

共創先導プロジェクト（共創促進研究）
「学術知デジタルライブラリの構築」基本計画

令和4年4月1日
人間文化研究機構

【プロジェクトの概要等】

① プロジェクトの概要

我が国の研究者・研究機関が世界諸地域で撮影・収録した写真・動画・音声資料は、特定の時代、地域の姿を記録した貴重な記録であると同時に、学術史を反映する歴史遺産でもある。国内の研究者・研究機関によって蓄積されてきた写真・動画・音声資料は天文学的な数量にのぼるが、それらの資料は、これまで、当該研究にのみ利用され、また、研究終了後の保存の手立ても十分に施されずにおかれ、他者が利用可能なかたちで公開されてこなかった。

本プロジェクトは、国立民族学博物館と国立国語研究所が国立情報学研究所と連携して、国内の大学・研究機関に属する研究者・研究機関の研究の過程で蓄積された写真画像、映像、音声資料等の統合的なデジタル化・データベース化とその高度統合化の進捗を図るものである。

国立民族学博物館（民博）拠点では、既に構築した写真画像、映像等の統合的なデジタル化・データベース化のシステムを適宜改善しつつ活用し、国内の大学・研究機関に属する研究者・研究機関を対象として、写真画像、映像等の統合的なデジタル化・データベース化の作業を支援して、研究に活用できるプラットフォームを提供する。この事業を通じて、当該研究者・研究機関の研究の進捗を図るとともに、そのうち公開可能なデータを国際的に共有化することで、分野の別を超えたオープン・サイエンスの基盤を構築することを目指している。

国立国語研究所（国語研）拠点では、国立民族学博物館の有する写真画像・映像のデジタル化・データベース化システムの開発・運用に係る知見を基礎としつつ、国立情報学研究所と連携して、音声・映像による言語の記録に共通のメタデータを付与し、言語資料として蓄積するシステムの開発に取り組み、そのシステムの有用性を確認した上で、他機関、個別研究者を対象とした言語記録の研究資源化支援活動への展開をはかる。

民博拠点での事業は、①デジタル化・データベース化支援活動と②ドキュメンテーション支援活動からなる。

- ① データベース化支援活動は、国立民族学博物館と国立国語研究所が国立情報学研究所と共同して、写真・動画・音声資料のデジタル化・データベース化システムの構築とその運用、資料情報の分析手法の開発にあたるものである。
- ② ドキュメンテーション支援活動は、国立民族学博物館所が、静止画と動画に必要な研究情報の付加の作業の支援にあたるものである。この作業には、画像資料への研究支援員による情報入力のほか、AIによる自動タグ付けの作業や資料情報の公開に付随する著作権や肖像権等の権利関係の処理の作業も含まれる。

事業の実施にあたっては、大学共同利用の観点から、4機構の内外に対して広くプロジェクトの公募をおこなう。

国語研拠点では、国立民族学博物館・国立情報学研究所と連携して音声・映像による言語の記録を

的関心から撮影・収録されたものであろうと、また自然科学的関心から撮・録られたものであろうと、対象となった時空間の瞬間の現実を切り取った資料として、撮影者・収録者の想定以上の情報量を潜在的に有している。それらが相互参照可能なかたちで情報リンクされたとすれば、蓄積された資料が媒体として有する価値は、他に類を見ないものとなる。本プロジェクトにおいては、相互参照が困難であった写真・動画・音声資料を統合的なデータベースとして整理することによって学術的な付加価値の高いビッグデータが生成される。これにより、個々の資料を対象とした研究に格段の進展が期待されることは言うに及ばず、これまで個別の学問体系や地域の特有のものとして処理されてきた問題を地球規模の時空間にかかる問題系としてとらえる異分野融合型の地域研究への展開が可能になり、分野横断的な知の開拓にも大きく貢献することになる。また、現地、調査対象国・地域への研究成果の還元という意味でも、本プロジェクトの意義は大きい。

④ デジタル化及びプラットフォーム構築による学術的貢献

人間文化研究機構においては、第4期中期目標・中期計画期間に向けて、人間文化研究創生センターを新たに設置し、「デジタル・ヒューマニティーズの推進」に注力することとしている。本プロジェクトは、人間文化研究の過程で生成された写真・動画・音声資料情報を、国立情報学研究所との機構間連携を通じて新たな知の基盤を「共創」というものであり、「デジタル・ヒューマニティーズの推進」の中核の一つを担う事業として位置づけられる。

国立民族学博物館では、平成28年度より令和3年度まで、科学研究費助成事業・新学術領域研究（研究領域提案型）『学術研究支援基盤形成』『地域研究に関する学術写真・動画資料情報の統合と高度化』プラットフォーム「地域研究画像デジタルライブラリ」（DiPLAS）の事業を実施し、進行中の科学研究費補助金助成事業を対象に、国内外で撮影された画像・動画資料のデジタル化・データベース化の作業の支援をおこなってきた。同事業もまた、国立情報学研究所との連携で実施して入る事業であるが、今回提案するプロジェクトは、同事業を通じて培われた高度な情報システムの技術と画像アーカイブ構築のノウハウを活用し、国内の大学・研究機関に属する研究者・研究機関に蓄積された写真・動画資料の共有化を図ろうとするものであり、分野を超えた大学の研究力強化の基盤構築に資する。

本プロジェクトの実施により、さまざまな写真・動画・音声資料を利用するためのプラットフォームが統合され、検索できるようになれば、20世紀以降の世界のありさま、人類のありようを、時空間を問わず横断的に見渡せる装置が確立され、世界各地の状況を知る上で質量ともに比類のないデジタルライブラリが構築され、地域研究に係るオープン・サイエンスの基盤が形成されることになる。また、このデジタルライブラリを集積し公開することは、これまでの日本の海外学術調査が撮影してきた写真・動画・音声資料の対象国・地域への還元に結びつく。情報・知識の現地との共有化は、21世紀の研究倫理に欠かせない要素であり、本事業の推進が、今後の海外学術調査における研究協力を得るのに大いに資することは間違いない。さらに、データベースに実装される高度な検索機能は、マクロとミクロ双方の視点から文明史や地球規模の変動を分析する可能性を開く。

⑤ 達成目標

民博拠点

- ・DIY型データベース機能をデータベースに実装し、データベースの共同編集を実現
- ・横断的検索を可能とするカレイドスコープの構築

言語資料として蓄積するシステムの開発に取り組むとともに、音声・動画資料に関するドキュメンテーション手法の精緻化、デジタル化手法の共有化をはかり、将来の言語記録の研究資源化支援活動に備える。

② プロジェクトの統括、運営体制

民博拠点における上記の①デジタル化支援活動と②ドキュメンテーション支援活動の統括と調整運営は、「プラットフォーム委員会」が担当する。①デジタル化・データベース化支援活動は、国立民族学博物館が国語学研究所と連携しつつ、国立情報学研究所と共同して、画像・動画資料のデジタル化を進め、データベース・システムの構築とその運用にあたるものである。②ドキュメンテーション支援活動では、国立民族学博物館が、分野を異にする研究支援担者の協力を得て、採択されたプロジェクトの資料調査を実施し、資料の特性に合わせたデータベース項目の整備をはかる。その後、国立民族学博物館において、利用許諾の取得もしくは著作権処理を行なううえで、技術支援員が、基本情報と画像内容に関わるテキスト情報の入力支援を実施する。

公募プロジェクトでは、以上のような作業を経て提供されるデータベースを用いて、関連情報を付加しつつ、科研の研究計画を遂行する。このようにして築かれたデータベースは、公開にむけてデータを付与していく、個々のデータの公開可否を点検したのち、公開可能なものから順次国際的に共有することとする。

国語研拠点では、国立民族学博物館・国立情報学研究所と連携して音声・映像による言語の記録を言語資料として蓄積するシステムの開発に取り組むとともに、関連分野の研究者との協同により、音声・映像による言語の記録に対するドキュメンテーション手法の精緻化、デジタル化手法の高度化をはかり、音声・動画資料の整備・公開を促進する。

両拠点は、プラットフォーム委員会と議論を共有しつつ、写真・動画・音声資料を統合的に閲覧し、横断的な検索を可能とするためのシステムを共同で開発する。また、プラットフォーム委員会と連携を図り、構築したデータベースを用いた分野横断的な共同研究を推進する。

③ 期待される学術的研究成果とその学術的・社会的意義

20世紀の中葉以降、科学研究費補助金による調査や、大学共同利用機関等の設置による共同研究によって毎年多くの学術調査が国内外で実施され、当該分野での顕著な成果をおさめてきた。これらの調査によって蓄積されてきた写真・動画・音声資料は、研究者個人による私的な資料も含めると天文学的な数量にのぼる。しかしながら、これらの資料は、1) 研究者個人や当該研究に専ら利用されることを想定していること、2) 自然科学分野の調査による資料と人文社会科学分野の調査による資料とで相互に利用することを想定していないこと、3) 当該資料に対する付加的情報（撮・録られた対象や場所、それに付随する状況）が必ずしも他者が利用可能なかたちで明示されていないこと等、さまざまな理由から、学術的な共有財とはなっていない。ひとつひとつの資料は、それが人文科学

- ・館外研究者をまじえた「プラットフォーム委員会」と民博 情報課の連携により、研究者コミュニティの意見を反映しつつ運営体制を強化
- ・すでに公開したデータベースの多言語化
- ・民博内外のポータル型データベースとの連携
- ・各年度10件ずつのデータベース構築
- ・各年度10件ずつのデータベース公開
- ・データベース活用に関する啓発と普及

国語研拠点

- ・音声・映像による言語の記録を言語資料として蓄積するシステムの開発
- ・音声・動画資料のドキュメンテーション手法の精緻化と普及
- ・音声・動画資料のデジタル化手法の高度化と普及
- ・音声・動画資料（主として国立国語研究所収集資料）の整備・公開

⑥ 6年間のロードマップ

※ 主要な研究成果の発信（国際会議、成果物等）を中心に記載

民博拠点

年度	取組内容
令和4年度	DIY型データベース機能の実装
令和5年度	DIY型データベースの運営体制確立
令和6年度	プラットフォーム委員会と民博情報課の分担と連携の見直し
令和7年度	データベース多言語化の準備、カレイドスコープの構築
令和8年度	データベース多言語化の準備、民博内外のデータベースとの連携の準備
令和9年度	データベース多言語化の実現、民博内外のデータベースとの連携の実現

国語研拠点

年度	取組内容
令和4年度	蓄積のシステムの調査、ドキュメンテーション手法精緻化に着手
令和5年度	蓄積システムの設計、ドキュメンテーション手法案の策定
令和6年度	蓄積システムの試作、ドキュメンテーション手法案の検証
令和7年度	蓄積システムの検証、ドキュメンテーション手法の完成
令和8年度	蓄積システムの運用開始、ドキュメンテーション手法の普及
令和9年度	蓄積システムの普及