

研究資源センターの概要

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と思われる。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。これらには、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である。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ウェブポータルを介して行われる。
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、データ整理及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. 放影研の資源を活用・共有し、契約や外部資金の獲得を促進するための管理体制を提供する。

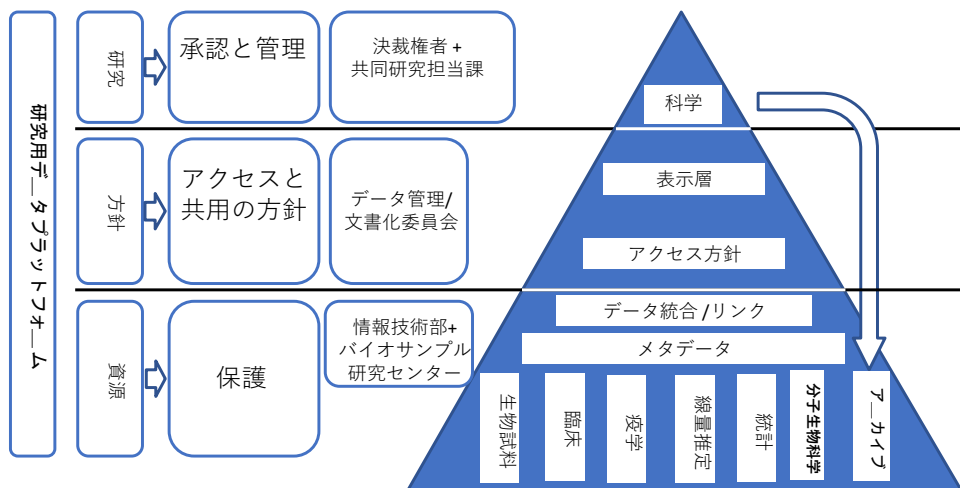


図1. 研究資源センターの機能的構造

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、デジタルデータ用のデータ管理システムや、放影研資料のデジタルスキャンした記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精査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成果物（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、「共同研究支援室」がデータやバイオサンプルを提供するための契約締結や手続きを支援し、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、解析スクリプト、論文原稿など）を、適切なメタデータとリンクを付け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。

職員配置、指揮系統、他の研究部との関係、予算などの研究資源センター（RRC）の構造については、まだ検討段階である。この計画の全体的な指導権は、2019年12月に設立された「研究資源センター設置準備委員会」の権限下にある。設置準備委員会には現在5つの小委員会があり、そのうちの1つが「研究資源センター設置要項ならびに運営要領起草小委員会」である。研究資源センターの構造に関する提言をしなければならないのは、この「起草」小委員会である。

RRCの構造については現在のところ決定されていないが、RRCには数名の常勤職員が配属され、IT開発やアーカイブを支援する部署及び共同研究支援室を持つことになろう。共同研究支援室は、放影研のデータやバイオサンプルの利用を希望する外部研究者の最初の連絡先となる。RRCは、資料目録にアクセスし、資料の保存やカタログ化を支援するために、放影研の全ての部署と協力関係を持つことが求められる。

正式な構造は決まっていなくても関わらず、情報技術部と主席研究員は過去2年間隔週で会議を行い、技術的な問題について検討し、必要とされるコンピュータインフラに関するパイロットプロジェクトならびに文書スキャン作業のパイロットプロジェクトを実施し、ソフトウェアインフラに関して外部業者と協議を重ねてきた。完了した具体的な作業と今後予定している作業については、次項に記載した。

## 2020年度業績

正式な組織構成構造や指導陣リーダーシップがまだ決まっていなかったが、2020年には技術面や管理面で多くの成果を挙げ、パイロットプロジェクト等も遂行した。

- **情報技術部長**：情報技術部（ITD）部長として小野博士が採用された。情報技術部長は2017年以降空席となっていた。新ITD部長は研究資源センター（RRC）と協力し、初期のデザイン決定を支援する。
- **新規採用プログラマー**：放影研は2名のプログラマーを新たに採用した。コンピュータ科学専攻の新卒者2名が4月に放影研に入所した。ITDでは10余年ぶりの新規採用となり、知識ならびに職員の年齢分布の両面における欠落部分を埋めた。RRCへの貢献が期待されている。
- **文部科学省の科学研究費補助金申請**：広島大学原爆放射線医科学研究所（RIRBM）との共同研究に対しへの文部科学省科学研究費補助金を申請した。この放影研とRIRBMの共

同申請（おそらくこの類の研究としては初）の内容は、一般市民や社会学者が興味を持つような歴史的資料を共同で保有し、一般市民がアクセスできるようにするためウェブベースの一般向けアーカイブシステムを構築するというものである。検索と表示のためのシングルインターフェースを備えた連携したデータベースを用いる設計になっている（各研究所はどの資料を提供するかについて独自に管理できる）。このシステムの設計から得られた知識は、RRCの設立、設計に含めるよう活用することができる。ウェブサイトの予定開設時期は、原爆投下から80周年にあたる2025年である（4月に申請結果発表）。放影研のPIはGrant主席研究員ならびに小野ITD部長である。

- 「研究資源センター設立準備委員会」、「研究資源センター設置要項ならびに運営要領起草小委員会」の発足

当該委員会発足以前は、RRCの設立に向けた取り組みは主席研究員の指示の下、ITD職員のみが行っていた。これは、放影研理事および放影研内の全研究部の協力が必要となるプロジェクトに向けた予備的な取り組みだった。RRC設立に向けた取り組みを正式なものとし、全研究部から協力を得るために、2019年12月、丹羽理事長とUllrich副理事長を共同委員長、Grant主席研究員と兒玉業務執行理事を共同副委員長とし、全研究部長、事務局長、2名の幹事から成る臨時の委員会を設立した。この正式な委員会により、RRCには十分な権限が与えられ、各研究部や事務局、その他の事務的なサポートを得て、RRCの実施プロセスを開始することができるはずである。研究資源センター設置要項ならびに運営要領起草小委員会は、指揮命令系統、リーダーシップ、予算、人事、RRCの内部構造など、放影研内でのRRCの位置付けを決定する責務を負う。これらを決定することは、RRCの設立を進め、運用を開始するために不可欠である。

- 「デジタル化小委員会」（小野ITD部長、丸茂図書資料課長が指揮する10名の委員会）  
「データ管理小委員会」（坂田疫学部副部長が指揮する13名の委員会）

この2つの委員会は、放影研で保存されている全ての資料の目録を作成し、デジタル化が必要な資料を特定するという第1段階の作業を終えた。データ管理小委員会は全研究部と会議や聞き取りを何度も行い、20種類の資料（デジタルファイル、ノート、地図、コード化マニュアル、紙のデータカード、研究室の写真など）のリストを特定、整理した。何百万枚もの紙資料、デジタル記録、フォルダーや箱で一杯のファイリングキャビネットに保存された資料等である。資料をさらに分類し保存の優先順位を決定するため、リストを現在、REDCapデータベース（オンラインデータベース）に移行中である。デジタル化小委員会はデータ管理小委員会と協力し、スキャンが必要な資料を特定した。

カルテのスキャン作業に関するパイロットプロジェクトを完了：外部の業者と契約を結び、広島研究所臨床研究部に保管されている紙のカルテをスキャンするために必要な労力を詳しく調査するため、パイロットプロジェクトが実施された。3週間の試行期間中に、様々なワークフロー方法、オーバーヘッドスキャナーとフラットベッドスキャナーの比較、様々な劣化状況のための留め具の取り扱い方法、サイズが不揃いの紙の取り扱いなどをテストし、その後報告書が作成され、放影研に提出された。パイロットプロジェクトにより、カルテには1,000万枚の紙が使われていると推定された。業者は、カルテの資料を全てデジタル化するには約500万ドルかかると結論づけた。25人のチームであれば、3年で当該業務を完了できる。25人より少ない人数でも同じ費用でデジタル化できるが、作業期間は長くなる。長崎研究所の資料をスキャンすると、総費用が約4割増えると推定されている。

- **“Gen3”データコモンズパイロットプロジェクトが進行中**：シカゴ大学のRobert Grossman博士の来所後、放影研はGen3システムを用いたパイロットプロジェクトを試みることにした。Gen3は、大規模データの保存と解析のためのオープンソースのソフトウェアプラットフォームである。このパイロットプロジェクトが検討された時点では、放影研は専門知識も人材も持ち合わせていなかった。Grant主席研究員は人脈を用い、テキサス大学メディカルセンターで疫学の博士号を取得中のコンピューター科学者を放影研でのインターンシップに採用した。当初の提案では、当該プログラマーは春にシカゴ大学で数週間の研修を受けた後、夏にはGen3の専任としてインストール作業を行うため広島研究所に来所することになっていた。しかし、新型コロナウイルス感染拡大の影響で、全ての来所は中止せざるを得なかった。彼の名誉のために言っておくと、そのコンピューター科学者は、夏の間および2020年12月にかけて、毎週Grant主席研究員とテレビ会議を行っていた。残念ながら、このパイロットプロジェクトでインストールは完了しなかった。これは、個人の過失ではなく、悪い状況が重なってしまったためである。複合的な複雑にした要因としては、Gen3システムは始まったばかりで文書化が不完全であったこと、Gen3の技術員が日々の業務に追われて放影研のGen3のインストールに十分な時間を割けなかったこと、大学院生がこのプロジェクトを遠隔で行おうとしたことなどが挙げられる。
- **技術チームの隔週会議**：Grant主席研究員、ITD指導陣、システム技術課および図書資料課の職員は、RRCの様々な業務について詳細に調査し実施計画を立てるため、2019年から隔週で会議を開催している。会議で協議される内容は、高度に概念的な議論から具体的な機能まで多岐に渡っている。技術的なニーズや放影研が望む機能を現在実践しているグループやソフトウェアプラットフォームを特定するため、かなりの時間を費やし調査した。このチームは技術的な問題に焦点を当てている。より大規模大きな「研究資源センター設立準備委員会」では、実施しなければならない多くの業務を特

定して実行するが、実際の技術的な実装作業はITDのチームが（外部の専門家の助けを借りて）行うことになるだろう。これらの会議は、正式な委員会の範囲を超えて継続する。

2020年に実現しなかったが、RRCの設立に関する取り組みは2017年に始まった。2020年以前に行ったその他の活動、成果を簡単に以下に述べる。

1. 世界中のデータやアーカイブの専門家を訪問。ITDの職員は、データ統合やアーカイブのための既存の専門知識をもともと持っていないため、選択肢や開発戦略について把握するため、大規模プロジェクトに携わっている様々な機関の専門家と連絡を取る必要がある。以下の専門家/機関と対面の会議を開いた：
  - Dave Thompson／ウェルカム図書館デジタルキュレーター
  - Stephen Chanock／米国国立がん研究所がん疫学・遺伝学部長
  - Anthony Philippakis／ブロード研究所最高データ責任者
  - James Cuff／ハーバード大学研究コンピューティング部門担当副学部長
  - Brian White／ハーバード大学研究コンピューティングITオペレーションマネージャー運用責任者
  - John Quackenbush／ダナ・ファーバー癌研究所 生物統計学・計算生物学担当部
  - Meir Stampfer／ハーバード公衆衛生大学院 看護師健康調査研究責任者(PI)
  - James Lacey／カリフォルニア教師研究 健康分析論分析学責任者
  - Tom Murphy／ミシガン大学政治・社会調査のための大学協会（ICPSR） コンピューターおよびネットワークサービス部長
  - Robert Grossman／シカゴ大学 生物科学部 研究情報科学最高責任者
  - 有賀 暢迪／国立科学博物館 理工学研究部
  - 杉本 重雄／筑波大学 図書館情報メディア研究科

2. James Cuff博士（ハーバード大学）の来所 2017年5月

放影研の研究上のニーズと現在のインフラを評価するために、ハーバード大学研究コンピューティング部門担当副学部長であるJames Cuff博士が3日間の日程で来所した。Cuff博士はITDならびに放影研の指導陣と協力し、問題を調査し論理的な手順段階に分類解するために「白書」の概要を作成した。

3. Robert Grossman博士（シカゴ大学）の来所 2019年1月

シカゴ大学を訪問した際、放影研のデータ統合のニーズに対してGen3システム

が良い解決策になるかどうかを議論したことを受け、生物科学部研究情報科学最高責任者のRobert Grossman博士が2日間来所した。Robert Grossman博士は、講演（「Gen3を使ったデータセンターの構築」）を行い、来所中のアーキビスト、ITD職員、放影研の指導陣と会議を開いた。会議終了後、Gen3システムを使用したパイロットプログラムを放影研で実施することが決定した。

#### 4. Christy Henshaw博士（ウェルカムトラスト）の来所 2019年3月

デジタルプロダクションマネージャーのHenshaw博士は、放影研で「ウェルカムコレクションにおける電子化デジタル化の取組み」と題した講演を行い、ウェルカムトラストのオンラインプレゼンスオンラインでの活動に関する戦略や業務上の技術的な詳細について語った。Henshaw博士は、放影研の図書資料課職員や、講演会に招待された地元のアーキビスト達とも会議を開いた。ウェルカムトラストが使用しているツールの多くはオープンソースであり、放影研のRRC設立に役立つかもしれない共同プロジェクトについて協議した。

#### 5. RRC支援用のサーバの購入

1月の会議でパイロットプロジェクトについて合意に至った後、技術的要件についての議論が始まった。Gen3チームは、最低でも少なくとも2台の高性能サーバを推奨した。放影研は3台のサーバを購入することを決定した（Gen3用に2台、検索やアーカイブを行うウェブサーバ支援用に1台）。夏に用度課が公開入札を開始した。2019年9月にコンピューターの購入、納品が完了した。

#### 6. Robert Grossman博士と技術者2名の来所

2019年11月、シカゴ大学のグロスマン博士と技術者2名が、新規購入したサーバに「Gen3データ commons」のパイロット版をインストールするために来所した。グロスマン博士は2日間、技術者は1週間滞在した。放影研で最高のシステム技術者がシカゴ大学の技術者と共に、使用するインストールツールや必要なネットワーク設定について理解するため1週間かけて共同作業を行った。インストールの初期作業は1週間で完了した。このシステムに関し、放影研は証明書に基づく第三者認証システム（Microsoft Azureインフラ）を用いた最新の認証システムを採用する必要があった。

#### 7. ヒトゲノムの共同研究のために、ビデオ監視および生体認証による入退室管理システム付きの安全なコンピューター室を2部屋設置

放影研ではヒトゲノムデータの導入取得を進めているが、そのための所内での計算機能力は十分ではない。この種の研究を行うためには、スーパーコンピュ

---

ーター（当初の目標は東北メディカル・メガバンクのスーパーコンピューター）にアクセスする必要がある。このコンピューターを使用するためには、認可された物理的スペースとコンピューターインフラが必要であった。建設業者と仕様書に基づき、現在放影研にはこのその特殊機能用機能のための専用の部屋室が2部屋設置されている。

8. オープンライブラリーサービスに関する会議や講演、日本におけるオープンデータ・イニシアチブ、オープンレポジトリに関する会議、専門家養成に関する講演講義等、会議や講義に参加した。
9. 白書の完成

