

日本史史料を対象とした研究データ基盤整備における課題

山田 太造（東京大学 史料編纂所／地震火山史料連携研究機構）

中村 覚・渋谷 綾子（東京大学 史料編纂所）

大向 一輝（東京大学 大学院人文社会系研究科）

井上 聡（東京大学 史料編纂所）

概要：人文科学を含む科学技術とイノベーションの創出の一体的・総合的な振興が急激に進められている。人文科学を取り巻く研究環境の急速な変化に対し、東京大学史料編纂所が推し進めている日本史史料を研究資源とした歴史研究データ基盤構築の状況について、具体的な取り組み内容を示すとともに、この基盤整備における課題について議論する。

キーワード：日本史、データ基盤、データ駆動型、研究資源

Challenges in Developing a Research Data Infrastructure for Japanese Historical Documents

Taizo YAMADA (Historiographical Institute/The Collaborative Research Organization for Historical Materials on Earthquakes and Volcanoes, the University of Tokyo)

Satoru NAKAMURA / Ayako SHIBUTANI (Historiographical Institute, the University of Tokyo)

Ikki OHMUKAI (Graduate School of Humanities and Sociology, the University of Tokyo)

Satoshi INOUE (Historiographical Institute, the University of Tokyo)

Abstract: The integrated and comprehensive promotion of science and technology, including the humanities, and the creation of innovation is being promoted at a rapid pace. In order to respond to the rapid changes in the research environment surrounding the humanities, we explain the specific efforts made by Historiographical Institute the University of Tokyo to construct a data infrastructure for historical research using Japanese historical materials as research resources, and discuss the issues involved in developing the infrastructure.

Keywords: Japanese History, Data Infrastructure, Data-Driven, Research Resource.

1. はじめに

「21世紀の石油」とも称される「データ」の存在への着目は高まり続け、情報社会を動かす基本的な資源[1]として認識され、客観的データに基づく意思決定を行うデータ駆動型社会への取り組みも各種マスメディア等で報道される状況下にある。また COVID-19 パンデミックが WHO により宣言された 2020 年 3 月以降はデジタル・トランスフォーメーション (DX) に対する取り組みも改めて重要視されている。国内 9 大学 2 研究所が連合して共同運営する全国共同利用に供するデータ科学、データ駆動科学、データ活用にフォーカスした高性能仮想化環境を提供するデータ活用社会創成プラットフォーム (mdx) が 2021 年 3 月より東京大学柏 II キャンパスにて稼働を開始[2]するなど、研究における DX への取り組みも加速している。

2021 年 4 月 1 日、「科学技術基本法」は「科学技術・イノベーション基本法」に法律名を変更し施行された[3]。(これまでは“科学技術の対象外”

として位置づけられていた) 人文科学を含む科学技術とイノベーションの創出の一体的・総合的な振興を進めるための法改正である。これに連動し、2021 年 3 月 26 日に閣議決定された第 6 期科学技術・イノベーション基本計画[4]には「自然科学のみならず人文・社会科学も含めた多様な「知」の創造と、「総合知」による現存の社会全体の再設計、さらには、これらを担う人材育成が避けては通れない」ことが盛り込まれた。さらにはオープンサイエンスとデータ駆動型研究等の推進を強調しており、「人文・社会科学の研究データの共有・利活用促進するデータプラットフォームについて、2022 年度までに我が国における人文・社会科学分野の研究データを一元的に検索できるシステム等の基盤を整備するとともに、それらの進捗等を踏まえた 2023 年度以降の方向性を定め、その方針に基づき人文・社会科学のデータプラットフォームの更なる強化に取り組む」ことが明記された。人文科学を取り巻く研究および研究データの扱いは、他の分野におけるそれと遜色なく、かつ、早急な整備が求められている。

日本学術振興会は2018年度より人文科学・社会科学研究に係るデータを分野や国を超えて共有・利活用する総合的なシステムを構築することを目的[5]として「人文学・社会科学データインフラストラクチャー構築推進事業」を開始し、2021年7月よりJDCat (Japan Data Catalog for the Humanities and Social Sciences)¹と呼ばれる人文学・社会科学分野データのメタデータの一括検索可能なデータカタログを公開した[6].

本稿では、このような人文科学を取り巻く研究環境の急速な変化に対し、東京大学史料編纂所(以下、史料編纂所)が推し進めている日本史史料を研究資源とした歴史研究データ基盤構築の状況について具体的な取り組み内容を示すとともに、この基盤整備における課題について議論する。

2. 日本史史料の収集と管理

COVID-19 パンデミックの影響により、史料編纂所だけではなく多くの史料所蔵機関での史料調査に関するサービスが停止してしまい、日本史のみならず、人文学における研究へ多大な影響を与えた。2019年度は約860万件だった史料画像へのアクセス数は、2020年度は約950万件と増加した。ウェブ上でのデータ提供・共有の促進により、このような状況においても研究を停止させない方法の1つとして挙げることができる。

ここでは、史料編纂所が行ってきた日本史史料収集の経緯とその方法、史料のデジタル化とその管理方法、および、実際の取り組み事例について述べる。

2. 1. 史料収集

1872年に史料編輯国史校正局が開設され「修史ハ萬世不朽ノ大典」から始まる明治天皇宸翰をもって三条実美がその総裁に任じられるところから明治の修史事業が開始され、これ以降の150年、史料編纂所は史料の調査と収集(史料採訪、もしくは、採訪)・編纂・出版という事業の柱を維持し続けている。史料編纂所は研究材料としての史料の収集・精確な読解・史料批判を歴史学研究の基礎として位置づけ、日本史史料研究を重視してきた。組織的かつ系統的に前近代日本史史料の採訪を1885年より本格化し約140年間にわたり、継続的に行ってきた。その成果は影写本・謄写本・写真帳などの複製史料として蓄積し、これらを公開することで研究者間での共有を可能にし、日本史学の発展や深化に大きく寄与してきた。この活動を通じて築き上げてきた、日本史史料コレクションとも言うべき史料編纂所所蔵史料の概要²を図1に示す。

種類	内訳	数量
史料	原本・写本類	200,355点
本所作成史料	影写本	7,105冊
	謄写本	22,705冊
	写真帳	45,872冊
	台紙付写真	23,222点
フィルム類 (複製本を含む)	マイクロフィルム	49,924リール
	シートフィルム	8,066タイトル
	乾板	9,000枚

図1 東京大学史料編纂所所蔵史料の概要
Figure 1 Summary of the historical materials in Historiographical Institute the University of Tokyo

2. 2. 史料画像のデジタル化

採訪において、1970年から2010年の間はマイクロカメラ撮影による史料収集を行っていた。このうち、採訪に関わるマイクロフィルム(11,393リール)およびシートフィルム(8,066タイトル)を対象に、2008年から2015年までの8年の歳月を経て、スキャニングを行い、史料画像を作成した。海外所在の日本関係史料を収載したマイクロフィルムは現在もスキャニングを続けている。

2000年から2008年の間、影写本や謄写本など史料編纂所が作成した写本類の大半を対象に、マイクロカメラによる撮影を行い、マイクロフィルムのスキャニングを経て、これらの史料画像を作成した。

2008年以降の史料編纂所が所蔵する原本史料撮影や2010年以降の採訪ではデジタルカメラを用いている。これらは現在も継続していることから、史料画像は増え続けている。2021年3月31日時点では、史料編纂所所蔵史料(写本類を含む)の画像は約950万点、採訪史料の画像は約1,080万点で、合計2,000万点以上の史料画像コレクションへと成長した。

史料編纂所所蔵史料に関する史料画像は、主に、史料編纂所歴史情報処理システム(SHIPS)のデータベース検索サービス(SHIPS DB)におけるHi-CAT(史料編纂所所蔵史料目録DB)を介して提供している。また、採訪史料の画像は、SHIP DBのHi-CAT Plusにより公開している。

2. 3. 史料画像公開のフロー

マイクロフィルムを用いた採訪からデジタルカメラを用いた採訪へと移り変わることが予想された2008年、史料編纂所は所内に組織したデジタル採訪小委員会にて、マイクロカメラ撮影による採訪の方式を踏襲しつつ、デジタルカメラへ変更した際へスムーズに移行するための採訪方式を提案・確立し、史料画像公開までのフローをモデル化(図2)した[7]。このフローは、史料画

¹ <https://jdcats.jp>

² 2016年3月31日時点

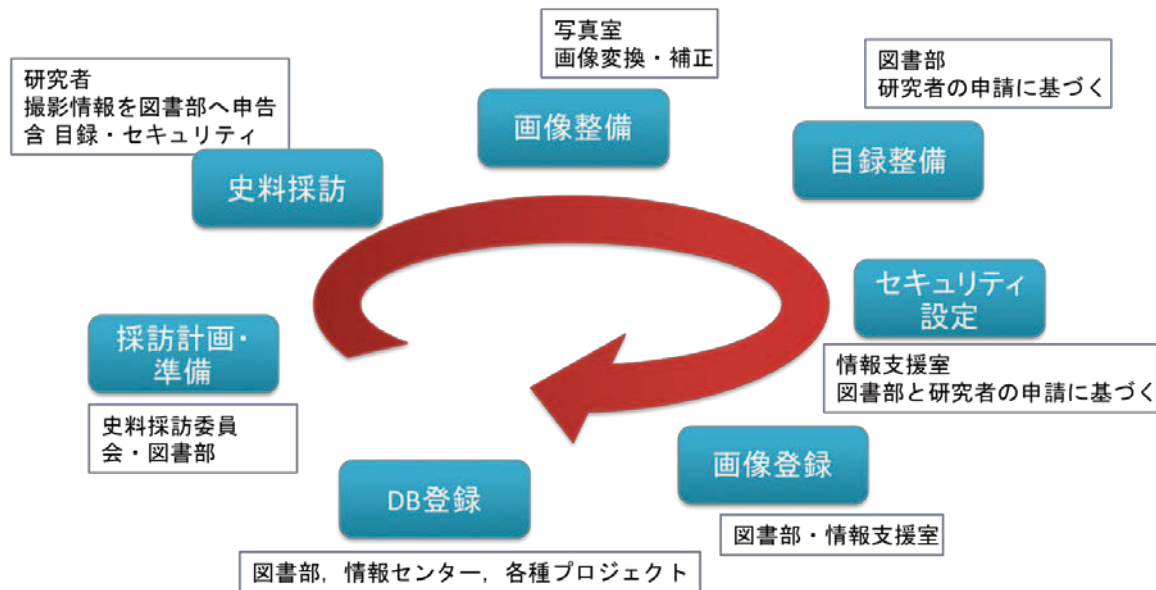


図 2 収集史料公開のフロー
Figure 2 The flow to release collected historical materials

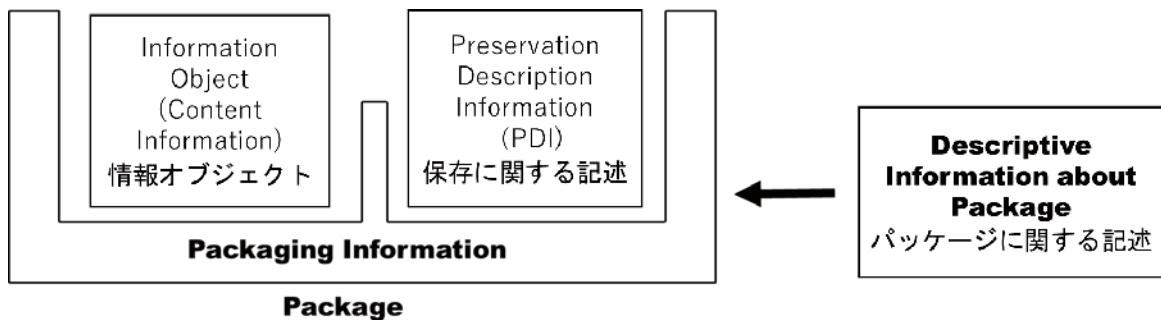


図 3 OAIS 参照モデルにおける情報パッケージ
Figure 3 Information Package in OAIS reference model

像公開までを史料編纂所の全教職員が関わりながら進めていく組織的な取り組みとしてモデル化した。この流れに従って処理された画像のみが共有され、永続的に管理されていく対象になる。さらに、デジタルカメラでの撮影方法、画像仕様等についても規約化を進めた。このモデル化や規約化には 2 年をかけて進め、2010 年に試行し、2011 年より本格的に導入した。

史料編纂所では年間 50 件を超える探訪を行っている。1 度の探訪で何箇所もの機関等へ訪れることから、年間 200 以上の DB 番号コードが発給される¹。史料画像がデータベースへ登録されるまでに 1 年以上かかることもある。また、このフローは 16 の段階²から構成されている。探訪史

¹ 探訪は下記のページで確認することができる:

<https://www.hi.u-tokyo.ac.jp/investigation>

² 具体的には、BD 番号申請、BD 番号発給、撮影画像受入、写真室完了 (画像調整・フォーマット変換)、画像受け入れチェック、画像整序、撮影情報申告、撮影

料画像を永続的に共有・管理していくためには、このフローの管理が求められることから、2012 年に探訪進捗管理システムと呼ばれる、このフローを管理するシステムを構築し、運用してきた[8]。さらに、史料編纂所所蔵の史料画像公開においてもこのフローを適用することを決定したことから、名称を「史料画像デジタル化進捗管理システム (以下、進捗管理システム)」に改めた。2021 年 10 月 29 日時点で、3,508 件の史料画像公開状態を管理している。

進捗管理システムは史料編纂所外の利用者はいないこともあり、独自モデルとして構築した。しかしながら、史料画像の長期利用・長期保存の観点を考慮していなかった。そこで OAIS 参照モデル (Reference model for an Open Archival

データ確認、IPTC 付与、所蔵史料目録登録、画像登録、セキュリティ設定、簡易検索目録データ登録、1 点目録加工、1 点目録作成、1 点目録登録の段階がある。図 2 ではこれらを集約し 7 段階で表現している。

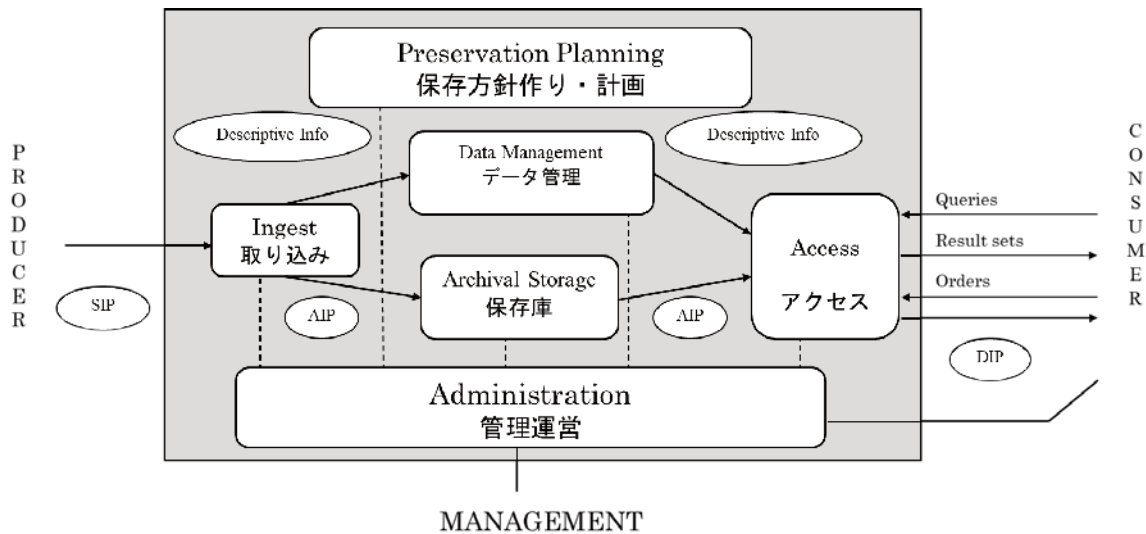


図 4 OAIS の機能要素
Figure 4 OAIS Functional Entities

Information System) [9,10]に従ったシステム再構築を行った. OAIS 参照モデルでは, 情報オブジェクト (Information Object) がデータオブジェクト (Data Object) とその表現情報 (Representation Information) から構成されており, 情報オブジェクトを保存するために情報パッケージ (Information Package) の概念を導入している. 情報パッケージは, 図 3 に示すように情報オブジェクト (Information Object) を内容情報 (Content Information) として保存するため, 保存記述情報 (Preservation Description Information, PDI) と併せてパッケージ化する. 情報パッケージに関する記述 (Descriptive Information about Package) を付与し, パッケージ化したオブジェクトを保存する. PDI は以下の 5 つの情報を持つ.

- 来歴 (Provenance) : コンテンツの由来・伝来. 保存の処理など履歴.
- コンテキスト (Context) : コンテンツとその環境との関係. 他のオブジェクトとの関係など.
- 参照 (Reference) : コンテンツの識別子.
- 不変性 (Fixity) : オブジェクトが変更されていないことを保証する情報.
- アクセス権 (Access Rights) : コンテンツへのアクセス制限・条件.

進捗管理システムを改修することで, PDI の登録・管理が可能になった. OIAS 参照モデルでは, 実運用において図 4 に示す機能要素のように, 保存方針計画 (Preservation Planning), (コンテンツの) 取り込み (Ingest), データ管理 (Data Management), 保存庫 (Archival Storage), アクセス (Access), 運用統括 (Administration) という 6 つの機能要素が定義されている. 史料編纂所では次のように対応している. これらの機能要素について, 史料編纂所では次のように対応してい

る: 保存方針計画は史料編纂所内に設置された小委員会によりなされ, 保存庫は進捗管理システムに相当し, データ管理およびアクセスは, 主に Hi-CAT および Hi-CAT Plus により行われている. 管理運営は図書部および前近代日本史情報国際センターにより行われている. 史料編纂所は, 図 2 で示したフローをモデル化し, 10 年を超えて運営してきたことにより, このフローを実証してきた.

2. 4. データ基盤としての取り組み

史料収集に関するこれまでの取り組みが評価され, 2019 年 10 月より史料編纂所は人文科学・社会科学データインフラストラクチャー構築推進事業における人文科学に関わる唯一の拠点として認定された. データアーカイブ機能の強化 (共有化), 海外発信・連携機能の強化 (国際化), データ間の連携を可能にする環境の整備 (連結化) について取り組んでいる[11].

共有化では, 主にデジタルデータの長期保存利用・長期保存の実現とデータ利用条件整備を行っている. 長期保存・長期利用は 2.3 節で述べたとおりである. 史料画像データ利用条件の整備として, 2019 年 4 月に, 史料編纂所が所蔵する原本等の史料画像データ¹であれば Creative Commons (CC) BY 相当, 史料編纂所出版物の版面画像データ²の画像データは CC BY-NC-SA と定めた³. その他の史料画像データは史料編纂所図書室へ申請が必要である. 史料編纂所所蔵史料だけでな

¹ 史料種別 (区分) が「貴重書 (原本・古写本類)」・「特殊蒐書」 (一部を除く) ・「写本」である画像が対象.

² 史料種別 (区分) が「本所出版物」である画像が対象.

³ <https://www.hi.u-tokyo.ac.jp/faq/reuse>

く、原蔵者との調整¹のもと、採訪史料の画像にも利用条件を設定しつつある。

国際化では、SHIPD DBの1つである維新史料綱要DBの英語版であるSummary database of the Ishin Shiryoを構築している。「維新史料研究の国際ハブ拠点形成」プロジェクトによる『維新史料綱要』の英訳化と史料用語・歴史用語の英訳グロッサリー作成の成果をもとに、データ登録している。2021年10月時点では、『維新史料綱要』全10巻中の3巻(約6,000件)の綱文を英訳している。

連結化では、IIIF Presentation APIを利用した史料画像の共有や他機関とのデータ連携・システム連携を行っている。2020年3月、IIIF Presentation APIを利用したHi-CAT Plusウェブ版を公開した。同時に、「天皇家・公家関係史料を中心とした研究」プロジェクトとの連携により、宮内庁書陵部所蔵の伏見宮家本・九条家本の画像(約567,720点)を公開した。2020年12月には山口県立山口図書館所蔵の今井似閑本の画像(約32,000点)を公開した。2021年7月、都城島津邸所蔵史料の画像(約11,000点)を公開した。他機関とのデータ連携・システム連携の具体例としては、史的な文字データベース連携検索システム²がある。2009年に公開した奈良文化財研究所とのくずし字連携システム『電子くずし字字典データベース』『木簡画像データベース・木簡字典』連携検索を進展させ、国文学研究資料館・国立国語研究所・京都大学・漢字字体規範史データセット保存会・台湾中央研究院と連携し、各機関の独自性を尊重した対等な連携検索、国際的な標準規格に準拠した枠組み、オープンデータ化も含めた利便性の向上を目的として、2020年3月にプロトタイプを公開し、2020年10月から正式公開するに至った。

3. データの流通

史料編纂所は2021年4月より「データ駆動型歴史情報研究基盤の構築」プロジェクトを開始した。これは日本史史料データをさらに100年にわたって維持・発展させるための研究基盤構築を目指しており、史料の目録・画像のみならず、本文・人名・地名・時間・イベント・文字など史料に関わるあらゆるデータをDNNやCNNなどのディープラーニングおよびLDAなどのトピックモデルなどのAI・機械学習の各種技術を組み合わせるなどにより分類し、分析しうる環境を整備し、さらに目録・画像等の素材データの洗練にもつなげていくデータ駆動型史料研究を目指している。データ駆動型研究基盤を目指す上で、データの整備、連携、および、利活用について整理していく必要があると考えている。

データ整備では以下について取り組む。

- AIによる文字認識のためのくずし字学習データセットの構築：AIを用いて文字認識し、ディープラーニングによって精度を高めるサポートツールの開発する。
- 文字認識AIによる構造化テキストの自動生成：100年以上1200冊近く刊行した基幹史料集と索引のデータ化する。史料原本・写本類の手書き史料の文字認識にも取り組む。
- 地名・人名等のシソーラス化・オントロジ化。これらは2節での取り組みを前提としており、さらに、編纂・出版という史料編纂所の研究成果を取り入れながら進めていくことになる。また、データ連携・利活用を意識し、汎用的なデータ整備を行う。この一例がJDCatへのデータ提供である。それ以外のシステムやツールでも利用可能にするため、汎用性のあるデータセットとしての公開を検討している。史料編纂所では、これまでに、人名・地名・文字をその出典とともに蓄積していく歴史情報リポジトリを構築している。ここに蓄積されたデータもデータセットの対象にする。

データ連携では、史料データに対する検索・関連付け・分類等の手法を検証し、この成果をSHIPSのミドルウェアに組み込み、実装していく。文献[7]はこの取り組みの先行研究と位置づけられる。データ連携の例を図5に示す。右上の点線で囲っている部分はデータ整備に相当する。異種のデータであっても、関連すると判断されるデータへつなげていき、そこからさらなる関連するデータへと数珠つなぎにしていくことができる仕組みを整える。歴史情報リポジトリはRDF Storeとして機能していることから、史料間、史料一人名間、史料-地名間データ間といったさまざまなデータ連携を行っていく。さらに、LDAを用いたトピックモデルやDNN・CNNなどのディープラーニング[12]を用いた分類結果にとるデータ関連付けを援用した結果を返戻できる仕組みを整える。

データ利活用は、SHIPS DB等で提供するアプリケーションとして位置づけられる。具体的な取り組みとして[13]があり、『倭寇図巻』や『正保琉球国絵図』などの絵図画像をもとに、IIIFやTEIを利用することで、画像・地名・テキスト・目録をデータ駆動的に結びつける手法を提示している。

現在、本プロジェクトの成果を反映したサービスを提供していく基盤を整備するべく、SHIPS DBの改修に着手しており、2021年度末に新たなSHIPS DBの公開を皮切りに、その後もSHIPS上にて成果を反映していく予定である。

¹ 原蔵者による設定を基本にしている。

² <https://mojiportal.nabunken.go.jp/>

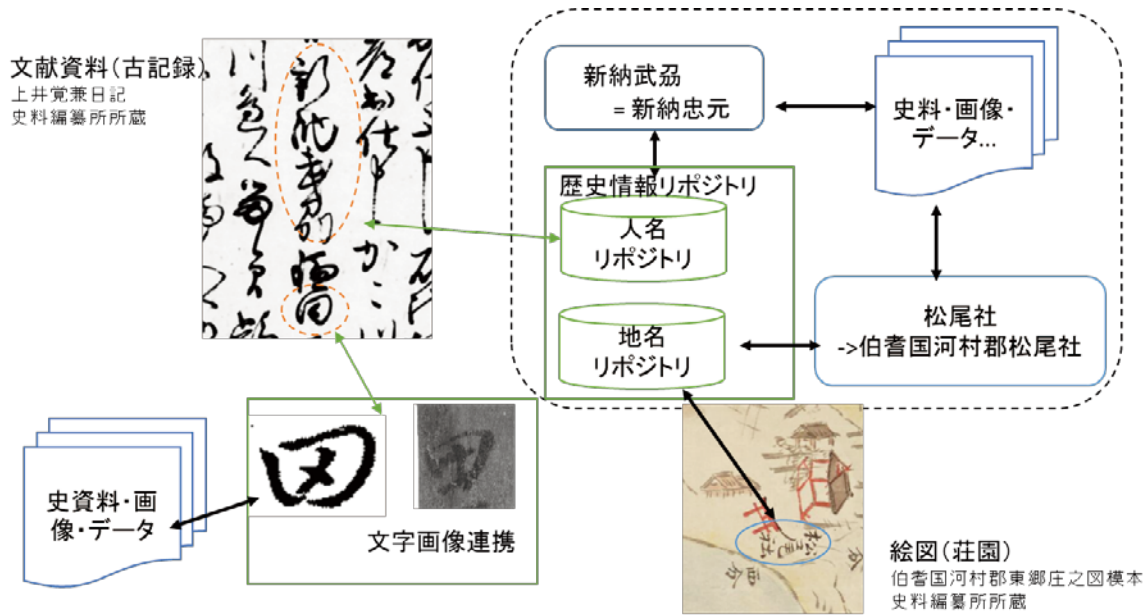


図 5 データ整備と連携の例
Figure 5 An example of data maintenance and collaboration

4. 課題と展望

ここでは、史料編纂所における史料データ提供基盤での取り組みにおける課題とその展望について述べる。

研究データ管理の議論は、地球科学や生命科学などが先駆けて実施されてきた。今では、さまざまな分野や観点にてデータ駆動型研究の推進がなされていく中、あらゆる分野で研究データの適切な公開が求められている。第6期科学技術・イノベーション基本計画から、これは人文学においても適用されていくことになる。ヨーロッパでの European Open Science Cloud やアメリカにける Big Data to Knowledge プロジェクトなどでは、研究データ公開基準として FAIR 原則 (FAIR Data Principles) [14]をデータ共有と管理の基盤としている。FAIR 原則では、Findable (見つけられる)、Accessible (アクセスできる)、Interoperable (相互運用できる)、Reusable (再利用できる) をキーとして、データ公開の実施方法を提唱している。FAIR 原則に従って、現在の史料編纂所での取り組みを検討する。

Findable では、(メタ) データがグローバルで永続的でユニークな識別子を持ち、データは豊富なメタデータで記述され、(メタ) データが検索可能な資源に登録あるいはインデックス化されることが求められる。JDCat への登録を経ることで、“ある程度”はクリアしていると考えているが、識別子付与という点は未だ達成していない。SHIPS 上でパーマリンクを設定してあるため、“限定的は世界”での識別子は保持しているが、DOI を付与するまでには至っていない。国文学研

究資料館新日本古典籍総合データベースでの DOI 付与の実装から、この付与は避けられないと考えている。

Accessible では、標準的な通信プロトコルで識別子を使って取得でき、データが利用不可能となったとしても、メタデータにはアクセスできることが求められる。進捗管理システムや Hi-CAT・Hi-CAT Plus での取り組みにより、データの長期保存・長期利用を行っており、現在改修中の SHIPS DB ではメタデータへの直接的なアクセスを可能にするための Web API を用意していることから、達成できると考えている。

Interoperable では、(メタ) データの知識表現として形式的で、アクセス可能であり、共有可能であり、広く適用可能な記述言語を使うことや、(メタ) データが他の (メタ) データへの適切な参照を持つことが求められる。これは史的文字データベース連携検索システムでの実証から達成しているものと考えている。

Reusable では、メタ (データ) が正確な関連属性を豊富に持つこと、たとえば、(メタ) データが、明確でアクセス可能なデータ利用ライセンスと共に公開され、その由来とつながっており、分野・コミュニティの標準を満たすことが求められる。これは進捗管理システムや収集史料公開のフローから達成しているものと考えている。

上記の FAIR 原則以外にも下記の課題があると考えている。

日本史研究者には史料編纂所の活動および提供している研究資源についてはよく知られているが、それ以外の研究者、人文科学に関わる研究者であっても、その利用方法は知られていない。

そのため、日本史に関わるデータがどこにあって、どのように利用できるものかが想像しづらいのが現状であろう。これは史料編纂所だけではなく、文書・記録等を所蔵している各機関に共通する課題である。史料編纂所では先にあげた人文学・社会科学データインフラストラクチャー構築推進事業への参加は大きな契機であり、他機関所蔵史料を史料編纂所所蔵史料とともに web 公開をしていくことでさらなる膨大な日本史史料コレクションとして構築していくことで日本史史料の Findable なアクセスポイントとして成長していく必要があると考えている。

高付加価値化が第6期科学技術・イノベーション基本計画では求められていることを鑑みると、異分野連携の“架け橋”にもなり得るような取り組みもチャレンジしていく必要がある。データ駆動型な取り組みを経て、史料データのさらなる成長も重要になると考えている。たとえば、料紙研究などモノとしての史料データの調査により、史料としてのデータがより高付加価値のあるものへと成長していくことが考えられる。

史料データの作成者のほとんどは日本史研究者であることから、日本史研究者にとってもその研究において利用可能なデータである必要もある。AI・機械学習を経て分析をし、日本史研究に役立てる上で、データサイエンスに関わる取り組みが不可欠である。しかしながら、日本史でのデータサイエンスは未熟であり、しかも周知されていない。データサイエンスの浸透を目指すためには人材が必要である。そのためには教育・若手育成も不可欠であると考えている。我々は日本史研究におけるデータサイエンスとしての取り組みにも寄与していきたい。

5. おわりに

本論文では、史料編纂所にて実践している、研究資源としての史料データを提供していくための基盤整備の取り組みを紹介した。

日本史史料データの収集・管理・分析等に関する研究を日本史研究において欠かせない研究基盤として位置づけるためには、システム構築のみならず利用者（日本史研究者）への呼びかけも非常に重要であると考えている。歴史学のみならず人文学ではデータが陳腐化することは稀であることから、データをいかに永続していくかも非常に重要であると考えている。

研究データ管理では、研究・編纂、研究成果公開といった研究フローにおいても十分なデータ提供・共有環境を目指していくことも必要である。研究・編纂・成果公開にて生成・共有されるデータが、研究資源たる史料データへフィードバックすることも十分にありえることから、史料データが循環する史料データ流通が実現すると考えている。この整備と実証を経ると、これは日本史研

究デジタルトランスフォーメーションの1つの解に成り得るはずだと考えている。

謝辞

本研究の一部は JSPS 科研費 18H03576, 18H05221, 20H00017, 20H00010, 21H04376, 21H04356, JSPS 人文学・社会科学データインフラストラクチャー構築推進事業「拠点機関におけるデータ共有基盤の構築・強化委託業務」、および、東京大学 FSI 事業「データ駆動型歴史情報研究基盤の構築」の助成を受けたものである。

参考文献

- [1] “第132回(2021年春季)東京大学公開講座「データ」講義要項”.
<https://www.u-tokyo.ac.jp/publiclectures/content/poster/kougiyoukou132.pdf>, (参照 2021-11-01)
- [2] “データ活用社会創成プラットフォーム mdx を導入”.
https://www.u-tokyo.ac.jp/focus/ja/press/z0310_00027.html, (参照 2021-11-01)
- [3] “科学技術基本法等の一部を改正する法律”.
https://www8.cao.go.jp/cstp/cst/kihonhou/kaisei_honbun.pdf, (参照 2021-11-01)
- [4] “科学技術・イノベーション基本計画”.
<https://www8.cao.go.jp/cstp/kihonkeikaku/6honbun.pdf>, (参照 2021-11-01)
- [5] “人文学・社会科学データインフラストラクチャー構築推進事業拠点機関におけるデータ共有基盤の構築・強化委託業務 令和元年度(2019年度)公募要領”.
https://www.jsps.go.jp/j-di/data/R1data_infrastructure_koubo.pdf, (参照 2021-11-01)
- [6] “人文学・社会科学総合データカタログ「JDCat」運用を開始”.
https://www.jsps.go.jp/j-di/data/JDCat_NII_20210716_ja.pdf, (参照 2021-11-01)
- [7] 山田太造ほか. 日本史史料データ流通基盤に向けた歴史データリポジトリの整備, じんもんこん 2019 論文集, 情報処理学会, pp.3-10, 2019.
- [8] 山田太造. 収集史料の体系化と永続的な利用に向けた管理, 研究報告人文科学とコンピュータ(CH), 情報処理学会, Vol.2018-CH-118, No.1, pp.1-4, 2018.
- [9] CCSDS. Reference model for an Open Archival Information System (OAIS),
<https://public.ccsds.org/pubs/650x0m2.pdf>, (参照 2021-11-01)
- [10] 杉本重雄. デジタルリソースの長期保存に関する概観. iPRES2017, 2017,
http://www-archive.cseas.kyoto-u.ac.jp/ipres2017.jp/wp-content/uploads/jtutorial_sugimoto.pdf (参照 2021-11-01)
- [11] 渋谷綾子ほか. 日本史史料の長期利用とデータ共有・連結化に向けたシステム環境整備, じんもんこん 2020 論文集, 情報処理学会, pp.23-30, 2020.

[12] 鬼塚洋輔ほか. 花押類似検索のための畳み込みオートエンコーダによる画像特徴抽出. じんもんこん 2018 論文集, 情報処理学会, pp.257-262, 2018.

[13] 中村覚ほか. データ駆動型歴史情報研究基盤の構築に向けた知識ベースの構築とその活用: 絵図史料を対象として. じんもんこん 2021 論文集, 情報処理学会, 2021 (tentative).

[14] 武田英明. 研究データ利活用に関する国内活動及び国際動向について. 第 2 回 SPARC Japan セミナー,

https://www.nii.ac.jp/sparc/event/2016/pdf/20161026_doc6.pdf (参照 2021-11-01).