOAR schema（ジェーピーコアスキーマ）による記述が求められます。大学図書館員が主体となって、データ入力や入力支援を行うとよいでしょう。

5.4.5 識別子の付与（DOI）

次に、管理用メタデータのうち、重要な要素である識別子について学びます。

DOI（ディーオーアイ）とは、コンテンツの電子データに付与される国際的な識別子です。

所定の様式でブラウザに入力すると、コンテンツの所在情報（URI：ユーアールアイ）に変換される仕組みを持っています。この仕組みを利用し、主に論文のリンク切れを防ぐために用いられています。DOIを研究データにも付与することで、論文同様に引用が可能となり、利便性が増すことが期待できます。

より詳しい内容については、参照にあげたハンドブックやガイドラインをご覧ください。

5.4.6 識別子の付与（ORCID）

管理用メタデータのもう一つの重要な要素として、人に付与する識別子であるORCID（オーキッド）があります。

ORCID（オーキッド）は、研究者や研究貢献者に対し、国際的にユニークな16桁のID（アイディー）を付与する取り組みを行う団体です。

ORCID ID（オーキッド　アイディー）を利用することで、研究者やその貢献者を一意に識別できるようになります。

これにより、業績の管理を一元化、簡略化することができます。このため、多くの出版者が、論文投稿時にORCID IDを提出することを義務化しています。

また、ORCID IDは、API（エーピーアイ）を通じて外部のシステムやサービスと連携が可能となります。日本では、JST（ジェーエスティー）のResearchmap（リサーチマップ）や、NII（エヌアイアイ）のKAKEN（カケン）との連携が実現しています。

5.4.7 研究データのメタデータの例

研究データのメタデータを具体的に見てみましょう。

これは生命科学分野の研究データを主に扱っているDryad（ドライアド）の検索結果です。

タイトル、作成者、公開日、DOI（ディーオーアイ）、抄録、キーワード等の情報のほか、研究データに特有な情報として地理的範囲、データ発行日、識別子、データと関連するコンテンツの情報、出版物の情報などのメタデータが付与されています。

5.4.8 研究データのメタデータ記述例（JPCOARスキーマに拠る）

こちらは、JPCOAR（ジェーピーコア）スキーマによるメタデータの記述例です。主要な情報に加えて、バージョンや、期間、地理的空間、また助成機関に関する情報も記載されています。

5.4.9 構造的メタデータによる連携

こちらは、構造的メタデータの活用事例です。

機関リポジトリのメタデータ入力において、構造的メタデータの入力が必須とされている事例は多くはありませんが、データの関連性の情報は、サービス向上に大変役立ちます。

構造的メタデータを活用したウッズホール海洋研究所のリポジトリ記述例を見てみましょう。論文とデータの関係が記載されていたり、データとデータの関係を記述しておくことで、論文の検証可能性の向上や、充実した検索サービスの展開に繋がります。

5.5 ライセンスの付与

ここからは、研究データへのライセンスの付与について学びます。ライセンスの付与は、研究データをデータ作成者が意図する形で適切に活用してもらうための重要なプロセスです。

5.5.1 研究データと著作権

まず、研究データと著作権の関係について確認しておきます。

日本では、著作権は著作権法により「思想又は感情を創作的に表現したものであって、文芸、学術、美術又は音楽の範囲に属するもの」と定義されます。研究データについては、判例において、データ自体は、集積行為に多額のコストをかけても、著作権法の保護対象にはならないとされており、単なる客観的事実やデータは、著作物とは見なされません。

但し、一定の考え方のもとにデータを整理・分析した場合には「創作性」が認められる場合もあります。例えば、地図やデータベースの著作物は、これに該当します。

5.5.2 ライセンシング

ライセンシング＝利用許諾とは、コンテンツの利用を認める際の条件、つまりライセンスを明示することを指します。ライセンシングにより対象コンテンツの利用条件が明確になり、法的な曖昧さを回避することで、コンテンツの再利用を促進することが期待できます。

公開された研究データにライセンスを付与することは、法的にどのような意味を持つのでしょうか。

一般的に、事実を収集したデータには創作性が認められないため、著作権法による保護がありません。したがって、特に断りがない場合は誰でも自由に使うことができます。

しかし、実際には著作物として認められるかどうかが微妙な場合もあり、データ利用者は、その判断に悩むことがあります。また異なる法的ルールをもつ外国のデータの利用の可否についても、曖昧な場合があります。

このような状態を解消するため、データ保持者がメタデータにライセンスを記載しておくことで、一律な利用が可能になります。

なおライセンスが付与されていない場合でも、その他の知的財産権によってデータが保護されている可能性があることにも、留意しておく必要があります。

5.5.3 ライセンス付与の原則とガイドライン

データのライセンシングに関する主な原則とガイドラインを示します。

データのあるべき姿を定めた原則である、FAIR（フェア）データ原則は、データの発見可能性、アクセス可能性、相互運用可能性、再利用可能性について言及しています。

また、研究データの法的相互運用性の向上を目的とした国際的なガイドラインが、2016年に発表されています。ガイドラインの中では、ここに示す6つの指針が示され、それぞれに実施のための具体的な手順が記述されています。

その他、このあと詳しく説明するクリエイティブ・コモンズ・ライセンスによるデータ・ライセンスに関するファクトシートなども公開されています。

これらの原則に加えて、データ管理計画に記載がある場合にはそれも参照しながら、付与すべきライセンスを決定するようにしましょう。

5.5.4 著者によるライセンス表示

データに付与するためのライセンスとして、Creative Commons License（クリエイティブ・コモンズ・ライセンス）とOpen Data Commons（オープン・データ・コモンズ）を紹介します。

クリエイティブ・コモンズ・ライセンスは、著作者が自らの著作物につき、4つの条件を組み合わせて利用条件を表現できるツールです。4つの条件は、作品のクレジットを表示することを意味する「表示」と、営利目的での利用をしないことを示す「非営利」、元の作品を改変しないことを示す「改変禁止」、元の作品と同じ組み合わせのライセンスで公開することを示す「継承」です。論文のライセンシングに広く使われています。

オープンデータコモンズは、事実情報由来のデータ及びデータベースにも適用可能なライセンスツールです。オープンデータの推進に役立ちます。

その他、ソフトウェアに特化したライセンスツールも公開されており、データに付随するソフトウェアについてライセンスを付与する場合に活用することができます。

データ登録先のリポジトリによって使用できるライセンスは異なりますので、事前に情報収集をすることが大切です。

ライセンスツールはさまざまなものが公開されていますが、細かい利用条件の設定は流通を阻害する可能性も高くなります。ライセンスを自由に選択できる場合には、ここに挙げたような、出来るだけ標準的なライセンスを付与することを推進していくとよいでしょう。

5.5.5 管理者側からの意思表示

ライセンシングは著作権者が行うことが原則となりますが、管理者が利用条件を確認できた場合に、それを示すためのツールもあります。

Creative Commons（クリエイティブコモンズ）によるパブリックドメインマークは、著作物を始めとする知的創作物について、知的財産権が発生していない状態を表示するマークです。何らかの理由で権利が発生しなかった場合、または権利が消滅した場合に用いることができます。ただし、権利が発生する要件、消滅する要件は国によって異なるため、ある国でパブリックドメインに置かれた著作物が、他の国では保護の対象となる場合もありますので、注意が必要です。

米国デジタル公共図書館と欧州の電子図書館、Europeana（ユーロピアーナ）は、オンライン文化資源の著作権のステータス交換を目的として、Rightsstatements（ライツステートメンツ）を創設しました。著作権のステータスを「In copyright(インコピーライト）＝著作権あり」、「No Copyright（ノーコピーライト）＝著作権なし」、「other（アザー）＝その他」に大別し、11種類のライセンスで表現できるようになっています。

5.5.6 リポジトリ毎に使えるライセンス例

ここでは、代表的なデータリポジトリ毎に使えるライセンスの例を紹介します。

Figshare（フィグシェア）、 Mendeley Data（メンデレーデータ）、Dryad（ドライアド）、PNAGAEA（パンゲア）などのリポジトリで使用できる、クリエイティブ・コモンズ・ライセンスを見てみましょう。それぞれのリポジトリによって使えるライセンスが異なることが分かります。

日本の研究機関が運営するリポジトリにおいても、それぞれ寄託規約や利用条件、データポリシーといった形で、リポジトリ毎に利用条件や、提供者が付与できる利用許諾の内容が規定されています。規定されている項目や、その内容はリポジトリによって異なります。

リポジトリの規定には、データ保持者にライセンシングの選択肢が広く認められている場合と、データリポジトリの定める利用規約に従わなければならない場合がある、という点に特に注意が必要です。

ライセンス付与の原則や法的な相互運用性に関するガイドラインを踏まえた上で、データをどのように使ってもらいたいか、という観点からライセンスを使い分けるとよいでしょう。

第6章　日常的な支援

6.1 第6章の概要

第6章では、日常的な支援サービスとして、研究データ管理研修の実施方法、ポータルサイトの設置概要と事例、相談窓口の設置概要と事例について学びます。

まず、「研究データ管理研修の実施方法」ではRDM支援者、研究者、大学院生それぞれへの研修方法を学びます。

支援者向け研修は、オンライン講座等による個人学習と、架空の研究者のプロフィールやストーリーを掲載した「シナリオ」に基づくワークショップ型の2種類を組み合わせましょう。

研究者向けの研修は、機密情報保護や研究データの保全、研究倫理などの所属機関のポリシーの文脈と、研究助成機関の要求や、研究データの種類、利用している研究支援ツールなど、研究者の研究活動の文脈の2つを踏まえて設計することが重要です。

大学院生向けの研修は、情報リテラシー教育の観点での位置づけが重要です。データを発見、保存、管理、再利用するためのスキルを総合した概念であるData Information Literacy（データ・インフォメーション・リテラシー）＝（DIL）ディー・アイ・エルの議論などは研修設計の参考になります。

次に、「ポータルサイトの設置」では、海外大学の事例を参照しながら、研究者が研究データ管理をおこなう上で最初に参照する場所であるポータルサイトに掲載すべき情報を学びます。ポータルサイトでは、RDMについての基本的な情報や、提供されているサービス内容、教育・研修ツール、ポリシーやガイドライン、リポジトリやデータカタログの情報、グッドプラクティスなどの情報を掲載し、研究データ管理とサービスへの理解を促進することが大切です。

そして、「相談窓口の設置」では、海外大学の事例も参照しながら、対面で個別の問い合わせを受け付ける相談窓口について学びます。実施体制のポイントは、全学体制で取り組みながら図書館員が相談窓口となり関連部署との連携を担い、機関の規模や提供サービスに応じた体制を備えることです。コンサルティングサービスを提供する際は、図書館ではデータライブラリアン、サブジェクトライブラリアン、リエゾンライブラリアンがスキルを活かして関与し、研究のライフサイクルの中で研究者の必要に応じたサービスを提供することが重要です。

最後に、「広報・アドボカシー」では、研究データ管理についての理解を促進する方法を学びます。具体的な方法として、ウェブサイトでの継続的な情報提供や、研究データ管理を簡潔に説明した資料の作成、ミニワークショップの開催や学部・学科の集まりでのブリーフィングの実施といった取り組みが挙げられます。

それでは、次の動画からそれぞれについて詳しく学んでいきましょう。

6.2 支援者向けの研修設計

はじめに、研究データ管理の支援者向けの研修設計について学びます。

6.2.1 データ管理支援者の研修に求められるもの

研究データ管理＝RDM（アールディーエム）の支援業務のあり方は、所属機関の環境や、構成員、制度のあり方に依存します。例えば、支援業務で扱うデータは、所属する研究者や学生の専門分野によって異なります。また、研究データの保全をどこまで支援できるかは、所属機関の備える情報システム基盤に左右されます。

このため、研究データ管理支援の研修は、組織の文脈を問わず、支援に必要な知識を習得するためのオンライン講座などの個人学習と、所属機関の環境の下でデータ管理の知識や技能をどう活かせるかを組織内で協議し、コンセンサスを形成するような組織内のワークショップ型研修の二種類を組み合わせることが望ましいでしょう。

次のパートでは、二種類の研修の具体的な内容を学びます。

6.2.2 オンライン学習による個人学習

個人学習はオンライン講座を受講することで行うことができます。

研究データ管理のオンライン学習教材は、世界中で開発されています。代表的なオンライン学習教材の例を紹介します。

支援者向けの学習教材「RDMRose（アールディーエム・ローズ）」は英国の非営利団体であるJISC（ジスク）による助成プロジェクトで開発されました。

英国のエディンバラ大学のオンラインチュートリアル「MANTRA（マントラ）」にはライブラリアン向けの自己研鑽用RDMトレーニングキットがあります。

国内ではオープンアクセスリポジトリ推進協会＝JPCOAR（ジェーピーコア）が開発した「RDMトレーニングツール」が公開されており、研究データ管理サービスを構築する際の足がかりを得ることができます。

6.2.3 組織内のワークショップ型研修

支援者向け研修のふたつ目は、ワークショップ型の研修です。

ワークショップ形式の研修を所属機関内で開催する場合には、シナリオベースの研修方式の導入が効果的です。

「シナリオベースの学習」とは、架空の研究者のプロフィールとストーリーを掲載した「シナリオ」に基づき、研究データ管理の具体的な課題や解決策、提供する支援サービスについてディスカッションを行う学習形式です。「シナリオベースの学習」を研修に導入することで、所属組織の文脈に沿った議論を活性化し、課題や解決策、提供すべき支援サービスについて組織内で共有することができます。

オーストラリアのグリフィス大学や米国のパーデュー大学のウェブサイトでは、こうしたワークショップ型研修の実践事例や、研修で使用した実際のシナリオが公開されているので参考にするとよいでしょう。

6.3 研究者向けの研修設計

次に、大学や研究機関に所属する研究者に対してどのような研修が必要となるかを学びます。

6.3.1 研究者向け研究データ管理研修に求められること

研究者に対する研修は、所属機関のポリシーと、研究者の研究活動の両方を踏まえる必要があります。

所属機関のポリシーについては、機関の機密情報保護の規則、研究データの保存のルール、研究倫理規定などを把握して、これらの規則に従った研究データ管理の方法を示すことが重要です。

研究活動の文脈の下では、研究助成機関によるデータ管理計画＝DMP（ディーエムピー）などの要求や、研究活動で扱われる研究データの形態、研究者が使用している研究支援ツールなどの実態を把握して、研修を設計する必要があります。

6.3.2 義務付けられた研修と自主的な研修

研究者に対する研修は、組織内の制度として義務付けられた研修なのか、研究者が自主的に参加する研修なのかによって設計内容が異なります。

ここからはそれぞれの研修を設計する際のポイントを学びます。

6.3.3 機関で義務付けられた研修：位置づけの明確化

機関や大学で義務付けられた研修では、研修の制度上の位置づけを明確にすることが重要です。

研究データ管理はそもそも、「研究公正の強化」と「オープンサイエンスの推進」の2つの役割を担っています。最初に、それぞれに関するポリシーを所属機関が制定しているかを確認する必要があります。

確認したポリシーをもとに、機関の研究データ管理の推進体制や、研究者に求める研究データ管理の要件を把握します。

その上で、研修がどこに位置づけられるかを踏まえて、セミナーやワークショップを企画することが重要です。

6.3.4 自主的に参加する研修：ワークシップの開催

研究者の自発的な参加を前提とした研修は、研究者それぞれの個別の状況に対応できるワークショップ形式が効果的です。

ワークショップの内容は、所属機関の持つ研究データ管理インフラによって変わります。例えば、所属機関が研究支援ツールやデータストレージを提供している場合は、それを利用した研究データ管理の方法を伝えるセミナーの開催が考えられます。所属機関がサービスを有していない場合には、外部の無料リポジトリや、無料の研究支援ツールに関するセミナーの開催が考えられます。

6.4 大学院生向けの研修設計

続いて、大学院生向けの研修の設計について学びます。

6.4.1 大学院生に対する研究データ管理教育に求められること

大学院生に対するデータ管理教育は、情報リテラシー教育の一部と位置付けて考えていく必要があります。

近年、情報リテラシーの中でも研究データに特化した概念として、 Data Information Literacy（データ・インフォメーション・リテラシー）＝（DIL）ディー・アイ・エルが提唱されています。DILは、データを発見、保存、管理、再利用するためのスキルを総合した概念であり、具体的に求められるスキルについてさまざまな議論がなされています。DILに関連する教材も開発されており、例えばここに示したDILガイドでは、図書館員によるDIL研修プログラムの設計方法に関する情報が提供されています。

こういったDIL教育で重要な点は、主に3つあります。一つ目は、「教員との連携をはかること」、二つ目は「年次や領域によって異なる大学院生のニーズを見極めて計画をすること」、三つ目は「学習の成果を継続的に評価すること」、があげられます。

6.4.2 カリキュラムへの導入

大学院生向けの研修の位置づけに続いて、実際に研究データ管理＝RDM（アールディーエム）を大学院生に習得してもらう方法を学びます。

方法の一つとして、大学院のカリキュラムへの導入があります。研究データ管理を大学院生のカリキュラムに導入することは、大学院生の研究スキル育成の観点からも重要です。

カリキュラムに研究データ管理研修を導入している事例として、オーストラリアのモナシュ大学博士課程における「専門的能力の開発」という授業の提供や、米国のマサチューセッツ医科大学における学生向けの共同データ管理カリキュラムの提供などがあります。

6.4.3 セミナー・ワークショップの開催

カリキュラムの導入の他に、セミナー・ワークショップを開催して研究データ管理を学んでもらう方法もあります。

研究データ管理の知識やスキルに対するニーズは、大学院生の研究状況によって異なります。大学院生がそれぞれのニーズに合ったセミナーやワークショップに参加することで、大学院生全体の研究データ管理能力の向上につながるでしょう。

セミナー・ワークショップの事例として、米国のイリノイ大学図書館の研究データサービスでは、データ管理の導入、簡単な研究データ管理、データドキュメンテーション、データワークフロー、データ共有などのワークショップを開催しています。

6.5 RDMポータルサイトの提供

ここからは、研究データ管理についての情報提供をおこなうポータルサイトの運営ついて学びます。

6.5.1 RDMポータルサイトとは

RDMポータルサイトとは、研究データ管理の基本情報や、研究者の責務、利用可能なサポートサービスや担当者の情報をまとめたウェブサイトです。

研究データ管理を担う研究者が最初に参照すべき場所であり、所属する機関や分野の要件に応じた効果的な研究データ管理を支援します。最新情報を掲載することで、研究者に限らず、さまざまなステークホルダーに対する広報の役割を果たすこともできます。

ポータルサイトに掲載すべき情報は、クレジットを表記すれば利用・改変が可能であることを意味する「CCBY（シー・シー・バイ）」ライセンスを付与して公開されているものもあります。新たにポータルサイトを開設する際や、ポータルサイトのコンテンツを拡充する際には、こういった公開情報を利用するとよいでしょう。

6.5.2 RDMポータルサイトの情報

続いて、ポータルサイトに、具体的にどのような情報を盛り込む必要があるのかを学びます。

まずは、研究データの基本情報として、研究データ管理の基本的な流れ、オープンサイエンスや研究公正といった「研究データ管理が求められる背景」や、資源の共有やイノベーションの促進といった「研究データ管理の目的」、研究の効率化や、研究者や機関の評価への貢献のような「研究データ管理のメリット」を提示するとよいでしょう。

次に、研究ライフサイクルの中で研究者が実施するデータの管理・保存・文書化・共有・公開等の実務について、機関や研究データ管理支援の担当部署が、どのような支援サービスを実施しているのか、といった情報も重要となります。

関係者が研究データ管理の理解を深められるように、さまざまな機関が作成する、研究データ管理に関する、教育・研修ツールの情報も掲載するとよいでしょう。

助成団体や出版者のポリシー・ガイドライン、また国の指針等は研究者や機関が実施すべき研究データ管理の指針となるので重要です。

適切な研究データ管理の結果、研究データの所在が明確になることが重要です。メタデータを登録しているデータカタログや、研究データを保存しているデータリポジトリのリストを公開することで、研究者が自身の要件にあった適切なインフラを選択することができます。

研究データ管理を先行的に実践し、ノウハウを蓄積しているさまざまな機関が存在します。より具体的に研究データ管理をイメージしてもらうための、このような機関グッドプラクティスや自機関の研究者の優れた取組みを共有することも有益な情報となります。

6.5.3 海外大学の事例

ここでは、世界の４つの大学のRDMポータルの事例を学びます。

大学の規模や置かれている文脈が異なるので、それぞれの事例を見ることで自身の機関に合ったポータルサイトのあり方が見えてくると思います。

6.5.4 例1 エディンバラ大学

エディンバラ大学のポータルサイト「研究データサービス」のページでは、大学が提供するサービスの概要を紹介しています。

「About（アバウト）」のページでは、情報部門の多様な専門家で構成する研究データサービスが、研究者や学生を支援するために、ツール、サポートおよび研修を提供し、データに関連する要求のすべてに対し、ワンストップでサービスを提供することを強調しています。

また、「The Researcher’s journey（ザ・リサーチャーズ・ジャーニー）」では、研究プロジェクトを研究前、研究中、研究後に大別し、それぞれの段階でどのようなツールやサービスが利用可能かを示しています。例えば、研究前にデータ管理計画を作成するにはDMPOnline（ディーエムピーオンライン）を利用すること、研究中にデータを検索するには図書館のコンサルティングサービスを利用すること、研究後にデータの発見性を高めるには、データカタログに登録することなどが挙げられています。

ここでは研究データ管理についての説明に加えて、大学の研究データ管理ポリシーの周知も行なっています。

6.5.4 例1 エディンバラ大学

「研究データサービス」のページでは、研究データ管理に関する基本情報および関連するサービスが紹介されています。

研究前のステップに該当する「Research data management planning（リサーチデータマネジメント・プラニング）」では、データ管理計画のメリット、記述方法とツールの紹介、助成団体による要件について説明しています。過去に助成を獲得したデータ管理計画のサンプルも請求できるようになっています。

研究中のステップに該当する「Working with research data（ワーキング・ウィズ・リサーチデータ）」では、研究データの発見、利用、保存、機微情報の扱い、バージョン管理、共同研究者とのデータ共有、電子ノートブックの利用について、その概要と利用できるツールや提供しているサービスと合わせて説明しています。データリポジトリのリストも掲載しています。

研究後のステップに該当する「Sharing & preserving research data（シェアリング・アンド・プリザービング・リサーチデータ）」では、エディンバラ大学のデータリポジトリであるDataShare（データシェア）、同じく長期保存ストレージのDataVault（データボールト）、データカタログのPure（ピュア―）といったツールの利用方法を説明しています。

6.5.4 例1 エディンバラ大学

研修情報のページには、研究データ管理と共有に関するMOOC（ムーク）や、オンライントレーニングツールMANTRA（マントラ）へのリンク、対面式のワークショップやセミナーの案内などの情報がまとめられています。

また研究データ管理サービスを周知するためにサービスの概要を紹介する動画も提供しています。

エディンバラ大学のポータルサイトの特徴としては、

・研究データ管理を研究前、研究中、研究後に大きく分類し、各段階で実施すべき手順について解説を行った上で、利用できるサービスやツールを紹介していること、

・ワークショップや学習ツールが項目として独立しており充実していること、

・担当部署の情報や相談窓口がサービスごとに示されており、わかりやすいこと、

などの点が挙げられます。

6.5.5 例2 イリノイ大学

イリノイ大学のポータルサイトは、図書館のウェブサイトの中に「研究データサ－ビス」として配置されています。

「How We Can Help（ハウ・ウィー・キャン・ヘルプ）」のページでは、研究データ管理に関する基本情報と、提供するサービスの全体像を紹介しています。

ここから、ワークショップや、広報誌Data Nudge（データナッジ）、データ管理計画、データの整理、データの保存・共有といった手続きの詳細にリンクすることができます。例えば、データ管理計画では、どのような計画を作成すればよいのか、ベストプラクティスを示すとともに、利用できるツールとしてDMPTool（ディーエムピーツール）を紹介しています。また、保存と共有では、共同研究者との間で安全にデータを共有できるツールや、さまざまなデータリポジトリの情報に加えて、長期保存のためのポリシーも提供しています。

データリポジトリ「Illinoi Data Bank（イリノイデータバンク）」や、データストレージの情報には、メインページからアクセスすることもできます。

イリノイ大学のポータルサイトの特徴は、研究データ管理のワークショップ、データ管理計画、データの組織化、データの保存といった重点的なサービスを軸に情報を提供している点です。担当部署の情報は、サービスごとではなく独立した項目として提供されています。

6.5.6 例3 モナシュ大学

モナシュ大学のポータルサイトでは、まず、大学における研究データ管理の位置づけや理念、ポリシーとガバナンス、達成状況や運用体制を紹介しています。

そのうえで、「ガイドライン」として、研究ライフサイクルで生じるデータ管理の課題に対応するために必要な情報を提供しています。データ計画、所有権、研究倫理、長期保存、フォーマット、ストレージ、セキュリティ、データの組織化、共有、再利用などが含まれます。例えば、データ計画については全ての研究者に作成することを推奨し、モナシュ大学の研究データ管理ポリシーを踏まえたチェックリストを提供しています。

さらに、研究データ管理に関する研修情報として、学生向けと教員向けに分けて、セミナーやワークショップなどの紹介もしています。

他にも、機関のデータリポジトリや、電子ラボノートなどの情報提供、相談窓口、データの保護や破棄、著作権などさまざまな情報が集約されています。

モナシュ大学のポータルサイトの特徴は、データ管理計画からデータの保存・共有にいたるまで、研究データ管理に係る代表的な手続きごとに情報を整理していることです。各ページは手続きの概要や提供しているサービスに加えて、関連する情報や学内の問合せ先を表示しています。

6.5.7 例4 ワーゲニンゲン大学

オランダのワーゲニンゲン大学のポータルサイトも、研究前、研究中、研究後、それぞれにおけるデータ管理の情報と、提供するサービスの内容を紹介しています。

研究前は、データ管理計画について説明し、ワーゲニンゲン大学や助成機関の要求内容を説明しています。

研究中では、データのドキュメンテーション、機密保護、データの共有やファイルフォーマット、ソースコード管理、ファイルやフォルダの管理、データストレージなどの情報を提供しています。

研究後では、データのライセンス、出版社のポリシー、データセットの公開やリポジトリの活用について説明しています。

ワーゲニンゲン大学の特徴としては、サービスのベースとなる、大学としてのポリシーについてのページも充実している点です。ポリシーの概要を示すとともに、ポリシーに沿って、各サービス内容が分かるように情報を整理しています。

ポータルサイトの右上には常に連絡窓口が表示され、担当者に簡単に連絡をすることができます。

事例に挙げた機関のポータルサイトに掲載されている情報の主な共通点としては、

・RDM（アールディーエム）に関する基本情報

・提供しているサービスの内容と担当部署の情報

・RDM（アールディーエム）に関する教育・研修ツール

・データリポジトリのリスト

・グッドプラクティス

が上げられます。

これらの情報を各機関の置かれた状況に合わせて整理して提供しています。

6.5.8 ナショナルサービス

ポータルサイトは、個別大学により設置されたものだけではなく、国レベルで設置された事例もあります。

英国のDigital Curation Center（デジタル・キュレーション・センター）、オーストラリアのAustralia National Data Service(オーストラリア・ナショナル・データ・サービス)、オランダのResearch Data Netherlands（リサーチ・データ・ネザランド）などが、国レベルで情報を提供しています。

ナショナルポータルサイトに掲載されている国内共通の情報を活用すれば、機関のポータルサイトは、機関独自の情報発信に注力することができます。ナショナルサービスの情報も、CCBY（シーシーバイ）で公開されていることが多いので、機関でポータルサイトを設置する際には参考にするとよいでしょう。

6.6 RDM相談窓口の開設

続いて、研究データ管理の相談窓口の開設について学びます。

6.6.1 RDM相談窓口とは

研修やポータルサイトの設置に加えて、日常的な支援サービスを構成するのが相談窓口におけるコンサルティングサービスです。

相談窓口とは、各機関の支援サービスの内容や研究者のニーズに合わせた研究データ管理のコンサルティングサービスを意味します。研究者がそれぞれの状況に応じて、具体的な相談ができる場で、重要なサービスのひとつです。

ここでは、コンサルティングサービスの内容や、実際のサービスの提供事例を学びます。

6.6.2 コンサルティングサービスの実施内容

実際のコンサルティングサービスは、データライブラリアンやリエゾンライブラリアン等のサービス提供部署による、研究データ管理への相談対応です。

これらのサービスは研究のプロセスに沿って提供され、研究者毎のニーズに沿ってカスタマイズされた形でのコンサルティングや支援をおこなうことが想定されます。

例えば、メタデータの作成や、データの準備及びリポジトリへの登録といったサービスがあります。

相談窓口の担当者の情報は、研究者が発見しやすいようにポータルサイトに掲載しましょう。

6.6.3 コンサルティングサービスの体制

コンサルティングサービスの体制づくりには大きく二つのポイントがあります。

一つ目のポイントは、サービスを図書館単独ではなく、学内の関連部署と連携して提供するという点です。図書館がサービス窓口となって、データライブラリアンが関わりながら、連携部署を紹介することで、より円滑な研究データ管理の支援をおこなうことが可能となります。

また二つ目のポイントとして、コンサルティングサービスをオンラインヘルプデスクや対面式で提供することが挙げられます。研究者が抱える具体的で個別な相談に、直接対応できるサービス提供が重要です。

6.6.3 コンサルティングサービスの体制

コンサルティングサービスの提供では、組織の体制に加えて、対応する職員のスキルも重要となります。

相談窓口でコンサルティングサービスを提供することが想定される「データライブラリアン」のスキル要件について、オーストラリア国立データサービスの掲げるスキル要件を見ていきましょう。

データライブラリアンには、「データ管理」、「メタデータ管理」、「データの利用」、「チェックリストや教材の作成」、「組織内外の情報源の紹介」といった、さまざまなスキルが求められます。

6.6.4 コンサルティングサービスの先行事例

ここでは、世界の４つの大学のコンサルティングサービスの事例を、サービス提供チームの構成と、提供するコンサルティングサービスに着目して学びます。

6.6.5 先行大学の事例 エディンバラ大学

エディンバラ大学は、RDMサービスをInformation Service Group（インフォメーションサービスグループ）が提供しています。同グループは図書館、情報技術、学習・教育技術、デジタル技術の研究センター（EDINA（エディナ））等で構成され、サービス母体は「Research Data Service（リサーチ・データ・サービス）」として一つに統合されています。

また、研究データ管理の向上を目的とするナショナルサービスであるDigital Curation Center（デジタル・キュレーション・センター）の本部がおかれています。

エディンバラ大学ではさまざまなサービスが提供されており、その中には研究データインフラ構築やサービス提供におけるアドバイス、データ検索・アクセス・使用・管理・分析の支援、各学部に所属する学術支援図書館員への相談などのサービスがあります。

電子メールサービスを活用しており、特定のニーズについて問合せ、相談することを奨励しています。

これらに加えて、本部を置くナショナルサービスのデジタル・キュレーション・センターでは、研究データ管理の支援者向けの相談対応をおこなっています。

6.6.6 先行大学の事例 イリノイ大学

イリノイ大学では、RDMサービスはResearch Data Service(リサーチ・データ・サービス）が提供しています。研究担当副学長室、学部長室、情報学科技術サービス、スパコン応用センターの共同運営で、図書館内に拠点を置き、ディレクター1名、プログラマ1名、キュレーター2名の4FTE（エフティーイー）相当で運用しています。

研究サイクルの全てのニーズを網羅し、「サービス範囲は無制限」としています。分野の専門性を持つ「Subject librarian（サブジェクトライブラリアン）」が組み込まれ、専門分野の視点から補足しています。

専門性を背景に提供しているサービスとして、データ管理計画＝DMP（ディー・エム・ピー）のドラフト版をチェックしてフィードバックしたり、適切なデータリポジトリの選択、データセットの整理、データの保存などについても支援を提供しています。

問合せ用のメールアドレスからサービスにアクセスすることも可能です。

6.6.7 先行大学の事例 モナシュ大学

モナシュ大学では、図書館が中心となって、学内のMonash eResearch Centre（モナシュ・イーリサーチセンター）、 Monash eSolutions（モナシュ・イーソリューションズ）、Monash Research Office（モナシュ・リサーチオフィス）と連携してRDMサービスを提供しています。

学内の専門知識を活かした分散型のアプローチをとっており、例えば、データの保存等の技術的な問合せはeResearch Center（イーリサーチセンター）に、著作権については図書館に、といったかたちで対応しています。

また、Subject librarian（サブジェクトライブラリアン）と研究インフラチームが協力して、ポリシー、ガイドライン、リポジトリ識別についてのガイダンスを提供しています。

さらに、コンサルティングサービスの入り口として、研究データ管理専用の問合せメールアドレスを用意しています。

6.6.8 先行大学の事例 ワーゲニンゲン大学

ワーゲニンゲン大学では、RDMサービスは図書館の上位にある研究基盤の一部に位置づけられています。サービスを提供するデータ管理サポートグループは各部局の専門家で構成しており、図書館が取りまとめ役をつとめ、ITサービス、文書管理、コーポレートガバナンス、法務サービス等によって構成されています。

提供されるサービスは、データ管理計画のアドバイス、データセットのドキュメント作成、メタデータ作成、外部リポジトリへの登録支援などで、RDMプロセスに関係する全ての課題について、スタッフの支援、アドバイスを求めることを推奨しています。

また、研究者のニーズを満たすために外部のソリューションも活用しており、例えば、ナショナルレベルのアーカイブであるDANS-EASY（ダンズ・イージー）やコンソーシアムレベルのデータリポジトリである4TU Repositories（フォーティーユー・リポジトリーズ）へのアクセスの仲介や、サービスの調整にも応じています。

ウェブサイトに常に表示されているコンタクトフォームからメールで問い合わせができます。

6.6.9 事例のポイント

最後に、これまで見てきた各大学の相談窓口サービスの事例から分かるポイントをまとめます。

相談窓口は全学体制で実施されており、それぞれのスキルをもった部署・人材が研究者を支援しています。図書館が相談の窓口となり担当各部署を仲介するケースが多くあり、データライブラリアン、サブジェクトライブラリアン、リエゾンライブラリアンといった専門性を持った図書館職員が深く関わっています。

提供するサービスは、機関の規模や置かれている状況に応じて異なり、それに対応してサービスの体制も異なっています。

サービス内容は、研究ライフサイクルに沿って、例えばDMP（ディーエムピー）の作成支援、適切なデータリポジトリの選択、デートセットの整理、メタデータの作成など、研究者の必要に応じてサービスを提供していることがわかります。

国レベルで研究データ管理に取り組むナショナルサービスにおいて、支援者の相談に応じるサービスも提供している例があります。

6.7 広報・アドボカシー

最後に、研究データ管理の必要性を周知し、理解を促すための「広報活動」や、「アドボカシー」について学びます。

6.7.1 広報・アドボカシーの重要性

研究データ管理＝RDM（アールディーエム）の必要性や重要性は、現状では、機関内に広く認知されているとは言えません。そのため、研究者や、機関のスタッフ、機関の経営層といったステークホルダーに対して、研究データ管理の意義や概要を伝える広報・アドボカシーの取り組みが重要です。

なお、これまで説明した支援サービスの中に、広報やアドボカシーの意味をもつサービスもありますが、ここでは、まだ説明をしていない取り組みを中心にとりあげます。

広報活動では、そもそも研究データ管理とは何か、また機関や研究助成機関、出版者のポリシーが何を求めているのか、そして研究データ管理支援サービスの存在の周知をおこないます。

アドボカシーでは、広報からさらに一歩踏み込み、研究コミュニティが研究データ管理について抱く疑問や懸念の解決に努めたり、連携・協力のため関連部署とのコミュニケーションの実施、またサービス実施に当たり不可欠となる予算要求の説明を経営層に対して行う、といったことが挙げられます。

6.7.2 広報・アドボカシーの方法

広報・アドボカシーの主要な実施内容について学びます。

まず、ウェブサイト上での継続的な情報提供です。先ほどポータルサイトの説明で述べたように、研究データ管理及びサービスに関する主要な項目について、ウェブ上で情報提供を行うことは、ステークホルダーへの広報やアドボカシーの役割も果たします。

広報・アドボカシー活動のための資料の作成も大切です。研究データ管理サービスについて説明したブリーフィングペーパーや、研究データ管理の基本情報や連絡先等を簡潔に説明したリーフレット、また学内等に掲示するポスターを作成することで、基本情報を迅速かつ簡単に提供することができます。

実際の資料の例として、英国のラフボロー大学で作成したリーフレットを見てみましょう。利用できる外部リソース、相談窓口のリスト、研究データ管理の概要がシンプルにまとめられています。

ミニワークショップの開催も広報・アドボカシーに有効です。ミニワークショップにはさまざまな形式がありますが、例えば新任スタッフ向けのオーダーメイドのワークショップなどが考えられます。

ワークショップの他にも、学部や学科のスタッフや研究者が集まる場でのブリーフィングを行い、研究データ管理や提供するサービスについての最新情報などを共有することも重要です。関係者に幅広く情報共有をしておくことで、理解を促進することはもちろん、必要な場合に、よりスムーズに研究者を支援サービスへと導くことができるでしょう。

6.7.3 広報・アドボカシーの事例: University College London（UCL）

実際の機関における広報・アドボカシーの事例を紹介します。

University College London（ユニバーシティカレッジロンドン）＝UCL（ユーシーエル）では、図書館のリエゾン・サポートサービス担当者2名が、「研究データ支援オフィサー」として広報・アドボカシーを担っています。

具体的な取り組みとしては、学部・学科の集まりにおいて、スタッフ向けに短いプレゼンテーションを行ったり、新任教職員のためのオリエンテーションにおいて、1時間から3時間程度のオーダーメイドのワークショップを提供しています。また、大学全体の研究データ管理の状況を把握するため、調査も実施しています。

さらに、オープンアクセス部門のライブラリアン向けにワークショップを開催し、研究の前段階から連携して研究データ管理への周知を促進しています。また、研究ITサービス部門と共同で、研究者向けのセミナーを定期的に開き、研究者へのさらなる周知を図っています。

このように広報・アドボカシーにおいては、担当者のみで行う活動に加えて、機関の状況に応じて他部門と連携しておこなう取り組みも有効です。