

センターの概要

研究資源センター（RRC）は、放影研のインフラの中核的な構成要素となる。放影研の戦略計画を推進するためには、RRC を上手く実現し運営することが不可欠である。

RRC には以下の3つの使命がある：

1. 放影研の研究資源を保護し、索引付けし、統合する。これには、データ、バイオサンプル目録、紙の記録、人工物、原稿、データセット、プログラミングスクリプト、その他歴史的に重要な論説などが含まれる。研究データへのアクセスは、対象者のプライバシーを保護するための明確なアクセス基準を備えた各種クラウド環境を介して行われる。
2. 全てのデータとバイオサンプルの目録とを統合することにより、放影研の研究遂行能力を高める。データ可視化、データ整理および分析のためのツールにより、アクセス手順を簡略化、標準化し、研究を促進する。
3. 所内の研究および契約や外部資金を含め共同研究プロジェクトを促進するための管理体制を提供する。

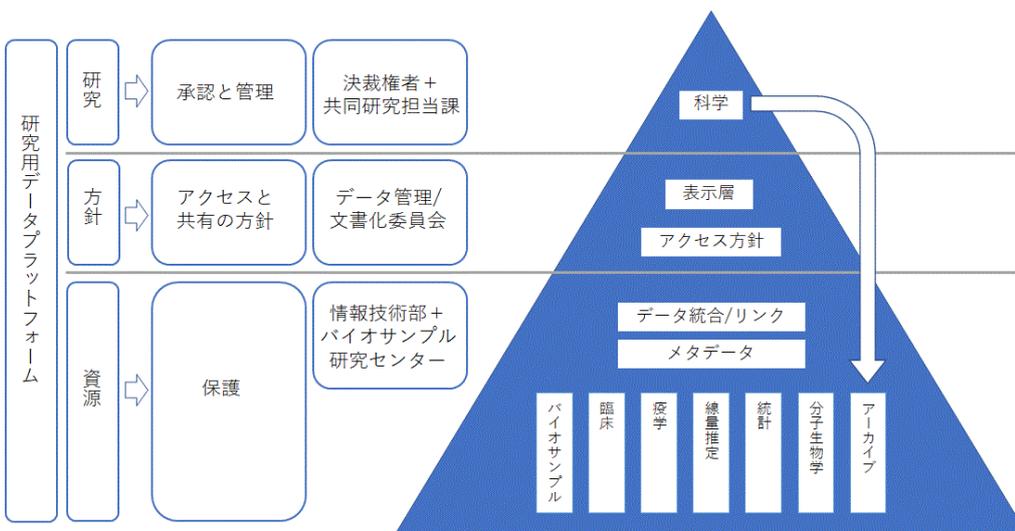


図1. 放影研の研究資源センターの機能的構造

図1を見ると、ピラミッドの最下層は保護層である。保護層には、デジタルデータの保存、統合、可視化のための枠組みが含まれる。デジタルデータ以外にも、この層には放影研資料のデジタルスキャンした記録や歴史的に重要な人工物の目録を全てカタログ化し索引付けするためのコンテンツ管理システム（CMS）が含まれる。データのコピーやスキャン画像は一元的に保管される（使命1）。全ての資料には各資料を説明する情報（メタデータ）が付くことにより、見つけられ、関連資料と適切に紐付けられ、検索可能になる。第2層は、「方針とアクセス」の層である。所内規則と最新の認証方法を用いて、権限を与えられた資料のみが閲覧可能となる。この層には、所内外の研究者が放影研のデータカタログを閲覧できるようなプレゼンテーションツールが含まれる。権限を付与されたユーザーは、これらのデータに基づく研究計画書や論文原稿を閲覧し、新たな研究を計画するためにデータを可視化し検索することができる（使命2）。最上層は研究層である。ここでは、承認された研究を実施できる。共同研究を行う場合は、「研究支援

室（ORS）」がデータを提供し、バイオサンプルの提供依頼を処理するための契約を履行するとともに手続きを支援し、研究を円滑に進める（使命3）。研究終了時には、当該研究に関連する全ての資料（データ、分析結果、解析スクリプト、論文原稿など）を適切なメタデータを付与しリンクを立て、アーカイブに保存し、将来の研究者がそれらの成果物を再利用できるようにする。

RRCを設立・管理するための人員や機能別ユニットは、研究資源センター運営委員会が暫定的に設定した。この臨時委員会は2021年11月に発足、2021年12月に第1回会合が開催され、2022年5月に第2回会合が開催された。小野ITD部長が委員長および田邊主席研究員が共同委員長として当該委員会を主導し、部長全員が委員を務める。前委員長のGrant主席研究員は2022年12月に退職した。この委員会は、2019年12月に設置された「研究資源センター設置準備委員会」という以前の臨時委員会に取って代わるものである。RRCの機能単位と指揮命令系統については図2を参照のこと。今年度よりRRCの組織体制を若干変更した。従来、事務局内にORSが設置されていたが、同室の役割を検討した結果、研究資源課に統合されることになった。

3つの主要な開発ニーズ

図1に示すような機能を持つRRCを実現するためには、大きく分けて3つの領域の開発を完了する必要がある。この3つの領域とは、「データフレームワーク」、「コンテンツ管理システム」、「研究支援室（ORS）」である。この機能と開発領域については、放影研の「技術チーム」が2017年以降隔週で会議を開催し綿密に検討してきた。このチームには、ITDの幹部職員とGrant主席研究員が含まれていた。現在、この機能は研究資源課に移管した。

データ管理フレームワーク：「データ管理フレームワーク」は放影研データに関して基盤となるべきコンピューター環境である。データの保存、アノテーション、統合が可能でなければならない。FAIRの原則に従って新しいデータを容易にシステムに追加できるよう、フレキシブルなものが理想的である。（コンピューターに詳しい管理担当者にこの作業を強いるのではなく）データを作成した研究員がアノテーションできれば理想的である。このシステムは日本語と英語の両方に対応し、ローカルサーバーまたはクラウド環境にインストールすることが可能で、データが研究所を「離れる」ことがないよう研究者がリモートで解析を実施できるならば、理想的である。また、認証された人だけがデータにアクセスできるよう、階層的で安全性の高い認証システムが必要である。またデータを可視化するとともにデータ探索のために検索できる、「データ可視化アプリケーション」が基礎データと相互作用できるようにする。

このようなシステムに関するパイロットプロジェクト（シカゴ大学が開発したGen3 データcommons）を2020年から2021年に1年以上かけて実施した。残念ながら、本章で後述する様々な理由により、このパイロットプロジェクトは成功しなかった。放影研は現在、適切なデータ管理フレームワークを模索している。現在、米国に本社を置くグループ企業であるSeven Bridges社のフレームワークを候補として「概念実証」による機能・性能の評価を計画している。

昨年度からデータ可視化ソフトウェアを用いたパイロットプロジェクトを実施しており、疫学部の杉山副部長から研究所内の実際のデータセットを使用した極めて実践的な可視化の例が提供された。そこで今年度、疫学部用にTableauのライセンス4つを追加購入した。Tableauを研究所全体で活用するには、より多くの時間と予算が必要である。

コンテンツ管理システム：コンテンツ管理システム（CMS）は、放影研内の全ての紙ベース、また

はテキストベースの資産を保存するシステムである。放影研の認可アクセス方針を厳守しながら、放影研資料へのアクセスを可能にするフレキシブルで頑健なセキュリティシステムが必要である。

より一般的な資料（出版物リストなど）については、現在、放影研には様々なリスト（研究計画書、業績報告書など）を閲覧するための多数のデータベースが存在するが、それらはしばしば独立したデータベースであり（すなわち、複数のサイトを検索して調べる必要がある）、放影研内のどこにも全文検索機能は存在しない。これらの問題（一元的なアクセスや全文検索が不可能）は、放影研の現在のコンピューター環境における大きな欠点である。放影研内の紙ベースの記録、放影研の全てのPDFファイル、検査記録のノート、質問票などをスキャンして保存できるCMSが、放影研には必要である。このシステムは日英に対応し、直感的な操作が可能であり、保存と再利用のために個人がローカルで保有された資産をアップロードできるものでなくてはならない。

RRCはCMSを特定し、当該CMSを専門とする外部業者を起用して、パイロットプロジェクトを策定した。すでに試験的な運用は終了しており、現在、システムの競争入札に向けて仕様書を作成中である。

研究支援室：放影研の研究活動は、管理面の障壁によって不必要に遅れが生じている。全ての業務が紙ベースで行われ、他の承認書と重複または矛盾する承認書も多い。さらに、外部の研究者が放影研の保存資料を閲覧したり、共同研究プロジェクトを検討したりするための打合せを始めるための「プレゼンス」情報が、放影研の外部ホームページにはない。研究支援室（ORS）は、これらの問題を一つ一つ解決することを目指す。放影研の申請書を徹底的に検討し、重複を避けるために書式を作り直すことを想定している。さらに、プロセス全体をオンラインシステムに移行すべきである。研究所内の意思決定全般を網羅する電子ワークフローシステムの開発・導入を計画しており、このシステムの導入により研究所内のあらゆる手続きが電子化・可視化される。

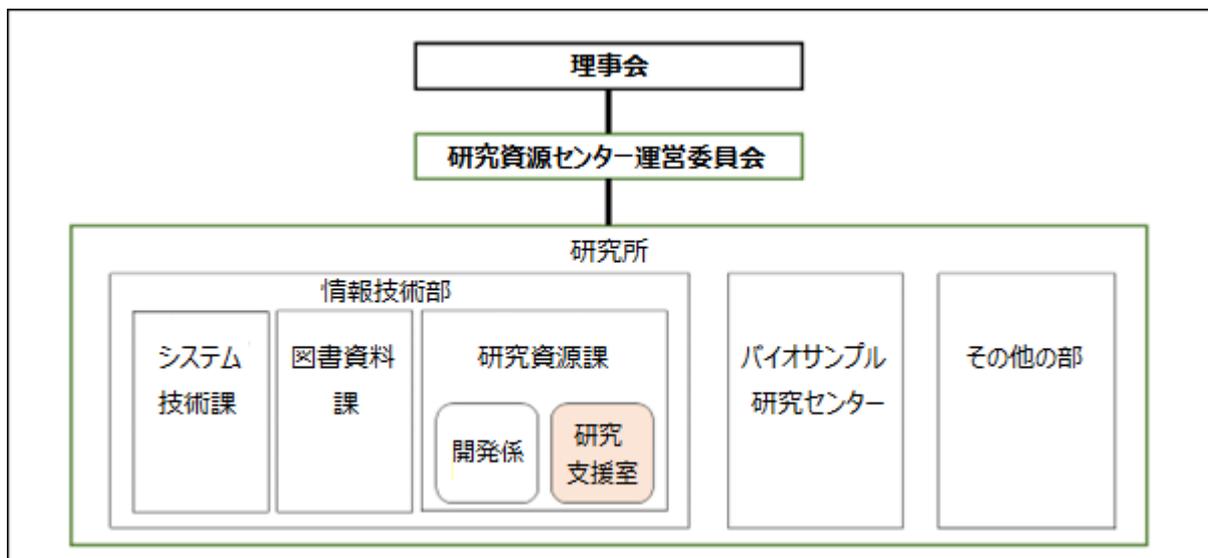


図 2. RRC の構造。これは、放影研の既存の部内に課が設置される 2 段階ステップの最初のステップである。第 2 のステップは、専任の人員と物理的なオフィスを備えた「独立した」センターとして RRC を設立することである。

2022年度業績

2022年の研究資源センターの主要な業績は、管理体制の再整備およびマイクロフィルムのデジタル化と染色体のネガフィルムのデジタル化計画であった。

当初、RRCに関する計画は、2019年12月に設立され、丹羽理事長およびUllrich副理事長が共同委員長を務めた「研究資源センター設置準備委員会」が主導していた。この臨時委員会は、全部門に情報を提供し、支持を取り付けるために設立された。この臨時委員会の中で、「ガイドライン起草」小委員会は、報告系統、リーダーシップ、予算、人員、RRCの内部構造など、放影研におけるRRCの位置づけを決定する役割を担った。第5回会議は2021年に開催され、その時までには目録作成やスキャニングの準備に関わるいくつかのプロジェクトを完了した。Eric Grant主席研究員を委員長、田邊主席研究員を副委員長とする新たな臨時委員会「研究資源センター運営委員会」（RRCOC）の設置に伴い、2021年11月にこの臨時委員会は廃止された。RRCOCの第1回会議は2021年12月に開催され、その後、2回の会議が開催された。2022年度末にGrant主席研究員が退職し、小野部長が委員長に就任した。Grant主席研究員の退職に伴い、昨年の計画の一部が変更された。

RRCは当初、物理的にも独立し専任の人員を有する放影研内の独立した「センター」となることを計画していた。現在でも長期的にはそうなる計画である。しかし設立を迅速に行うため、まず放影研の既存の部署にいくつかの新しい課を設立し、その後それらの課をRRCとしてまとめる「2段階」プロセスが採用された。「概要」で述べたように、RRCが機能するために必要な主な開発領域が3つ存在し、それは、1) データ統合、データのタグ付け、データ可視化、データセットの整理のためのデータフレームワークの設置および運用、2) 放影研内の全ての文書を一元管理し、（日英両方での）全文検索を可能にするコンテンツ管理システム、3) 全ての申請／承認手続きを合理化し、これらの手続きをオンラインシステムに移行させるとともに、RRCのオンラインプレゼンスを確立することを任務とする研究支援室、である。これらの課題を達成するために、RRCOCは以下の3つの課を新たに設立する（または既存の課を当該目的に充てる）ことを承認した。

- 研究資源課（RRS）：この新しい課は、情報技術部内に設置される技術面を担当する課で、小野部長が主導する予定である。この課は、データのタグ付け、データの視覚化、研究用の新しいデータセットの整理を含む放影研のデータを一元管理するためのフレキシブルなシステムを提供する「データ管理フレームワーク」の設置を監督することを任務とする。RRSでは、まずRRCのプログラミング作業を指揮する上級プログラマーを契約スタッフとして採用する。2023年度には、新しいプログラマーをRRSに採用する予定である。さらに、ITDのシステム技術課の職員2名にRRSのプログラマーを兼任させ、データ統合と開発作業を支援することにより、RRSを更に強化する予定である。
- 図書資料課：これは放影研に既存の課だが、新しいコンテンツ管理システム（CMS）に論文、研究計画書、業績報告書などを識別し、アップロードし、カタログ化する役割を追加する。また、外部業者とも協力して、放影研のニーズに合わせてCMSをカスタマイズする。この課は小野部長に直属する。
- 研究支援室（ORS）：以前、ORSは放影研事務局内に設置されていた。しかし以下の理由により、ORSは研究資源課に組み込まれる。1つ目の理由は、指揮命令系統を一元化するためである。2つ目の理由は、放影研戦略計画の観点からORSの役割が再検討されたためである。ORSの役割は、各種事務手続きを支援する、または取りまとめるだけでなく、機関の研究能力を強化する人的資源であるユニバーシティ・リサーチ・アドミニストレーター（URA）のように研究活動を促進することである。例えば、ORSは知的財産の管理・活用や、種々の外部研究資金源

の多様化に対応することになる。さらに、意思決定プロセスのデジタル化も計画している。これらのプロセスには、各種申請書の提出などの事務手続きだけでなく、放影研における研究に関する意思決定フロー全体が含まれる。この計画を実行し役割を遂行するためには、技術チームと非常に密接に連携し、一元化された指揮命令系統の下で活動する必要がある。

RRCOCがRRCの今後の体制や人員に関する要件を設定したことに加え、いくつかのプロジェクトが進行中である。

- **RRC設立に向けた5か年スケジュール：**RRCにおける技術開発のスケジュールを作成した。このスケジュールに含まれる主要な項目は、技術的發展に必要な放影研のインフラに関するいくつかの整備（外部通信回線の強化など）、現在放影研内で手作業と紙ベースで行われている多くのプロセスの自動化と合理化を可能にする電子ワークフローシステムの導入である。その他の主な項目としては、上記の業務を扱う新しい課の設立スケジュールも含まれる。RRCは2024年度末までに外部向けのオンラインプレゼンスを確立し、2025年度末までにフル稼働する予定である。
- **スキャニングセンターの実用化に向けたプロジェクトの始動：**昨年度構築したスキャニングセンターは現在実用化の段階に入っている。スキャナーにログインすると、全てのスキャンデータは光学式文字認識機能を通して送信され、新しく作成されたPDFファイルは作成者の名前により一元的に保存される。今年度は、ITDに提出する紙ベースの書類を、スキャニングセンターを利用して電子的に送信できる電子フローを確立した。これにより書類提出、特に遠く離れた長崎研究所からの提出が効率化される。またこの構築は、近い将来導入される電子ワークフローシステムのプロトタイプとなる。
- **コンテンツ管理システムに関する予備調査：**放影研はコンテンツ管理のためのオープンソースソフトウェアを特定し、そのソフトウェアを専門とする日本で有数のプログラミング企業を探し出した。そして、放影研のコンテンツ管理のためのパイロットシステムを開発するため当該企業と契約を締結した。作業は進んでいるが、放影研の設計仕様は相当難易度が高いものであることが判明した。その仕様とは、放影研のシングルサインオンシステムと互換性のある階層的なアクセスシステム、日本語および英語での完全な対応が可能であること、さまざまな資料の間の関係性を示す複合的なリンク、フレキシブルであること、カスタマイズ可能であることなどである。現在、システムの要件定義がほぼ完了し、システム調達に向けた仕様書を作成中である。
- **データ可視化に関する予備調査：**放影研は昨年度、データ可視化ソフトウェア「Tableau」のライセンスを3つ購入した。これはデータ可視化分野の先進的なソフトウェアであり、放影研のデータを探索するための「フロントエンド」として使用することができる。どんな複雑なソフトウェアでもそうだが、習得にはかなりの時間がかかる。このソフトウェアは有望と思われるが、成功するには、放影研で使用するためのパッケージの習得および開発に専念する人員が必要である。購入したライセンスは試験運用のため疫学部提供された。これにより、疫学部の杉山副部長から研究所内の実際のデータセットを使用した極めて実践的な可視化の例が提供された。そのため、疫学部用に Tableau のライセンスを4つ追加購入した。追加分のライセンスは、疫学部の腫瘍組織登録室および解析室に配布された。
- **劣化したマイクロフィルムのデジタル化：**原簿管理課および臨床研究部に保管されてい

る数十万枚のマイクロフィルムは、数十年にわたる不適切な保管状態により劣化が進んでいた。これらのフィルムを特定し、外部業者が全てデジタル化した。復元率はおおむね良好で、自動化装置でデジタル化できないほど劣化したリールは4、5本のみであった。2020年度から2021年度にかけて、約100万件のデジタル化ファイルが作成され、独立した電子ストレージに保管された。

- **Seven Bridges**社のデータ管理フレームワークについて評価が進行中：データ保存と管理ツールに適したフレームワークを特定するために、Grant 主席研究員と小野部長は、昨年6月に米国のIT企業に出張した。現在、研究所内の実際のデータを使用して、候補となるシステムを試験運用するための概念実証について計画している。概念実証用の一次データとして、所内のマウスの単細胞RNA配列データを使用する。
- **Grant** 主席研究員の退職：非常に残念なことだが、昨年12月にGrant 主席研究員が退職した。Grant 主席研究員、ITDの幹部職員、システム技術課と図書資料課のメンバーは、RRCの種々の任務について詳しく調べ、実施計画に着手するために、2019年に隔週の会議を開始した。高レベルの概念的な議論から、特定の機能に至るまで議論が行われた。計画遂行の途中で主要なスタッフが抜けた影響は極めて大きい。しかし、Grant 主席研究員は非常勤の研究員として放影研に協力し、我々は引き続き彼の指導の下にRRCの使命を遂行する努力を行う。