

東京大学史料編纂所研究成果報告二〇二四―一

日本近世史料学の再構築

基幹史料集の多角的利用環境形成と社会連携を通じて

日本学術振興会科学研究費補助金基盤(B) 22H00692/23K21964

■ 目次 ■

目次	2
第一部 研究の概要	3
研究のねらいと研究方法	4
研究組織	5
各年度の成果概要	6
第二部 研究の成果	8
(1) シンポジウム / 国際研究集会	9
‘22 幕末維新时期研究と井伊家史料	9
‘23 近世初期における「御国」と「公儀」	13
‘24 デジタル化と歴史研究の未来	17
[趣旨説明] シンポジウム「デジタル化と歴史研究の未来」開催に当たって ...	18
[報告]「編纂知」の DX—大日本維新史料『井伊家史料』・『松平昭休往復書翰留』 編纂とデータベース構築の経験から—	19
[報告] 社会科学から見た歴史データ	23
[報告] 経済史研究における基盤的数量データの公開・活用	28
[報告] 歴史資料のデジタル化を支える技術	32
[コメント] 人文学における研究資源のデジタル化と持続可能性	36
[コメント] コメント	38
ラウンドテーブル・質疑応答	39
(2) デジタライジング研究会	41
研究会の概要	41
報告の記録	41
(3) 公開データ	48
1. 近世史編纂支援データベースの拡充	48
井伊家史料／松平昭休往復書翰留／細川家史料／市中取締類集	
2. 幕末維新史料・横断検索システム	51
(4) 論文・口頭報告等	52

PART

研究の概要

1

■ 研究のねらいと研究方法 ■

研究代表者 杉本 史子

■ 1. 研究のねらい

古文書学が権利文書の真偽判定と切り離せないかたちで形成されてきた西欧に対し、日本においては、近代的古文書学の構築が、近代国家をはじめとする公私の史料編纂と一体不可分のかたちでなされ、発展してきた。ここでいう史料編纂とは、史料の収集と解析の上で研究活用可能な状態に整えた史料集の刊行に至る、一連の過程のことである。

近代日本における代表的史料編纂事業は、明治政府が、日本の正史を叙述する修史事業を行う史料編輯国史校正局を1869（明治2）年に開設したことに始まる。同事業は、現在の東京大学史料編纂所へと受け継がれ、既に約1200冊もの史料集が出版されている。こうした史料編纂事業における、史料の収集・解析作業、及び研究活用に資する史料集としての版面校正作業などを通じて、史料そのものの緻密な分析が行われ、古文書学の知見が経験知というかたちで膨大かつ精緻に蓄積されてきた。個々の史料群の固有の構造、それを構成する個々の史料の個性を、紙面という限られたスペースと、本という秩序のうえにどのように表現するかについての経験知が豊富に蓄積されている。しかし、それらの「編纂知」は、体系化された成果として、学会・社会に示されたことはなかった。

また、近世古文書学構築における固有の困難さとして、古代・中世の古文書学の見方では、近世の、質的にも身分的にも拡大し多様化した史料を体系化できない点があげられる。かつて高木昭作は、このような状況に対して「近世史研究にも古文書学は必要である」（稲垣泰彦他編『中世・近世の国家と社会』東京大学出版会、1986）との警鐘を鳴らした。しかし、高木が指摘した状況は、こんにちでもなお克服されたとはいえない。

史料編纂に関わる知の蓄積を生かし、近世史料学の独自の体系化をいかに構築するか、これが本研究の核心的「問い」である。

■ 2. 方法

本研究では、近世史料研究を深化させ、史料そのものについての開かれた議論をも可能とするために、データベースの整備を行った。

研究基軸の中心にすえたのは、東京大学史料編纂所で公開中の「近世史編纂支援データベース」（以下「近世史DB」）の拡充・発展である。このデータベースは、国内の諸機関の学術データベースのデザインに関わってきた横山伊徳氏らが、第一に文書の作成主体と文書蓄積との関係についての史料学・アーカイブズ学研究成果をふまえ、第二に多様な刊行史料集の特性を十分考慮しながら、東京大学史料編纂所近世史料部門が編纂してきた基幹的史料集について効果的に検索できることをめざしたものであり、これ自体が研究成果の結実といえる。

本研究においては、次の視点から、この研究成果を次の段階へと深化・発展させることを企図した。紙媒体の史料集は、その編纂者が掲載内容を「選択」し、史料名称・掲載順序・掲載体裁の整理・決定といった一連のモノ化のプロセスによって「構成」したものであり、史料集はその

編纂者の史料理解を示したものといえる。一方で、「選択」されなかった史料情報が必然的に生じてしまうことや、編纂者による「構成」以外の視点での史料検討が難しくなってしまう側面がある。そのような紙の史料集の内容をデジタルデータとすることで、この「選択」の限界を打ち破り、「構成」に規定されたものを超えた多角的・複眼的な情報をそこに付加することができる。この過程では、デジタルデータもまた、「選択」「標準化」の操作のもとに生成されることが念頭におかれなければならない。

このように、本研究では、単なるデジタル化ではなく、紙史料集とデジタル化されたデータのふたつを両輪のように動かしていき、このレベルを実現したうえでの、研究・討論・再解釈化を企図した。この過程では、紙媒体の史料集の「選択」「構成」の再解釈も図られる。本研究でのデジタルライゼーションとは、この研究過程全体を指している。

■ 研究組織 ■

■ 研究代表者

杉本 史子（東京大学名誉教授・東洋文庫専任研究員、〈採択時〉東京大学史料編纂所教授）

■ 研究分担者

鶴田 啓（東京大学名誉教授、〈採択時〉東京大学史料編纂所教授）

箱石 大（東京大学史料編纂所教授）

林 晃弘（東京大学史料編纂所准教授、〈採択時〉東京大学史料編纂所助教）

立石 了（東京大学史料編纂所助教）

中村 覚（東京大学史料編纂所助教）

母利 美和（京都女子大学文学部教授）

稲葉 継陽（熊本大学永青文庫研究センター教授）

菊地 智博（東京大学史料編纂所助教〈2023-2024 年度〉）

■ 研究協力者

渡辺 恒一（彦根城博物館学芸員）

山田 太造（東京大学史料編纂所准教授）

高垣 亜矢（東京大学史料編纂所学術専門職員〈2022-2024 年度〉）

竹村 到（東京大学史料編纂所学術専門職員〈2022-2023 年度〉）

林 大樹（東京大学史料編纂所学術専門職員〈2022 年度〉）

菊地 智博（東京大学史料編纂所学術専門職員〈2022 年度〉）

黒滝 香奈（東京大学史料編纂所学術専門職員〈2022-2023 年度〉）

世川 祐多（東京大学史料編纂所学術専門職員〈2023-2024 年度〉）

高橋 直大（東京大学史料編纂所学術専門職員〈2024 年度〉）

■ 各年度の成果概要 ■

立石 了

各年度について、本研究の3つの柱、①デジタルライゼーションによる史料研究、②史料所蔵機関との連携強化による社会連携、③研究交流による研究成果発信の3者に即して概要を記す。

■ 2022 年度

①においては、データベース登載用データの作成と公開、データベース改良方針の策定、新データベースの構築・公開を行った。まずデータベース登載用データの作成と公開については、史料編纂所ホームページから公開されている「近世史編纂支援データベース」に登載するデータとして、『大日本維新史料 類纂之部 井伊家史料』索引型データの作成を行い、12,151レコードのデータを登録・修正して公開したほか、『大日本近世史料 細川家史料』第18巻の本文データを作成・公開した。また、『井伊家史料』の未刊行原稿381件のデジタルデータ作成も行った。データベース改良については、③の成果でもあるデジタルライジング研究会での知見をもとに、方針の策定を行った。さらに、同研究会でのOCR技術活用に関する報告をもとに、幕末維新期の史料集を横断的に検索できる「幕末維新史料・横断検索システム」の構築・公開を実現した。

②については、公益財団法人永青文庫と同文庫所蔵史料のデジタル公開方法について協議を深め、公開の準備作業として目録データの作成を進めたほか、同文庫所蔵史料に関するシンポジウムの開催について、熊本大学永青文庫研究センターと協議した。また、彦根城博物館と協力のうえ、③の成果である国際研究集会「幕末維新史研究と井伊家史料」を開催した。

③については、まず10月に国際研究集会「幕末維新史研究と井伊家史料」をオンラインにて開催し、基幹的史料集である『井伊家史料』についての理解を深め、最新の研究成果を共有・発信した。12月には史料編纂所・同所維新史料研究国際ハブ拠点形成プロジェクト主催の国際研究集会「維新史料研究と国際発信」開催に協力して、同プロジェクトが進める幕末維新期貴重史料の研究資源化や幕末維新史研究の英訳グロッサリー研究といった成果を共有した。さらに、デジタル技術やその課題に関する理解を深めるために、デジタルライジング研究会を4回開催し、最新の知見を共有しただけでなく、①で触れたデータベース改良方針の策定や新データベースの構築・公開へとつなげた。

■ 2023 年度

①については、史料編纂所近世史編纂支援データベースを充実すべく、前年度に引き続きデータの作成・登録・公開を進めたほか、小規模な改良を実施し、今後のさらなる改良方針の策定を行った。近世史編纂支援データベースへのデータ登載に関しては、『大日本維新史料 類纂之部 井伊家史料』について、17,051レコードの索引データを作成・登録・修正して公開したほか、『大日本近世史料 細川家史料』第17巻の本文データを作成・公開し、さらに『大日本近世史料 市中取締類集』の目次データ896件を作成した。『井伊家史料』については、前年度に引き続き未刊行原稿について843件のデータ化も行った。データベースの改良については、索引型のインターフェースを改善する改良を実施した。また、③の活動の一つであるデジタルライジング研究会で得

た知見をもとに、データベース改良方針を策定した。

②については、公益財団法人永青文庫と同文庫所蔵史料のデジタル公開方法について協議を深め、公開の準備作業として目録データの作成を進めたほか、③と関わって同文庫所蔵史料に関するシンポジウム「近世初期における「御国」と「公儀」」を熊本大学永青文庫研究センターと共同して開催した。

③については、熊本大学永青文庫研究センターと共に9月にシンポジウム「近世初期における「御国」と「公儀」」を対面・オンラインのハイブリッドにて開催し、中世から近世にわたる非常に貴重で稀有な史料群である永青文庫史料についての理解を深め、最新の研究成果を共有・発信した。また前年度に引き続き、史料編纂所・維新史料研究国際ハブ拠点形成プロジェクトとの協力関係を深め、12月には史料編纂所・同プロジェクト主催の国際研究集会「維新史料研究と国際発信」開催に協力したほか、同プロジェクトが進める幕末維新时期貴重史料の研究資源化や幕末維新史研究の英訳グロッサリー研究といった成果を共有した。さらに、昨年度開催の国際研究集会「幕末維新史研究と井伊家史料」の成果を『東京大学史料編纂所研究紀要』第34号に掲載して、その成果を学会に広く発信した。そして、前年度に引き続きデジタライジング研究会を4回開催して、デジタル技術やその課題に関する理解を深める、最新の知見を共有した。

■ 2024 年度

①については、まず史料編纂所近世史編纂支援データベースを充実すべく、前年度に引き続きデータの作成・登録・公開を進めた。具体的に追加したデータとしては、『大日本維新史料 類纂之部 井伊家史料』について、14,330件のデータを作成・登載したほか、新たに『大日本維新史料 類纂之部 松平昭休往復書翰留』のデータ3,222件を作成・登載した。さらに、本科研で得た知見を還元すべく、昨年度に策定したデータベース改良方針に従い、ユーザーインターフェースを中心としたデータベースの改修を実施した。

②については、昨年度に引き続き公益財団法人永青文庫と同文庫所蔵史料のデジタル公開方法について協議を深め、公開の準備作業としてこれまで本所が撮影してきた永青文庫の史料画像1点ごとの目録データの作成を進めた。

③については、デジタライジング研究会を開催する中で積み重ねてきた知見を広く学会に発信・還元するために、9月にシンポジウム「デジタル化と歴史研究の未来 ―人文学・社会科学の協働―」をハイブリッド形式にて開催し、歴史学関係者のみならず、経済学や情報学、図書館・資料館やデジタルアーカイブ関係の方々の参加を得て、学際的に成果を共有することができた。そして、前年度に引き続きデジタライジング研究会を3回開催して、最新の知見を共有し、デジタル技術やその課題に関する理解を深め、また今後のデータベースの在り方について議論した。

PART

研究の成果

2

1

シンポジウム/国際研究集会

本章では、本科研において 2022 年度～ 2024 年度に実施した 3 回のシンポジウム・国際研究集会の概要と成果について紹介する。

■ '22 幕末維新时期研究と井伊家史料 ■

■ 概要

- 開催日程：2022 年 10 月 1 日（土）13:30 ～ 18:00
- 開催方式：オンライン（Zoom ミーティング）
- 参加者数：138 名
- 主催：東京大学史料編纂所
- 共催：JSPS 科学研究費補助金基盤研究（B）「日本近世史科学の再構築—基幹史料集の多角的利用環境形成と社会連携を通じて」／東京大学史料編纂所維新史料研究国際ハブ拠点形成プロジェクト
- 協力：JSPS 科学研究費補助金基盤研究（A）「在外日本関係史料の調査と貴重史料の研究資源化による維新史料研究国際ハブ拠点の形成」（20H00023）

■ プログラム

司会挨拶（箱石大）

所長挨拶（本郷恵子）

趣旨説明（杉本史子）

13:45-15:45 第一部

菊地久氏（北海学園大学名誉教授）「幕末の幕府政治と井伊直弼」

朴薫氏（ソウル大学校教授）「幕末期幕府権力をどう見るか—「幕閣」という謎」

16:00-18:00 第二部

宮地正人氏（東京大学名誉教授）「『井伊家史料』編纂と私」

母利美和氏（京都女子大学教授）「『井伊家史料』と明治維新」

閉会挨拶（杉本史子）

■ 関連リンク

» 開催報告（UTokyo Focus）

https://www.u-tokyo.ac.jp/focus/ja/articles/z0206_00021.html

» 『東京大学史料編纂所紀要』34 号（2024）：本国際研究集会を原稿化し掲載

<https://www.hi.u-tokyo.ac.jp/publication/kiyo/34>

開催の記録

立石 了

2022 年 10 月 1 日（土）、国際研究集会「幕末維新时期研究と井伊家史料」を開催した。

この研究集会は、『大日本維新史料 類纂之部 井伊家史料』（以下『井伊家史料』）の出版が 2019 年の第三十巻で完結したことを記念して企画された。もともとは 2020 年 3 月に開催を予定していたものの、新型コロナウイルス感染症の拡大によって中止を余儀なくされ、今回オンラインに形式を変更して開催したものである。

幕末維新史料 類纂之部 井伊家史料 完結記念 国際研究集会

2022年10月1日(土)
13:30 ~ 18:00 (予定)
オンライン開催 (Zoom) 予定

13:30開演 (13:00開場)
所長挨拶 本郷恵子
趣旨説明 杉本史子

幕末の幕府政治と井伊直弼
菊地 久 北海学園大学名誉教授

**幕末期幕府権力をどう見るか
—‘幕閣’— というなど**
朴 薫 ソウル大学校教授
《15:45-16:00 休憩》

『井伊家史料』編纂時に学んだこと
宮地正人 東京大学名誉教授

『井伊家史料』と明治維新
母利美和 京都女子大学教授

主催：東京大学史料編纂所

共催：科学研究費補助金・基盤研究 (B)
「日本近世史科学の再構築—基幹史料集の多角的利用環境形成と社会連携を通じて」(22H00692)
東京大学史料編纂所 維新史料研究国際ハブ拠点形成プロジェクト
協力：科学研究費補助金・基盤研究 (A)
「在外日本関係史料の調査と貴重史料の研究資源化による維新史料研究国際ハブ拠点の形成」(20H00023)

参加のお申し込みは、下記 QR コードまたは URL より、専用フォームにご登録ください。

登録された方へ、集会の数日前をめどに、事務局よりレジュメ・配信 URL 等をメールにてお送りします。

問い合わせ先：
(事務局・菊地〈学術専門職員〉)

『井伊家史料』は、現在国指定の重要文化財として彦根城博物館に所蔵されている「彦根藩井伊家文書」を編纂・出版したものである。「彦根藩井伊家文書」は、江戸時代を通じて彦根 30 万石を領した井伊家に伝来した文書群であり、江戸時代中期から幕末を中心として、井伊直弼大老在任中の文書のほか、藩主家文書、彦根藩政文書といったものからなる約 2,7000 点の文書が残されている。長らく未公開であったが、1959 年から刊行が開始された『井伊家文書』によって井伊直弼関係史料が、そして 1978 年度から彦根市が行った史料調査報告書によってその全体像が、学会・社会に共有されることとなった。

この研究集会は、『井伊家史料』についての理解を深め、広く活用されることで、幕末維新时期研究の発展につなげていくことを目的として、二部構成とした。第一部では、『井伊家史料』をはじめとする幕末の諸史料から 19 世紀半ばの日本政治のありよう及び特質を問い直し、第二部では『井伊家史料』について編纂者としての立場、及び井伊家・彦根藩研究者としての立場から見つめるものである。第一部では北海学園大学名誉教授の菊地久氏とソウル大学校人文大学東洋史学科教授の朴薫氏とから、第二部では東京大学名誉教授の宮地正人氏と京都女子大学文学部教授の母利美和氏とから、それぞれご講演いただいた。司会は本所教授の箱石大氏が担当した。

研究集会の冒頭では、本郷恵子所長からの挨拶、本所教授の杉本史子氏による趣旨説明が行われたのち、各講演が行われた。

まず第一部の菊地氏による「幕末の幕府政治と井伊直弼」では、研究史を整理しながら自身の研究の出発点、問題意識に触れたのち、第一に、表方における合議分掌と奥向きでの裁可承認という多極分権的な幕府の意思決定システムにおいて、井伊直弼やそれに先行する阿部正弘・堀田正睦がどのように政治的リーダーシップを発揮したのか、第二に、諸侯諮問に見える合議分掌という幕府の政治システムの間口開放は、政治課題に対する最適解抽出方法ではあっても繰り返されるとかえってリーダーシップの弱体化を招きかねず、その性格をどう評価するのか、という二つの論点が提示された。そして、井伊直弼について、そのリーダーシップの発揮方法が、既存の制度を前提としてその枠内でできることを追求する制度的リアリズム、そして制度が崩壊しつつあり残骸となりつつある中で発揮される状況下の中のリアリズムという政治的リアリズムの二タイプのどちらでも評価できないことを指摘した。

続く朴氏の「幕末期幕府権力をどう見るか―「幕閣」という謎」では、一般に変革期の研究においては開明派と守旧派との対立構造が暗黙裡に設定され、歴史の勝者となる前者に注目が集まりやすい傾向があり、幕末維新时期についても、開明派が幕政に参与する機会を得て改革を試みたものの、守旧派の反発によって失敗したことで明治維新を迎えるという叙述が多くみられ、薩長に代表されるいわゆる雄藩や井伊直弼については研究蓄積が多くみられるのに対して彼らを挫折させた幕閣に関する研究が薄いという課題が提示された。そのような挫折の具体例として徳川斉昭の幕政参与や井伊直弼の大老就任を挙げ、立場や意見は異なるものの、共に幕閣に対して意思を押し通すことが難しい状況が繰り返されていたことが指摘される。こうした「守旧派」幕閣が強固に存在し続けたことの一因は、藩政と比較すると主君親裁がなく、激論が度々交わされるような御前会議もなかったことに求めうるのではないかという見通しが示された。そして大身の家門や譜代大名、雄藩の外様大名を差し置いて、小身の譜代大名が老中として幕政を専担する体制自体、中国や朝鮮と大きく異なっており、小身の老中集団の手に余る国家的危機に際して幕政参与や大老、政事総裁職が幕政に参加するものの旧来の幕閣が権力を維持できた事情を明らかにすることで、東アジアの諸国家と異なる近世日本権力の特徴が見いだせるのではないかと論じた。

休憩をはさんで行われた第二部の宮地氏による『井伊家史料』編纂と私』では、実際に『井伊家史料』の編纂に携わった経験から、新たな近世国家論や幕府像を持つに至った経緯、そして今後の課題が提示された。史料編纂所入所前に近代史を専攻し、近世における天皇・朝廷を古代の残滓、或いは単なる宗教的権威と見ていたところに、近世政治史でも最も難しい諸課題が噴出して安政五年の『井伊家史料』編纂の担当を命じられたことで、それまで幕末特有の問題と認識していた公武合体というあり方が、幕末に限定されるものではなく、幕府は朝廷と合体することによって国家権力たり得ていたのではないかと考えるようになり、自身の近世史研究の出発点となったことが紹介された。そして、公家史料にあたり、そこから公武合体を考えていくなかで、全国の社寺への位階受領の事例から、朝廷は一つの国家を成り立たせる不可欠の要素であり、最幕末に分裂していくまで、朝幕関係を前提として初めて幕府が国家として存在していたという考えに至ったこと、さらに『井伊家史料』編纂に従事するなかで、幕府が情報収集・集約するシステムの強固さに触れ、「幕藩制国家論」が提示する二重国家論に疑問を抱くようになり、幕府そのものが国家権力であるという見方をするようになったことが示された。そのうえで、幕末に公議輿論という考え方が広まっていくなかで、幕府がどのように最後まで権力を維持しようとしたか、

そのためにどのような形を展開していこうとしていたか、そこに幕臣・旗本のどのような層がリーダーシップをとって遂行しようとしていたかが残された課題として提示された。

最後の母利氏による『井伊家史料』と明治維新においては、18年間彦根城博物館の学芸員として「彦根藩井伊家史料」に携わられた経験から、明治維新以降の伝来過程とそこにおける改竄のあり方が明らかにされた。まず「彦根藩井伊家文書」の概要や大まかな伝来過程、直弼関係史料における文書の切り取りや断簡について紹介されたのち、明治14年以降の井伊直弼顕彰運動の盛り上がりと建碑挫折によって、直弼の事績を編纂物として世に広める方針が定められたこと、そこで「直弼伝」の略伝として編纂された「磯打浪」で孝明天皇や三条実美、松平忠固罷免に関わる史料・記事が削除・改竄されていること、「磯打浪」の脱稿後まもなく明治政府の修史館に提出された「公用方秘録」や編集写本史料にも同じ改竄が見られることが紹介された。そして、直弼没後や直弼伝編纂の過程で収集・筆写された文書が「彦根藩井伊家史料」の一部として今日に伝来していることで、本来藩主家ではなく藩士家に伝来していたであろう文書も一体的に見ることができ、さらに「彦根藩井伊家文書」に井伊直弼の受給文書だけでなく直弼が文書を発給するにあたっての草稿も伝来していることから、これらを取録する『井伊家史料』は幕末の政治過程について双方のやり取りを明確に分析することが可能な史料集であるという意義を有することも示された。

第一部の最後には発表者同士での質疑が、第二部の最後には発表者同士に加え参加者からも質疑が行われた。第一部においては、菊地氏発表に関して井伊直弼権力の評価について専制・専断の内実があったかどうかや、阿部正弘政権以降の幕府の対外政策決定過程の評価をめぐって議論されたほか、朴氏発表について東アジアの中で幕府権力を見る際に、「家門」の政治への関わりをどう捉えるのかについて、特に朝鮮と中国の支配層が家門として同一視されるのか、中国では科挙によって支配層に上昇していくが、その母体の郷紳層は朝鮮の両班の如き家門といえるのかが議論された。第二部においては、まず発表者同士の質疑として、宮地氏発表における幕府や近世国家の位置づけについて、及び母利氏発表における井伊直弼顕彰運動の論理に関して基礎的な確認が行われたのち、安政の大獄における意思決定過程や大老就任前の井伊直弼からの老中への交渉ルートについて議論が行われた。続いて参加者から、史料編纂所における史料集編纂と江戸時代の和学講談所以来の伝統との関わりについてや、井伊直弼顕彰運動における史料改竄問題に関連して修史館への史料提出の経緯や顕彰運動前後での井伊家側での記述の差異について、それぞれ疑問が提示され、発表者が応答した。

最後に、長く『井伊家史料』の編纂に携わってきた維新史料第一室教授の杉本氏より閉会の挨拶が行われ、本研究集会で示された『井伊家史料』の意義をより発展させることのできるツールとしての本所データベース「近世史編纂支援データベース」を紹介し、『井伊家文書』のさらなる多面的活用を願って研究集会を終えた。

本研究集会は、4時間半にわたる長丁場であったにもかかわらず、当日は138名もの参加者を得ることができたうえに、発表者同士の活発な議論が行われ、さらには全てを取り上げられない程に参加者から多くの質疑が寄せられ、大変に充実した研究集会となった。

(『東京大学史料編纂所報』58号、2023年より転載・改変)

■ ‘23 近世初期における「御国」と「公儀」 ■ —永青文庫細川家文書から—

■ 概要

■ 開催日程：2023 年 9 月 16 日（土）13:00 ～ 17:00

■ 開催方式：ハイブリッド

対面会場：熊本大学工学部百周年記念館

オンライン：Zoom ミーティング

■ 参加者数：146 名（オンライン 108 名、対面参加 38 名）

■ 主催：熊本大学永青文庫研究センター・東京大学史料編纂所

■ 共催：JSPS 科学研究費補助金基盤研究（B）「日本近世史料学の再構築—基幹史料集の多角的利用環境形成と社会連携を通じて」／JSPS 科学研究費補助金基盤研究（C）「永青文庫資料と「くずし字 AI-OCR」の活用による 17 世紀社会論・公儀権力形成史の再構築」（23K00842）／JSPS 科学研究費補助金若手研究「日本近世における政教関係の形成と確立」（21K13090）／熊本史学会

■ プログラム

・林晃弘氏（東京大学史料編纂所准教授）

「細川忠利の『公儀御書案文』と近世政治史研究」

・稲葉継陽氏（熊本大学永青文庫研究センター教授）

「元和～寛永期における「御国」統治と文書・記録 —ポスト戦国世代の領国支配—」

コメント 三宅正浩氏（京都大学文学研究科准教授）

■ 関連リンク

» 開催報告（熊本大学永青文庫研究センター）

<https://eisei.kumamoto-u.ac.jp/event/2023/09/post-40.php>

» 報告記事（『熊本大学新聞デジタル』2023 年 9 月 16 日）

<https://note.com/kmdishinbunsya/n/n2cd4884631e8>

» 報告記事（『熊本日日新聞』ウェブサイト、2023 年 9 月 17 日）

<https://kumanichi.com/articles/1174449>

■ 開催の記録

立石 了

2023 年 9 月 16 日（土）、シンポジウム「近世初期における「御国」と「公儀」 —永青文庫細川家文書から—」を開催した。

公益財団法人永青文庫が所蔵する細川家文書は、「和泉上守護家」の古文書をはじめ、確実な自筆文書を含む織田信長発給文書 59 通、そして全国的に屈指の規模を誇る近世藩政史料といった、中世から近世にかけての長期にわたり、かつ内容的にも多岐にわたる、非常に貴重な史料群である。

本シンポジウムの主催者である史料編纂所では、長きにわたり本史料群の調査を許され、『大日本近世史料』というシリーズで史料集を刊行し、近世初期の政治史研究において様々に活用されてきた。また、同じく本シンポジウムの主催者である熊本大学永青文庫研究センターでは、熊本大学附属図書館寄託分の悉皆調査と総目録を作成し、約 58,000 点に及ぶ全体像を明らかにしてきたうえに、「永青文庫叢書」というシリーズで 10 冊の史料集を刊行し、これまで知られてこなかった新史料を含め、興味深く重要な史料を広く学会に提供してきた。このような、着実な調査の実施と研究の基礎・基盤となる史料集の刊行によって、研究環境の充実が着実に進んできた状況にある。

今回のシンポジウムは、こうした取り組みのなかで蓄積してきた知見をもとに、近世初期の領国支配や幕府・諸大名との関係を検討し、「永青文庫細川家文書」の豊かな世界とそこから明らかになる研究成果とを広く発信して、今後の熊本藩研究や近世政治史研究をより前進させていくために企画・開催した。シンポジウムの構成は、「永青文庫細川家文書」の史料を検討した研究報告二本とコメント、そして報告者からのコメントへのリプライと参加者・報告者間の質疑応答とした。報告は史料編纂所の林晃弘准教授と永青文庫研究センターの稲葉継陽教授が行い、京都大学文学研究科の三宅正浩准教授がコメントに立った。司会は永青文庫研究センターの今村直樹准教授が担当した。また、開催方法としては熊本大学工学部百周年記念館での会場参加と Zoom ミーティングによるオンライン参加とを併用したハイブリッド開催であった。

研究集会の冒頭では、本郷恵子史料編纂所長からの挨拶が行われたのち、各報告が行われた。

まず林氏による「細川忠利の『公儀御書案文』と近世政治史研究」では、『大日本近世史料 細川家史料』一六巻以降の底本である『公儀御書案文』という史料を検討し、その史料性格を明らかにするとともに、そこに収録された細川忠利書状案に見ることができる幕府中枢の人物や近隣大名とのやり取りから、幕藩関係や大名間関係、すなわち『公儀御書案文』のタイトルの「公儀」と、領国支配との関係を明らかにして、近世政治史研究における書状留の有用性を改めて示した。具体的な内容としては、まず『公儀御書案文』の基礎的な情報を示したのち、収録された案文と現存する実際に発給された書状原本とを照合して、基本的に右筆が作成した文書のうち留書に収めるべきと判断されたもの、という『公儀御書案文』に収録される書状案文の性格を明らかにした。そのうえで、『公儀御書案文』に残る注記から、「奥」の空間からもたらされる「御案紙」をもとに記録されるほか、「御案紙」がもたらされずに実際に発給された文書を写したり、忠利御前で作成・発給されたために「御案紙」が存在せずに留められなかったりといった、具体的な作成方法を復元した。そして、『公儀御書案文』に収録された書状案文の中から、まず春日局宛書状案を取り上げて、幕府中枢にいた人物との情報交換と細川忠利の意思決定との関係を示した。次に寛永 15 年 9 月 13 日付キリシタン高札をめぐる大名間の情報交換と幕府法令の大名領国への浸透のあり方を明らかにした。

次に稲葉氏の「元和～寛永期における「御国」統治と文書・記録 ―ポスト戦国世代の領国支配―」では、永青文庫史料において元和期より「御国」という言葉が頻出するようになり、寛永末期の細川家当主代替わりに際しては「御家」と結びついた「御国家」が家老・奉行衆の方向の対象として挙げられるに至ることに注目し、「御国」(＝領国)支配の構造、特に「御家」の構成員である知行取家臣層が「御国」の再生産に奉公することを必然化させる統治の構造について、文書・

記録の作成・機能という点から考察し、「御国」支配に「御家」の主要構成員が平均的に関与する状況が、奉公の対象としての「御国家」概念成立の基礎的条件であったことを指摘した。具体的には、まず元和7（1621）年の細川忠興から同忠利への代替わりを契機として作成された文書・記録を検討し、忠利が百姓・庄屋層やその利害関係を反映した郡奉行からの上申文書を基軸にして「御国」統治の基礎を把握し、整除することに一定の成功をみたことを示した。次に忠利が参勤して国元不在の時の史料を検討して、国元での奉行衆・家老衆合議での審議とその報告・伺に対する遠隔地にいる忠利の最終決裁という絶え間ない交渉によって領国支配の重要案件についての意思決定が行われていたこと、さらに奉行所の記録から、「御国」支配に携わる数十の役職に任命された家臣が主体的に関与する体制が形成され、これが奉公の対象としての「御国」・「御国家」という意識を家臣団に早くから植え付けた要因であったことを指摘した。そして最後に、忠利在国中の史料について、まず上申・裁可文書の分析から、忠利が無下に却下したものではなく、また参勤準備中で一度裁可を停止した際にも、奉行衆や家老衆の要求によって「御国」の再生産に関わる決裁行為を継続させられていたこと、次に口頭命令記録である「奉書」の検討から、「御国」支配上の懸案について惣奉行や郡奉行を中心とした関係者の総意を前提として忠利の裁可が行われていること、また主従制に係る案件でも、「御国」での評価や家中世論が忠利の裁可に影響を与えていることを示した。

コメントに立った三宅氏からは、まず林報告について、『公儀御書案文』の史料性格を解明したうえで、その近世政治史研究での用い方の事例を紹介した意義を高く評価したうえで、『公儀御書案文』の作成段階や収録基準、選別方法の解明という追求すべき新たな課題も示されたことを指摘した。そこで、阿波徳島藩蜂須賀家文書の『草案』という類似した史料を比較対象として取り上げ、控である『公儀御書案文』に対して『草案』は本来下書であったものが徐々に控としての性格を強めていったものということを確認したうえで、同時期のそれぞれの蜂須賀宛・細川宛の史料件数を比較すると『草案』のほうが5～6倍ほど多く収録されていることを示した。そして、寛永期が武家社会の中で書状が飛び交っている時代であるものの、そのやり取りの相手には偏りがあり、そうしたネットワークを念頭に置きながら当時の武家社会を考えていく必要があることを指摘した。次に稲葉報告について、元和～寛永初期の細川家の家中・領国支配の有り様を文書から具体的かつ詳細に提示し、当該期の大名家における政治的意思決定過程を復元したものと評価したうえで、この時期の「藩政」の実態をこれほど詳細に復元できる細川家文書は稀で貴重な事例であることを指摘した。そして、大名家において家老合議を軸とした機構が形成され、また繰り返される飢饉的状况に大名・家臣が危機感を共有し、寛永飢饉を契機として地方支配機構が整備されていくなど、寛永期に近世大名家の政治機構が確立されるという一般的理解に対し、元和期からの事例が提示される本報告をどのように位置づけるべきか、という課題を提示した。また、細川家文書ほどには復元できなくとも、同様に「藩政」が展開されていたと思しき事例が土佐山内家にあることから、細川家と同様の展開が他の大名家でもあった可能性を示した。さらに、「御国」という言葉が当時どのように認識されていたのか、特に17世紀半ばに池田光政や藤堂高次が示した認識との関係について、そして細川家の藩政機構における役職・人数などの定着時期と今回の事例との関係といった検討課題や、史料の豊富さから大名家における意思決定構造の推移の具体的復元が期待できる点に触れたうえで、今回示された元和～寛永期の細川家における「藩

政」の実態を、中世・戦国段階や17世紀半ば以降のあり方と比較した時に、歴史の流れのなかでどのような段階と位置づけ、理解すべきかが最も大きなこれからの課題であると述べた。

最後に設けられた質疑応答の時間では、まず報告者からコメントへのリプライがあり、その後参加者からの質問に答える形で進化した。まず林氏から三宅氏へは『公儀御書案文』と『草案』との比較について、稲葉氏からは「藩政」の位置づけや「御国」概念、「御国」意識についてを中心に、それぞれ返答があった。

続く一般の参加者からの質疑においては、まず林報告に対して『公儀御書案文』の性格についてや「公儀」の語が意味するところ、キリシタン高札の目的や本報告での位置づけについて議論が行われた。稲葉報告に対しては、家中世論と細川忠利の意思決定との関係や「藩政」が展開される対象領域、地方知行制との関わりについて議論された。

全体討論では、まずコメントの三宅氏に対して、稲葉報告を時代の流れにどう位置づけるかという問題に関連して、時代の転換期の捉え方についての質問があった。その後、細川家での文書控の蓄積の前提や戦国期・織豊期からの連続性、控文書群の具体的利用例、控文書群の作成・保管の様相、奉行所における文書群蓄積の意味といった文書作成・蓄積に関する議論が行われ、最後に本シンポジウムのテーマに関連して、『公儀御書案文』の「公儀」の語の意味や収められた文書群が「公儀」と分類された意味、「御国」と「公儀」という語の関係をどうとらえるか、について議論が深められた。

質疑の締めくくりとして、コメントの三宅氏からは、二本の報告を聞き、またコメントを依頼されたことでの新たな発見や検討課題はもちろんとし、質疑の場で提起された「御国」や「公儀」といった史料用語の意味をどこまで考えるべきか、他方で当時の人々がどこまで意識して用いていたのか、改めて考えていきたい、林氏からは、約58,000点に及ぶ細川家文書の検討は大変な作業であるものの、今回のシンポジウムで一端が示されたような史料同士の関係性が明らかになり、史料がより利活用されていくことを願っている、稲葉氏からは、「御国」支配の体系性の復元はまだまだ不十分である一方で、永青文庫史料は「御国」支配と「公儀」との関連が元和期からどのように展開して17世紀半ばにつながっていくのかを体系的に追える史料群であることやそれを追求していくという長いスパンの課題を再認識できて感謝している、司会の今村氏からは、近世後期の研究者としての観点に立つと、近世中後期にも他大名家とのやり取りの案文集が作成されているように、永青文庫史料は通時代的に検討できる極めて良質な史料群であり、今回のシンポジウムの射程をさらに伸ばして考えていくという課題が示されたように思う、とそれぞれの登壇者から感想が述べられた。

最後に、熊本大学永青文庫研究センター長である稲葉氏から閉会の挨拶が行われ、熊本藩家老家文書である松井家文書に代表される周辺史料群も合わせた、永青文庫史料のさらなる活用や研究の進展、成果発信を誓って、本シンポジウムが終了した。

本シンポジウムは、会場での対面参加者・オンラインでの参加者合わせて150名弱もの参加者を得ることができたうえに、当初予定の4時間を20分超過するほど、発表者同士の活発な議論や多くの質疑が重ねられ、大変に充実したシンポジウムとなった。

(『東京大学史料編纂所報』59号、2024年より転載・改変)

■ ‘24 デジタル化と歴史研究の未来 ■

— 人文学・社会科学の協働 —

■ 概要

■ 開催日程：2024 年 9 月 7 日（土）13:00 ～ 17:00

■ 開催方式：ハイブリッド形式

対面会場：東京大学史料編纂所大会議室

オンライン：Zoom ミーティング

■ 参加者数：81 名（オンライン 65 名、現地参加 16 名）

■ 主催：東京大学史料編纂所

■ 共催：JSPS 人文学・社会科学データインフラストラクチャー強化事業（JPJS00320231001）

／ JSPS 科学研究費基盤研究（B）「日本近世史料学の再構築 — 基幹史料集の多角的利用環境形成と社会連携を通じて」／ JSPS 科学研究費基盤研究（B）「長期歴史ビッグデータを用いた日本の政治機構と都市構造の持続性の分析」（23K25516）／ JSPS 科学研究費基盤研究（S）「史料データセンシングに基づく日本列島記憶継承モデルの確立」（24H00011）／ JSPS 科学研究費基盤研究（B）「幕末維新时期史料研究の学術資源化推進と国際共有基盤の新展開」（24K00108）／ JSPS 科学研究費基盤研究（B）「日本型人材選抜システムの定量分析」（24K00265）／ JSPS 科学研究費基盤研究（B）「近世日本経済のデータ駆動型分析に向けた基礎的研究：物価・賃金、金利、貨幣両替相場」（24K00277）

■ プログラム

13:00-13:05 杉本史子（東洋文庫）[趣旨説明](#)

13:05-13:35 菊地智博（東京大学）

[「編纂知の DX—『井伊家史料』『松平昭休往復書翰留』編纂とデータベース構築の経験から」](#)

13:35-14:20 山崎潤一氏（神戸大学〈当時〉）[「社会科学から見た歴史データ」](#)

14:30-15:00 高槻泰郎氏（神戸大学）[「経済史研究における基盤的数量データの公開・活用」](#)

15:00-15:40 中村覚（東京大学）[「歴史資料のデジタル化を支える技術」](#)

15:50-16:20 コメント [山田太造氏](#)（東京大学）・[中村雄祐氏](#)（東京大学）

16:20-16:40 [ラウンドテーブル](#)

16:40-17:00 [質疑応答](#)

■ 関連リンク

» 開催報告（カレントアウェアネス -E）

<https://current.ndl.go.jp/e2743>

以下、各報告・コメントを報告者により原稿化した上で収録する。

趣旨説明

シンポジウム「デジタル化と歴史研究の未来」
開催に当たって

杉本 史子（東京大学名誉教授、東洋文庫専任研究員）

本シンポジウムは、本科研で開催してきたデジタライジング研究会の成果に基づき、企画された。紙史料集とデジタルデータとの有機的結合のもとに、歴史情報を公開するありかたを考察するものである。

以下、関係データベースを紹介しながら、各報告について、その概要を述べたい。

■ 神戸大学経済経営研究所「近世経済データベース」

<https://www.rieb.kobe-u.ac.jp/project/kinsei-db/>

神戸大学経済経営研究所と公益財団法人三井文庫が共同で作成・運営するもので、18世紀中期から明治初年にわたる、日々の米価、金銭比価、そして天候を採録する。高槻泰郎報告は、この経験の上に立ち、経済史研究における数量データのデジタル化について考える。

■ 東京大学史料編纂所「近世史編纂支援データベース」

<https://wwwap.hi.u-tokyo.ac.jp/ships/w30/search>

東京大学史料編纂所（古代～明治維新期の研究）の近世史料部門で蓄積されてきた編纂知を、紙史料集+デジタルデータとして公開することで広く学会・社会で共有し、編纂に還元するという、知の循環構造を形づくろうとするもの。本科研では、このデータベースのデータ拡充と機能改善をおこなってきた。

「編纂」とは、綿密な史料分析の上に立ち、活用可能な状態に整えた情報として公開するための、一連の過程のことを意味している。菊地智博報告が、編纂過程のデジタル化の試みについての、成果と課題を提示する。

■ 東京大学史料編纂所「幕末維新史料・横断検索システム」

<https://ishin.lab.hi.u-tokyo.ac.jp/>

史料編纂所前近代日本史情報国際センター HI-Lab から公開中。これまで刊行されてきた幕末維新期史料の横断的検索を可能とする、幕末維新史研究の新たな可能性を開くもの。このデータベースは、本科研のデジタライジング研究会における中村覚報告をきっかけにし、箱石大のサポートのもと、明治維新史学会例会での報告を経て、完成した。

中村覚報告は、この検索システムをはじめとする、歴史資料のデジタル化を支える技術について提言する。

そして、山崎潤一報告は、デジタル化をめぐり、研究分野を横断する開かれた議論への展望を、社会科学の立場から検討する。

これらの報告に対し、歴史史料のデジタル化を主導してきた経験を踏まえた山田太造コメント、文化資源学・人文諸学を見渡す広い立場からの中村雄祐コメントをお願いした。

本シンポジウムが、デジタル化と歴史研究の、実りある未来のための、確実な基礎作りとなることができれば幸いである。

報告 「編纂知」の DX

—大日本維新史料『井伊家史料』・『松平昭休往復書翰留』編纂とデータベース構築の経験から—

菊地 智博（東京大学史料編纂所）

■ はじめに

史料集編纂の過程では、様々な暗黙知・経験知が編纂機関に蓄積されてきた。本科研ではそれを「編纂知」として把握し、その析出および研究やデジタルデータへの応用を一つの目標としている。本報告では、編纂過程における編纂知の実例、さらにデジタルデータ作成における応用例を紹介する。また、「史料集編纂」の将来像への見通しにも言及したい。

■ 1. 編纂知とは何か

編纂知とは何か。これを具体的に述べるならば、〈原史料の構成要素を理解し、不可逆に分解・再構成して版面に落とし込む技術〉ということになろう。史料の解説・読解におけるノウハウのみならず、モノとしての史料や、文書自体の構造、また史料に含まれる人名に関する情報等について、多様な暗黙知が史料編纂所を含む編纂機関には蓄積されている。この「知」は言語化され、また研究に活用されるべき貴重な財産であるといえる。

ところで近年、紙の史料集を「デジタル」へ移行する流れが強まりつつある。出版不況や、研究活動自体がデジタル媒体によって行われるなか、この流れはごく自然なものである。史料編纂所においても、『九世紀編年史料』（データベース「[MIDOH](#)」にて公開）のように、史料集をデータベースに切り替えた例がある。今後「史料集編纂」というものを考える上では、作ったものを後からデジタル化するというよりは、デジタル化を当初から「前提」として考える必要があるだろう。

一方では、様々な史料所蔵機関において原史料の撮影画像をデジタル公開する取り組みも進みつつある。史料編纂所においても「データインフラ強化事業」によって所蔵史料のみならず他機関の画像公開ハブとしての機能を整えつつある。さらに、こうしてデジタル公開された史料を深層学習 AI による OCR にかかけ、容易に判読することも可能になった。代表的な取り組みとしては、国立国会図書館における古典籍 OCR や TOPPAN 株式会社による「ふみのは」等が挙げられ、大量処理のみならず、個人端末においてスタンドアローンに実行することも可能になりつつある。精度にまだ一定の課題があるとはいえ、画像さえあれば全文検索が簡単に実行できる状況において、果たして史料集やデータベースの作成にどれほどの必要性があるのか、という疑問も生じる。

結論を先に述べると、AI-OCR による全文検索は、史料集編纂やデータベース作成の価値を損なうものではない。文脈に即したテキスト理解がなければ史料を解釈することができないのは当然であり、また画像と OCR 文字列だけでは「モノとしての史料」のあり方——例えば端裏書や付箋、目下や脇付といったものの機能を踏まえた史料の理解は困難である。史料中では人名や地名の記され方も極めて多様であり、特に人名については官途名や符丁、偽名などの存在ゆえ単純な検索でカバーできるものではない。史料集は、注釈等の形でこれらの情報を補い、読者へ適切な補助線を提供してきた。「デジタル」へ編纂知を応用する意義が、ここに生じるのである。

ただし、編纂知の応用において注意すべきは編纂という行為の不可逆性である。史料という三次元立体を二次元の版面表現へ落とし込む過程では、どうしてもモノ実体における文脈を捨象せ

に挙げられるだろう。

『井伊家史料』に限らず史料編纂所の史料集とは、こうした経験知の集積によって成り立っているのである。

ここで、その活用例として「[近世史編纂支援データベース](#)」における『井伊家史料』索引型 DB を紹介したい。「近世史編纂支援データベース」は近世史料部門の編纂した史料集を構成要素ごとに検索可能にしたデータベースであり、特に人名注部分に関するものが「索引型」である。

『井伊家史料』索引型 DB とは『井伊家史料』全 30 巻に付された人名注、幕末政治史に関する膨大な人名情報を活用した人名索引であり、約 50 年間の編纂史において蓄積した人名に関する「編纂知」の結晶である。ここでは版面に表示した人名注情報に加え、刊行時には省略したもの・編纂後の研究進展によって比定可能になった人物などの情報を付加しており、例えば同じ史料中で二回目に出現し「同人」などと呼称したり隠語や符丁で述べられたりする場合もデータとして登録している。このことによって、単なる索引に終わらない人名データセットの形成が意図されている。

『井伊家史料』索引型 DB は本科研の資金によって 2024 年度末までに 23 巻分のデータが登録される見込みである。

■ 3. 史料集編纂の DX とその将来像

2023 年度より、史料編纂所では『大日本維新史料』の新シリーズとして『松平昭休往復書簡留』の編纂・刊行を開始した。松平昭休は水戸藩主・徳川斉昭の九男で、一時忍藩松平家に養子入りし世子となるものの安政大獄の影響で廃嫡、その後岡山藩池田家に養子入りして藩主・池田茂政と名乗り幕末政局に関わった人物である。この昭休が岡山藩養子入の直前まで、鳥取藩主池田慶徳など各藩に移った兄弟姉妹などと交わした書簡を写し留めたものが『松平昭休往復書簡留』である。原本は史料編纂所に所蔵され、所蔵資料目録 DB「Hi-CAT」で原本画像が全てインターネット公開されている。

本シリーズの刊行にあたっては、編纂作業のデジタル・トランスフォーメーション (DX) を試みた。『井伊家史料』編纂終盤から導入された電子入稿をさらに一歩進め、原稿データに「編纂知」を反映したタグをあらかじめ付与、人名や日付はもちろん、差出や宛所、付箋などといった文書特有の構造をマークアップしている。こうして機械可読性を高めることにより、組版とデータベースへの二次利用を図ろうというものである。

技術的な詳細は中村覚氏報告も参照されたい。簡単に具体的なワークフローを述べると、Word ファイルとしてタグ付き原稿を作成、ここから機械変換により組版用の TeX 作成と DB 用の人名抽出を行い、校正を相互に反映してゆく形とした。中村氏の報告に述べられる通り、将来的には TEI (Text Encoding Initiative) XML ファイルの自動作成と提供も本書目に限らず進めてゆく見込みである。

編纂知を考慮した構造化テキストの提供は、何を実現するのだろうか。機械処理性を高めるといって、大量処理による統計分析や、生成 AI の学習などというような流行の手法がまず思い浮かぶが、しかし意義はそれらにとどまるものではない。

第一に、機械にも史料の適切な取り扱いが可能となる。例えば「差出」などといった史料の特定の構造のみに限定した検索が可能になるとともに、付箋などの付加情報を本文と区別して扱う

こともできるだろう。これらはいわゆる棒テキストによる検索や処理では不可能であり、利用者の利便性を大きく高めるものであろう。

第二に、バリアフリー・国際化への貢献が挙げられる。例えば視覚障害者の用いる読み上げソフトに対して史料構造を適切に反映した読み上げをさせることが可能になる。また日本語話者以外に向け、史料の参考訳を自動翻訳により提供するということも見通せるだろう。ここには、史料編纂所が作成した日英グロッサリー・データベースも役割を果たしうる。このように、編纂知に基づく適切な構造化は、人間にも機械にもやさしい史料集を実現することができる。

ところで史料編纂所では、史料集の版面画像を「[版面ギャラリー](#)」でインターネット公開している。この版面画像やデータベース、史料画像の活用の展望を最後に紹介したい。実は『松平昭休往復書簡留』には伝来過程で発生した錯簡が確認されており、編纂に際して冊子の解体と検討・修正・再編綴を行った。その各段階で撮影した画像はいずれも「Hi-CAT」で公開されているが、現時点ではバラバラの状態である。これらを、版面も含め紐づけることが将来的には必要となる。作成済みの構造化テキストを基礎に、版面と史料画像にアノテーションによる構造化を施すことで、相互のリンクやデータベースとの接続を可能とすることが展望される。また、索引 DB の活用としては、版面画像へのアノテーション付与によりデータベースへの導線を確保した上で、さらに他種人名データベースへのリンクや、他機関を含む関連史料へのリンクも構想される。(当日の報告では、そのデモを用意し参加者に公開した)

■ おわりに

以上、「編纂知」と呼ばれる暗黙知・経験知について、事例に即した紹介と、そのデジタル・トランスフォーメーションにおける応用を論じた。近年一足飛びに進むデジタル化において、紙史料集は時代遅れのものと見なされつつある。しかしながらそこには、近世に端を発する「編纂」の伝統と経験知の膨大な蓄積が存在する。本のスキャンや OCR だけではすくい上げることのできないそれらを、適切にデジタルの世界へ持ち込むことは急務と考えられる。

本論において述べたようにデジタル画像の公開と AI-OCR の発展はあたかも原史料をそのまま誰でも刊行史料のように扱うことを可能としている。だが、適切な補助線もなく原史料にあたることは、訓練された研究者ならまだしも、恣意的あるいは不適切な解釈を生む危険性を大いにはらんでいる。さらに注意すべきは、その「利用者」がもはや人間だけではないことだ。単にインターネットに公開されただけの史料や歴史データが、適切・正当な扱いを受けることを期待するのは、あまりにナイーブである。そのような態度は、史料の価値をかえって損ねるものではないか。

編纂史料集は、編纂知に則った責任ある解釈を提示することで社会に役割を果たしてきた。現代にあって、紙の世界に留まることはいわばその責任の放棄である。史料の画像やテキストのみならず、「編纂知」を積極的にデジタルデータやデジタル規格に持ち込み、それによって「責任ある解釈」をあらゆる人間と機械に提示することが必要だ。先に述べた「モノ」実体をハブとした多種データの編成の実現が、ここに大きな役割を果たすことになるだろう。そしてこのことは、ハイブリッド化された史料集やその編纂自体の価値をも高めるものである。人間と機械知性の合同作業によって行われることになるだろう将来の歴史学に向けて、真のデジタル・トランスフォーメーションの姿を見極めることが、編纂する者に問われている。

報告 社会科学から見た歴史データ

山崎 潤一（京都大学大学院経済学研究科）

■ 1. はじめに

本題に入る前に、どのような視点から書かれた原稿なのかを明らかにするために、少し私のこれまでの研究歴を紹介したい。私は、学部から修士までは東京大学に在籍し、その後ロンドンスクールオブエコノミクス（LSE）に留学をし、2018年に経済学の PhD を取得した。経済学業界では、修士の一年目にいわゆるコアコースと呼ばれる必修の講義があり、大学が違ってもある程度標準化された内容が教えられている。そういう意味で経済学業界の極めてスタンダードな教育を受けてきた人間だと思われる。

もともと私は開発経済学の実証研究に興味があり、学部時代からそうした分野で研究をしたいと思っていた。ちょうどその頃、開発経済学分野では歴史データを用いた研究が数多くトップジャーナルに載るようになった（そしてその嚆矢となった Acemoglu, Johnson, and Robinson(2001) は先日のノーベル賞の受賞理由となった）。イギリスに渡って数年経ったころだったか、友人と何か歴史を題材に研究ができないかという話をし、実際取り掛かって以来日本経済史のデータを使った研究をするようになった。

私は漢字の暗記、特に人名が苦手なこともあり、もともと所謂“歴史嫌い”である。高校の日本史や大学の日本経済史の授業を取ったことがない。おかげで日本経済史の基礎的文献をロンドン大学東洋アフリカ研究学院の図書館で一人読む羽目になった。

こういう経緯があるので、なぜこのような“大転換”が私に起きたのかを説明することと、経済学が歴史に対して持つ眼差しを説明することと、実はほぼ同じことであると考えている。

■ 2. なぜ経済学（者）が歴史に興味を持つのか

遅塚「史学概論」によれば、経済学は政策提言を目的としていることが、歴史学との大きな違いであると述べられている。この理解自体には大きな異議はないが、ではなぜそんな経済学が歴史に興味を持ちうるのだろうか。私の狭義の専門の開発経済学を例に、2つ理由を挙げようと思う。

第一に、多くの経済現象は長期的なプロセスによって生まれてくるので、今起きている事象を理解するには歴史が不可欠ということが挙げられる。例えば、国レベルの経済発展の指標である「一人あたり国内総生産」一つをとっても、過去に高い経済発展を示した国は今でも過去に高い経済発展を示すことが多い。一方で、過去にはあまり発展していなかった国でも、その後発展を遂げた国と、そうでない国があり、その理由を探究することは経済政策の提言にとって重要な分析となるのである。こうした潮流は以前から存在したが、特に Acemoglu, Johnson, and Robinson(2001) 以降、所有権や法の支配といった制度が果たす役割に注目が集まり、さらに研究成果が蓄積されている（Koyama and Rubin (2022)）。

第二に挙げられるのは、実証経済学において、歴史は実験室という考え方が広まったことである。これは開発経済学に限らず、労働経済学や公共経済学など応用分野に関する実証研究分野に広がったもので、関連する先駆的研究を行なった研究者たちは 2021 年にノーベル経済学賞を受賞している。この説明は少し専門的になるので、分量を割いて説明しようと思う。

経済学の実証研究では、物事の因果関係を数値的に示すことが最も重要であるといっても過言

ではないだろう。例えば、大学に通うことでどれだけその人の賃金上がるのか、と行ったことが因果関係と呼ばれるものである。こうした因果関係を知らなければ、大学進学を政策的に促すべきかの、判断に迷うからだ。さて、それを分析しようと思うと、大卒の人とそうでない人の賃金を比べるのが自然な方法ではあるが、残念ながらそれでは因果関係はわからない。これらの二つのタイプの人たちの間には、おそらく大卒かどうか以外にあらゆる違いが存在するからである。例えば大卒の人は家庭環境が良かったり、元々能力が高かったりするかもしれない。なので、両者を比較しても色々な効果が混ざったものしかわからず、純粋な大学進学の影響は見えにくいのだ。

因果効果を知りたい際の有力な方法として、ランダム化比較実験法があげられる。例えば、ある薬の効果を知りたい場合、ランダムに薬を服用する人と偽薬を服用する人を選び、グループ間で薬を飲んだかどうか以外の違いを排除することで、因果効果を調べることができる。しかし、多くの経済政策においては、ランダム化を行うことが政治的・倫理的に難しい。そこで、過去に実施された政策や経済環境の変化を利用し、その中にランダム的要素を見出して因果関係を探る試み、自然実験的アプローチ、が取られている。

具体的な例を挙げよう。アメリカはベトナム戦争中、ベトナムを空爆することで共産主義勢力の進展を抑えようとした。しかし、この空爆は地元住民の反感を買い、共産主義勢力の支持を増やす可能性が指摘されていた。どちらが正しいかはデータを分析してみなければ分からないが、空爆を受けた地域とそうでない地域の間には、さまざまな違いがあると考えられる。そこで Dell and Querubin(2018) ではアメリカ軍の空爆戦略を再検証した。結果として、アメリカ軍はベトナムでサーベイを行い、統計モデルに基づいて空爆を優先的に行うべき場所を点数化してランクづけをし、とある点数を境に五段階に分類したことがわかった。例えば点数がとある点より上だと A ランク、下だと B ランク、といった具合である。この点数の切れ目を使って、たまたま空爆されやすい地域とされにくい地域の綺麗な比較が可能で、結果共産主義勢力をサポートする人が増えてしまったという結論を得た。

つまり、自然実験的アプローチをとる経済学者にとっては、自らの手で政策をランダムに実施するということが不可能な場合、既に世の中に起きた事柄の中から、ランダム的要素を探すことになる。それが自国でも他国でも良いし、現在でも過去でも良い。そこで歴史が出てくるわけである。実際、自分の興味のある政策が直近の過去に起きることは稀かもしれず（コロナ禍の際にスペイン風邪まで遡って議論がされたのは記憶に新しい）、自然と過去に戻ることになる。こうした潮流は経済学のみならず、ポリティカルサイエンスなど他の社会科学にも見られ、こうした意味での歴史への“需要”は高まったといって良いだろう（Charnysh et al. 2023）。この例以外にもナポレオンの海上封鎖、新大陸発見による作物交換、スエズ運河の閉鎖などを使って興味深い自然実験論文がある（Juhász 2018, Nunn and Qian 2011, and Feyrer 2021）。

そして、一つ見過ごされがちな重要な点は、Dell and Querubin (2018) の例でも明らかなように、そもそも興味のある政策がどう決まっていたのかという背景に関する非統計的な知識に依拠しているということである。また、さらに興味深いことに、Dell and Querubin(2018) では点数自体は既に記録として失われていたが、サーベイ記録と空軍が使った統計モデルはわかっていたため、サーベイ記録を電子化したのち研究者たちが再推計した点数を分析に使っている。つまり

充実したアーカイブや記述的分析も当然ながら欠かせないものなのだ。

■ 3. デジタル化といった時に社会科学者は何を期待しているのか

さて、そんな社会科学者にとって、資料のデジタル化は福音である。また歴史学や人文学一般でも近年デジタルヒューマニティーなどの盛り上がりもあり、デジタル化の波が来ているように感じる。しかしデジタル化、といったときに想像するものが、社会科学系と人文系で異なることがある。本節ではその違いを私見ながら簡単にまとめた。

人文学では、アーカイブの作成など、オリジナルの史料をなるべく原型を保った形で保存し公開することが重要視されているように感じる。その価値は誰もが認めるところではあるが、経済学でデジタル化といった場合、紙から画像などのデジタルファイルに変換することではなく、統計分析に適した行列形式（つまりエクセル表のような形）で保存されたデータを作成することを指すことが多い。

経済学や社会科学一般にで使いやすいデータの例として思いつくのは Integrated Public Use Microdata Series (IPUMS) が整備した各国のセンサスデータである。ここでは各国の人口センサスで調査された年齢や教育、職業といった変数を、国際間もしくは時代間比較可能な形で再編集し、容易に国際間比較ができるように加工されたものがダウンロードできる。これにより、研究者は手間を労することなく、データ分析を行うことができる（残念ながらなぜか日本のデータは収録されていないが）。

さらに、こうした変数に関する解釈の明確化や標準化のみならず、データ形式自体にも一定の規則が守られていることが重要だ。データサイエンスの業界ではよく tidy data（整然データ）と呼ばれ、データベース理論の業界では正規化と呼ばれるような、データに関する規則がある。厳密なルールはさまざまなものがあるが、ごくごく基本的なものとして、一つのセルには一つの値が入る、というものがある。しかしこれも守られていないことが散見される。こうした問題は学問分野だけではなく、一般社会でも神エクセル問題として、日本社会の IT 化を阻害している要因として指摘されており、総務省が資料を作成し具体的な改善案を提案するところである（総務省 2020）。

なお、経済学者がデジタル化を行う場合、紙で出版されている活字の資料を使うことが多いが、そこにも落とし穴が存在する。紙で出版された場合、本来残されるべき編集情報が欠落していることがあるからだ。例えば、右図のような例がそれにあたる。これは大日本近世史料柳営補任の人名索引なのだが、同一の名前“荒井精兵衛”を持つ人物が二人おり、p102 に双方共が登場しているようだ。この場合どちらの人物が p187 にも登場する人物なのか、この索引を見るだけではわからず、人の目で本編を確認の上特定しないといけない手間が発生する。本来ならば、柳営補任の各職歴記録全てに職歴 ID を振り、その職歴 ID と個人の ID との対照表が示されるべきである。紙資料での出版であるからか職歴記録のページ数という曖昧なオブジェクトと紐付けたからこそ起こる現象である。

荒井郁之助	⑤二〇七・二二三・二二三・二二三・二二三
荒井十兵衛	①二六
荒井甚之丞	和泉守 因幡守 伊勢守
荒井甚之丞	①元五⑤三三⑥二八・二二三・二三五
荒井助太郎	④二六
荒井精兵衛	⑥三三
荒井精兵衛	①〇三
荒井精兵衛	①〇三⑤二七

そうして考えると、本来ならば、こうした編集の過程で付与された情報はデジタルデータとして残し、あくまで閲覧方法の一つとして紙媒体での出版があったり、また別の閲覧方法として構造化されたデータもある、という捉え方が自然に思えてくる。紙では疎かにされたり、表現しにくいタイプの情報もある。国の統計年鑑のような統計資料に限って言えば、紙をめくって統計表を読んでいる人がどれだけいるのかも疑問であるし、データはやはりデータとして整備し利用するのが理想的だろう。

■ 4. 進行中のサムライプロジェクトの紹介

私は共同研究者らと江戸幕府の行政を知る上で重要な史料である、寛政重修諸家譜と柳営補任のデジタル化、構造化データ化に取り組んでいる。これらは旗本の家系図と職歴の記録に対応し、個人がどのような家に属し、どのような職歴を辿ったかを捕捉するための貴重な史料である。また幕府側からの観点から見れば、混乱の戦国時代から戦争のリスクは低くなった江戸時代において、軍事から行政への転換があったのか、どのようにしてそれを促したのかといった、組織の変遷を見たかを分析できる貴重な史料であり、開発経済学 / 政治経済学的にも重要である。我々はこれを通称サムライプロジェクトとしてこの数年間進めてきた。

既に OCR や手入力によりおおかたのテキストの電子データ化は終えており、柳営補任と寛政重修諸家譜それぞれにおいて、個々の記述に対する個人 ID の割り振りもほぼ完了している。つまり、個人に関してどのような家に生まれ、養子としてどの家に行ったのかも追跡できるし、またどのような役職を経験したのかを全て芋づる式に把握できるデータとなっている。まだ双方のデータを接続するところまでは作業は済んでいないが、構造化されたデータとして整備し、学術研究に用いること、また将来的には公開をしデータの細かい改善をしながら研究コミュニティに対して公共財を提供できればと思っている。

■ 5. おわりに

経済学と歴史学は、従来遠い分野として捉えられてきたと思う。もちろんその目的や、それに根付いた研究手法には違いは歴然として存在している。しかし、それを断絶というよりも、グラデーションのようなものだと考えた方が良く、研究者個人の中では異なる関心や分析手法が同居するケースもあるだろう。さらには、研究手法の違いも、不変のものではなく、機械学習のような新技術の登場によって、変化するものでもある。

共著論文を書くという意味においての分野横断的共同研究は、職業的研究者にとって雑誌の選定など色々と乗り越えるべき高いハードルが多くある。しかしデータや史料の整備といったより基礎的な部分に関しては、経済学者や社会学者はより幅広いデータの分析方法や、統計的な利用を念頭においた整備に関する助言によって、そのデータや史料の利用価値を上げることができるだろう。また、歴史学者は従来から行ってきたように、経済学者が見落としがちなデータの解釈の問題や多義性などを指摘し、編纂をより正確に行うことができる。そうした協力—それは研究チームを組んだ陽な協力であっても、引用関係を通した陰な協力であっても良いが—によって、データ整備の好循環が生まれ、さまざまな研究成果が蓄積されていくのが、健全な姿だと思う。

■ 参考文献

- Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91, no. 5 (December 1, 2001): 1369–1401.
- Charnysh, Volha, Eugene Finkel, and Scott Gehlbach. "Historical Political Economy: Past, Present, and Future." *Annual Review of Political Science* 26, no. Volume 26, 2023 (June 15, 2023): 175–91. <https://doi.org/10.1146/annurev-polisci-051921-102440>.
- Dell, Melissa, and Pablo Querubin. "Nation building through foreign intervention: Evidence from discontinuities in military strategies." *The Quarterly Journal of Economics* 133, no. 2 (2018): 701–764.
- Feyrer, James. "Trade and Income—Exploiting Time Series in Geography." *American Economic Journal: Applied Economics* 11, no. 4 (October 1, 2019): 1–35. <https://doi.org/10.1257/app.20170616>.
- Juhász, Réka. "Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade." *American Economic Review* 108, no. 11 (November 2018): 3339–76. <https://doi.org/10.1257/aer.20151730>.
- Koyama, Mark, and Jared Rubin. *How the World Became Rich: The Historical Origins of Economic Growth*. Cambridge, UK ; Medford, MA: Polity, 2022. (和訳本 「経済成長」の起源: 豊かな国、停滞する国、貧しい国, 2024, 草思社, 秋山勝訳)
- Nunn, N., and N. Qian. "The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment." *The Quarterly Journal of Economics* 126, no. 2 (May 1, 2011): 593–650. <https://doi.org/10.1093/qje/qjr009>.
- 総務省, "統計表における機械判読可能なデータ作成に関する表記方法", 2020 https://www.soumu.go.jp/menu_news/s-news/01toukatsu01_02000186.html

報告 経済史研究における基盤的数量データの公開・活用

高槻 泰郎（神戸大学経済経営研究所）

はじめに

近年、自然科学や社会科学の研究において、歴史を実験室として検証することへの関心が高まっている（Diamond and Robinson 2011）。自然科学や社会科学における重要な問題の中には、科学的手法の典型である「管理された実験室での実験」では答えられないものもある。たとえ実験が可能であっても、非道徳である、または違法とみなされる場合もある。歴史上の出来事を自然実験として扱う研究は、こうした制約を乗り越えるものとして期待が集まっていると言える。

この流れにあって、ユニークで信頼性の高い歴史上のデータに対する需要は高まりを見せている。報告者が専攻する近世日本経済史研究も例外ではないはずだが、基盤的な数量データ、例えば生産量、人口、物価、金利といったデータのデジタル化が十分に進んでいないこともあり、近世日本を題材とする数量的な研究、特に日本語を母語としない研究者による研究は、決して多いとは言えない状況にある。横文字の研究が増えること、それ自体が望ましいと主張するつもりは毛頭ないが、国内外の、より多くの研究者が近世日本に関心を持ち、分析の俎上に載せることは、近世日本経済史研究に裨益する面も大きいと考えている。

そこで本報告では、史料から復元した数量的な情報をデジタル化して公開するに際して解決すべき課題を共有するものである。ここでは近世日本の経済史に関わるデータ（主として物価と貨幣両替相場）を想定して論じる。

1. 近世日本経済史研究における基盤的数量データの公開状況

近世日本経済史を研究する上での基盤的数量データの多くは、学術書ないし各種辞典などの「紙」の上にまとめられている¹。これらの内、既にデジタル化されて無償公開されているものを挙げると以下の通りである。

第一に、三井家編纂室（1916a、1916b）が提供する大坂の米相場、貨幣両替相場、天候、一部イベントについては、神戸大学経済経営研究所と公益財団法人三井文庫が共同で作成・管理している「近世経済データベース」においてデジタルデータとして公開されている²。このデータベースは、三井家編纂室（1916a、1916b）をデジタル化したものではなく、典拠史料である三井大坂両替店「日記録」（公益財団法人三井文庫蔵）に遡及してデジタル化したものである³。

第二に、岩橋（1981）が提供する日本各地の年次米価、および三井文庫（1989）が提供する月次の商品価格の一部については、UC Davis の提供するデータベース“Global Price and Income

1 具体的には、三井家編纂室 1916a、1916b、鈴木 1935、三井文庫 1952、1989、京都大学近世物価史研究会 1962、宮本 1963、小葉田他 1966、鶴岡 1972、岩橋 1981、山崎 1983、深尾他 2017、高島 2017 などである。

2 神戸大学経済経営研究所「近世経済データベース」（<https://www.rieb.kobe-u.ac.jp/project/kinsei-db/>）、2024 年 12 月 6 日閲覧。本データベースの作成を担当したのは、村和明（公益財団法人三井文庫、当時）、高槻泰郎（神戸大学経済経営研究所）、藤尾隆志（神戸大学経済経営研究所研究補佐員、当時）、尾脇秀和（神戸大学経済経営研究所研究員）の 4 名である。

3 この過程で、三井家編纂室（1916a、1916b）には複数箇所誤りが見られることが判明した。したがって、近世大坂の米相場、貨幣両替相場を参照する者は、今後「近世経済データベース」を参照することを推奨する。

Database (GPID)"⁴にてデジタルデータが公開されている。

管見の限り、上記以外については、現時点では「紙」の上には存在しないデータであり、データ利用者によるデータの再打ち込みコスト（時間的、経済的コストに加えて入力ミスを確認・検証するコストも含む）が不可避免的に発生する。加えて、「紙」の上に掲示されたデータは、本来復元できるデータから摘記されたものであることが多く、紙幅という物理的制約によって落とされている情報が存在することもあり得る。

例えば、鶴岡（1972）は、国文学研究資料館所蔵「近江国蒲生郡鏡村玉尾家文書」所収の「万相場日記」という史料から、宝暦期から安政期（18世紀中期から19世紀半まで）の大坂米価、大津米価など月次データ（毎月月初の価格を採録）を掲載しているが、原本史料には、特に寛政期（主として1790年代）以降、月次以上の頻度で価格を記録している上に、鶴岡（1972）が掲示するよりも多くの商品価格を記録している（高槻 2012、140-145 頁）。作業者の時間的制約も手伝っていたであろうが、文字通りの「紙幅」に納めるために、採録する銘柄や頻度が調整されたことは事実であり、またこの割愛が認識されないまま、鶴岡（1972）の提供する米価データが様々な形で参照されていたこともまた事実である。

また、新保（1978）は、貨幣両替相場を視認性の高い一覧表として掲示しているが（表 4-2）、これも紙幅の関係であろう、貨幣改鋳前後に掲示されている両替相場が、新貨と旧貨、いずれの数値なのかが不明瞭になってしまっている。一部は備考欄で付記されているが、新保（1978）が参照した原史料（「両替年代記」など）に遡らないと、この点を厳密にすることはできない仕様になっている。

上記のような制約から自由なデジタル化されたデータの公開を推し進めることにより、国内外の研究者にとって利便性の高い環境を構築していくことが求められる。

■ 2. 歴史的数値データをデジタル化するメリットと解決されるべき課題

2025 年度応募分より、科学研究費補助金などの公的資金を受けて行われた研究については、成果として得られた査読付き学術論文及び根拠データの公開が義務づけられることを踏まえれば、今後、歴史的数値データがデジタルデータとして公開される機会は増えていくであろう⁵。しかし、そこには当然メリットとコストがある。そしてそれは、経済史研究に限定されないものである。まずメリットについて検討すると、上記の通り、紙幅という物理的制約を受けないという点、そして利用者に金銭的費用が発生しない（書籍購入代・複写代、データ打ち込み費用が発生しない）という点に加えて、公開後の修正が容易であることが重要である。

データ作成にはヒューマンエラーがつきものであり、例えば三井家編纂室（1916a、1916b）の誤りを訂正した「近世経済データベース」もまた、ヒューマンエラーから自由ではない。この点、デジタルデータであれば、バージョン情報を付記して改訂していくことも可能で、利用者には最新バージョンの利用と、論文へのバージョン明記を求めることで対応が可能である。無論、前提

4 <https://gpih.ucdavis.edu/Datafilelist.htm#Asia>（2024 年 12 月 6 日閲覧）。

5 「学術論文等の即時オープンアクセスの実現に向けた基本方針」（2024 年 2 月 16 日、統合イノベーション戦略推進会議決定、https://www8.cao.go.jp/cstp/oa_240216.pdf、2024 年 12 月 6 日閲覧）。

6 筆者が鶴岡（1972）の整備した大坂と大津の月次米価データについて、原史料に遡って追証を行った際、鶴岡（1972）には、くずし字で墨書された漢数字の「六」と「壺」および「三」と「壺」の識別に揺れが見ら

として典拠史料への遡及が可能な状態、それもできればオンライン上で可能な状態になっていることが求められる。

しかし、それには当然ながらデータサーバー購入費用、維持・管理費用といった費用が発生する。現状では、各研究機関（場合によっては研究者個人）がこれらの費用を負担しており、科学研究費補助金（研究成果公開促進費）では十分にカバーされていない。例えば、公開後の保守・運営費を同費目の科研費から支出することは不可能である。属人的管理が行われている場合、当該研究者の移籍・退職によって管理が行き届かないこともあり得る。最悪の場合、データの公開自体が持続できない可能性もある。競争的資金を受けて掲載された査読付き学術論文で典拠となったデータのみならず、潜在的に研究資料となり得る歴史的数値データを、しかるべき公共機関に集約して恒久的に管理する体制の構築に向けた議論が必要ではないだろうか。

デジタルデータ公開に向けて解決すべき課題は、費用面に限らない。例えば上述の GPID では、岩橋（1981）の付表をデジタル化した年次米価データを公開しているが、解題が不十分なまま公開されている。例えば、寛永元年（1624）から天和3年（1683）までの大坂米価とされている米価は、実は京都米価であることは記されていない（無論、岩橋 1981 は明記している）。必要にして十分な解題を伴わずにデータだけが一人歩きする状況を避けるためにも、上記に提案した公的機関による審査体制を構築することも一案であろう。

最後に、データベースを構築している当事者として、データベース構築、それ自体が研究者個人の業績に必ずしも結びついていない（少なくとも、査読付き学術論文や書籍出版ほどには評価されない）という点を指摘しておきたい。ウェブ上にデータを公開するよりも、書籍の形で出版した方が業績として認知されやすい、という状況があるとも仄聞する。

インパクトファクターの高い国際査読誌への論文掲載を国策として推奨するのであれば、インパクトファクターの高い国際査読誌への論文掲載には不可欠となる、ユニークで精度の高い良質のデータを国際発信した者に対しても、しかるべき評価が与えられるべきであるが、少なくとも現状では「奇特の行為」の範疇に収まっている。データベース構築、公開の費用を負担している者が適切に評価されるような体制が構築されることを強く望んで、この報告を終える。

■ 謝辞

本報告は、JSPS 科研費（24K00277、24K00107、23K25538、23K25370、21K18444）の助成を受けたものである。

れることが判明した。しかし、これは筆者とて同様であって、原史料に当たって検証しても、これらの識別が不可能である場合が少なくなかった。こうした場合、別の読み方が排除されない、ということを示すために、例えば「不確定フラグ」なる欄を用意して、そこに「1」を入力する、などの対応がデジタルデータの場合は可能になる。分析者は、フラグに「1」が入力されているデータを除いて分析する場合と、含めて分析する場合とで結果が変わるかどうかを検証できる。また、明らかな数字の誤りがある場合（同時代人による書き間違い、写し間違い等）、それを訂正したデータ系列と、あえてそのままにした系列の両方を掲示することも同様に可能である。

■ 参考文献

- 岩橋勝（1981）『近世日本物価史の研究—近世米価の構造と変動—』大原新生社。
- 小葉田淳・豊田武・寶月圭吾・森克己監修（1966）『読史総覧』人物往来社。
- 京都大学近世物価史研究会（1962）『15～17世紀における物価変動の研究—日本近世物価史研究 1—』読史会。
- 新保博（1978）『近世の物価と経済発展—前工業化社会への数量的接近—』東洋経済新報社。
- 鈴木直二（1935）『大阪に於ける幕末米価変動史』四海書房（後、国書刊行会より1977年に再版）。
- 高島正憲（2017）『経済成長の日本史』名古屋大学出版会。
- 高槻泰郎（2012）『近世米市場の形成と展開—幕府司法と堂島米会所の発展—』名古屋大学出版会。
- 鶴岡実枝子（1972）「近世米穀取引市場としての大津—湖東農村商人の相場帳の紹介（二）—」『史料館研究紀要』第5号、19-209頁。
- 深尾京司・中村尚史・中林真幸編（2017）『岩波講座 日本経済の歴史 第2巻 近世 16世紀末から19世紀前半』岩波書店。
- 三井家編纂室編（1916a）『自天明七年至明治四年大阪金銀米錢并為替日々相場表 巻1』。
- 三井家編纂室編（1916b）『自天明七年至明治四年大阪金銀米錢并為替日々相場表 巻2』。
- 三井文庫編（1952）『近世後期における主要物価の動態』日本学術振興会。
- 三井文庫編（1989）『近世後期における主要物価の動態』東京大学出版会。
- 宮本又次（1963）『近世大阪の物価と利子—日本近世物価史研究 3—』同氏発行。
- 山崎隆三（1983）『近世物価史研究』塙書房。
- Diamond, Jared and James A. Robinson (2012), *Natural Experiments of History*, Harvard University Press (ジャレド・ダイヤモンド, ジェイムズ・A・ロビンソン (2018)『歴史は実験できるのか』慶應義塾大学出版会)。

報告

歴史資料のデジタル化を支える技術

中村 寛（東京大学史料編纂所）

本稿では、歴史資料のデジタル化とそれを支える技術について説明する。歴史研究における史料のデジタル化は、単なる技術的な進展にとどまらず、研究手法や史料へのアクセス、保存の方法に変革をもたらす重要なプロセスである。特に、デジタル技術の進展により、これまで研究者のみが利用していた史料が、広範なユーザ層にも提供され、利用可能となる。このことは、史料へのアクセスを促進し、研究の効率化や拡充に寄与する点で重要である。

本稿では、歴史資料のデジタル化において中核的な役割を果たす2つの技術、TEI（Text Encoding Initiative）およびOCR（光学文字認識）を中心に取り上げ、それらの技術の具体的な利点や、東京大学史料編纂所で行われているデジタル化プロジェクトを事例に、その応用可能性について考察する。

■ 1. 歴史資料のデジタル化の意義と背景

歴史資料のデジタル化は、学問的および社会的に重要なプロセスであり、その意義は広範囲にわたる。従来、紙の史料は時間の経過とともに劣化し、その閲覧には場所や人数の制限があった。しかし、デジタル技術の進展により、史料をスキャンしデジタル形式に変換することが可能となり、誰でもどこからでもアクセスできる環境が整いつつある。これにより、研究の進展が加速し、史料の利用方法にも変化をもたらしている。

デジタル化はまた、史料の保存においても重要な役割を果たしている。紙の史料は物理的な劣化のリスクがあるが、デジタルデータは適切な管理のもとで長期にわたって保存することが可能であり、さらに容易にコピーや複製ができる。これにより、災害や事故による史料の消失リスクが軽減される。このように、デジタル化は技術的な進展のみならず、歴史資料の保護や永続的な保存においても重要な役割を果たしている。

■ 2. TEI（Text Encoding Initiative）とは

TEI（Text Encoding Initiative）は、人文学分野におけるテキスト史料のデジタル化に使用される標準規格である。1987年に設立されたこのイニシアチブは、人文学者、情報技術者、図書館員などが協力して策定したもので、主に歴史的文献や文学史料のデジタル化を目的としている。その最大の特徴は、テキストを単にデジタル化するのではなく、その構造や内容、注釈、メタデータを詳細に記述できる点にある。

TEI ガイドラインは、XML 形式を用いてテキストを構造化し、史料の内部構造や解釈に関する詳細な情報を記録することが可能である。例えば、文献の章や段落、注釈、解説、執筆者の情報、信頼性の程度などが含まれ、研究者にとって有用なメタデータを提供することができる。これにより、史料は単純なテキストデータ以上の価値を持つことになり、研究に役立つリソースとなる。

■ 3. TEI の応用例と視覚化ツール

TEI によるデジタル化の実例として、1556年に明の使者蔣洲が対馬の宗氏に倭寇禁圧を求めた

公文書である『蔣洲咨文』のデジタル化プロジェクトが挙げられる。この史料を TEI に基づいてデジタル化することで、テキストのみならず、関連する地図や画像、固有名詞のリストなども一元的に管理し、視覚的に表示することが可能となった。



TEI は、単なるテキストの保存に留まらず、史料に関する多次元的な情報を含むことができる。このため、視覚化ツールを用いることで、史料のコンテキストや構造の直感的な理解を助けることができる。視覚化ツールは、史料に付加されたメタデータを基に、検索や表示のカスタマイズが可能である。

4. OCR（光学文字認識）の役割

次に、歴史資料のデジタル化において重要な技術である OCR（光学文字認識）について説明する。OCR は、紙の文書をスキャンし、その画像データからテキストデータを抽出する技術であり、特に大量の歴史資料に対して効率的なデジタル化を可能にする。

東京大学史料編纂所や関連するプロジェクトでは、国立国会図書館が提供する NDLOCR アプリケーションを用いて、これまでの出版物をデジタル化している。例えば、寛政重修諸家譜といった膨大な人物情報を含む史料は、手作業での入力が困難であるが、OCR 技術を用いることで、迅速かつ効率的なデジタル化が進められている。

5. OCR 技術における課題とその解決策

OCR 技術は非常に有用であるが、特に歴史資料の場合、いくつかの課題がある。手書き文字や異体字、特殊な書体などが使用されている場合、標準的な OCR では正確な文字認識が難しい。また、複雑なレイアウトを持つ史料では、単純に文字を読み取るだけでは不十分な場合がある。

こうした課題を解決するためには、レイアウト認識やカラム認識といった高度な技術が求められる。近年では、AI や機械学習を用いた OCR 技術が進展しており、手書き文字の認識精度が向上している。さらに、YOLO などの物体検出モデルを用いることで、ページ内の構成要素を正確に識別し、複雑なレイアウトに対応することが可能になっている。

6. デジタル技術の応用事例とその効果

これらのデジタル技術の応用により、歴史資料の管理と活用を支援することができる。東京大

学史料編纂所が提供する史料集版面検索システムでは、出版物のOCRテキストを検索・閲覧することができる。また、Ngram Viewerを用いることで、複数の史料からキーワードを抽出し、その出現頻度や時期ごとの傾向を視覚的に分析することが可能となる。

こうしたツールは、従来の手法では見過ごされがちだった情報の掘り起こしを助け、新たな視点からの研究を促進する可能性がある。

地震

詳細検索

検索オプション①

☒ 通常検索
 ☐ 正規表現
 ☐ 曖昧フレーズ検索1
 ☐ 曖昧フレーズ検索2

検索オプション②

☒ AND検索
 ☐ OR検索

年: 1847

クリア

検索結果: 72

1 / 4

20 件

適合度 降順

4 弘化4年2月～同年3月 p.689

第一階層: 『大日本維新史料 編年之部』 第二階層: 『大日本維新史料 編年之部』 1編 史料名: 『大日本維新史料 編年之部』 1編 4 弘化4年2月～同年3月

○信越地震記事ニモ略々同一内容ノ記事アリ。

在町破損所御坐候、遠在し儀は、右し通御座候、此段御届申上候、以上、私在所信濃國松本、去ル廿四日夜四ツ時頃大地震強、翌廿五日ち爲差儀は無之、間遠ニ相、(信越地震記事)、未相分兼申候、先此段御届申上候、以上、以上

松本藩主戸田光則届書、三月廿七日在所日附、松平丹波守、男拾壹人、一死人六十七人内七人焼死、松本領分地震異變取調

三月廿七日在所日附松平丹波守、○帝國圖書館所蔵本、○中村、六所蔵本、胡路、信州松本地、方地震

4 弘化4年2月～同年3月 p.709

第一階層: 『大日本維新史料 編年之部』 第二階層: 『大日本維新史料 編年之部』 1編 史料名: 『大日本維新史料 編年之部』 1編 4 弘化4年2月～同年3月

ニ有之、夫ち不絶震動いし、折る地震發シ、翌廿五日卯中刻、漸く相、上ケ可申を難斗、同廿六日に至り候なも

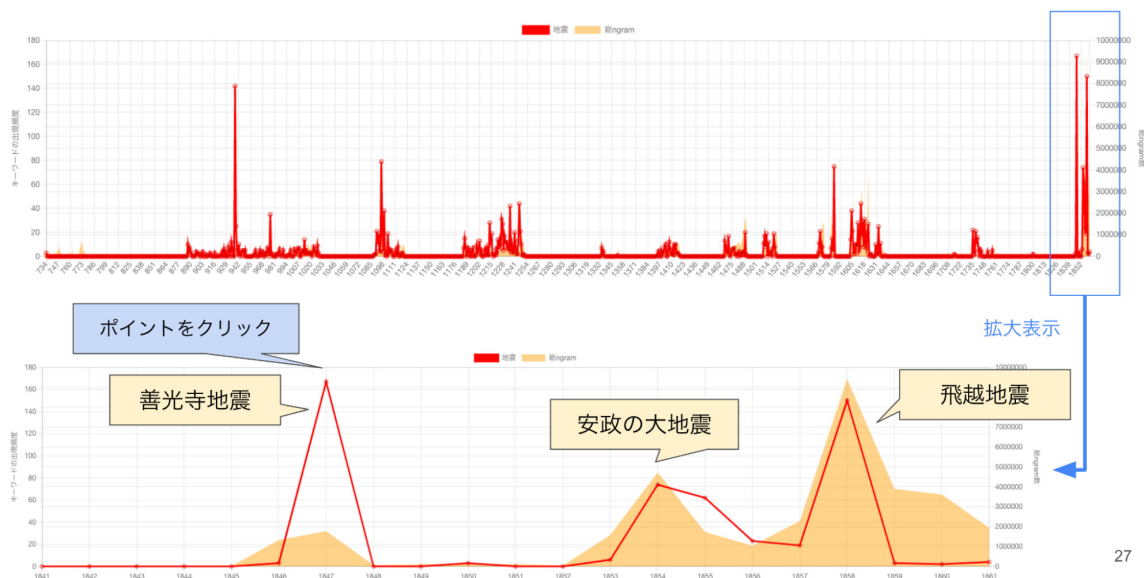
井人馬死亡等も有之趣、其上篠井川筋兩岸拔崩、水中埋押、水吐兼田方水溢追々畑方に押、弘化四末年三月大地震

し儀ニ付御届書、出雲守領分信濃國高井郡・水内郡み内村、去月廿四日亥刻頃大地震二な、人家震ヒ潰、

三月晦日地震し儀ニ付御届、堀出雲守家來、四月四日、奥田爲太郎、四月四日堀出雲守家來、○胡路

、隔、震島御届書對、○胡路、○信濃國大地、被害アリ、信州中野地、方地震

Ngram Viewerで「地震」を検索



7. クラウドソーシングとデジタルプリザベーション

デジタル化の進展は、クラウドソーシングやデジタルプリザベーションにも大きな影響を与えている。クラウドソーシングにより、多くの人々が歴史資料の校訂やデータ整理に参加でき、研究の速度が格段に向上している。市民参加型のプロジェクトは、データの拡充と同時に、歴史的

知識を広く普及させる手段としても重要な役割を果たしている。

一方、デジタルプリザーベーションは、歴史資料を将来にわたって保存し続けるための技術である。デジタル資料は、オンラインで広くアクセスされることで、学術的および文化的価値が共有されることが期待されている。しかし、長期保存は単なるデータのバックアップとは異なり、データの真正性と完全性を維持することが求められる。

■ 8. 今後の展望と課題

今後、デジタル化技術のさらなる進展により、歴史資料のデジタル化とその活用は一層進化すると考えられる。VR や 3D 技術を用いた史料の視覚的再現、機械学習を用いた高度な史料分析が可能になれば、歴史研究において新たなアプローチが生まれる可能性がある。

しかし、技術の進展に伴い、デジタル化された史料の保存やデータの信頼性、プライバシー、著作権問題など、新たな課題も浮上する。これらの問題を解決しつつ、より多くの人々が史料にアクセスし、活用できる環境を整備することが求められる。

■ 9. まとめ

デジタル技術は、歴史資料の保存、共有、活用を支える重要な要素となっている。TEI や OCR などの技術は、従来の史料管理方法に変化をもたらし、研究や教育の分野で新たな可能性を生む可能性がある。本稿では、これらの技術が歴史資料のデジタル化にどのように貢献しているかを、具体的な事例とともに考察した。

デジタル化技術の進展により、歴史資料の保存や利用はますます進化し、史料の未来への伝承がより確実なものとなるであろう。

■ 関連リンク

» 「蔣洲咨文」テキストビューア（倭寇図巻デジタルアーカイブ）

<https://www.hi.u-tokyo.ac.jp/collection/digitalgallery/wakozukan/tei/>

» 策彦周良文集

<https://www.hi.u-tokyo.ac.jp/collection/digitalgallery/sakugenbunshu/>

» 東京大学史料編纂所・史料集版面検索（HI Lab）

<https://dip-s.lab.hi.u-tokyo.ac.jp/>

» 東京大学史料編纂所・史料集版面検索 Ngram Viewer

<https://dip-s.lab.hi.u-tokyo.ac.jp/ngram>

コメント 人文学における研究資源のデジタル化と持続可能性

山田 太造（東京大学史料編纂所）

1. はじめに

この十数年、人文学を取り巻く研究環境において、デジタル化は急速に進展し、重要な要素として定着してきた。特に、歴史学のような文献史料を扱う分野では、紙媒体の史料を撮影・スキャニングしてデジタル画像を作成することが一般的となっている。国文学研究資料館では、2014年度から2023年度にかけて、「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」を通じて約30万点の史料を対象に3000万コマを超えるデジタル画像を蓄積した。また、東京大学史料編纂所（以下、史料編纂所）でも、2008年に開始したJSPS科研費基盤研究（S）「史料デジタル収集の体系化に基づく歴史オントロジー構築の研究」年以降2100万コマを超える史料画像アーカイブを構築している。

デジタル化は、単にアナログからデジタルへの変換（digitization）のみならず、デジタル技術を活用して新たな研究手法や社会的仕組みを創出する「デジタルトランスフォーメーション（DX）」も含むものとなった。研究DXが実現することで、研究活動の各プロセスが変革し、従来の研究スタイルが進化すると同時に、新たな研究モデルが生まれる可能性がある。これにより、デジタル化は単なる技術的取り組みに留まらず、研究の在り方そのものを変える鍵となっている。

2. デジタルデータの作成

DXを実現するには、まず「digitization」を通じてデジタルデータを作成し、それを基に「digitalization」を進める必要がある。例えば、歴史資料の画像を作成し、それにOCR処理を施してテキストデータを出力することを考える。この解決手法の1つとして、画データをIIIF（International Image Interoperability Framework）に準拠した形式（例：Tiled Pyramid TIFF）に変換し、IIIF Presentation APIを用いて画像データを提示する。また、テキストデータはTEI（Text Encoding Initiative）/XMLを用いて整備し、ウェブ上での公開を行う。このようなプロセスにおいては、目的に応じてデータを作成するキュレーションツールが重要な役割を担う。また、キュレーションツール自体の開発・改善もリサーチエンジニアリングの一環として進める必要があり、デジタルアーカイブ構築の重要な要件の1つとして位置づけられる。

3. データ利用環境の整備

作成したデジタルデータを活用するには、「digitalization」による利用環境の整備が欠かせない。データの可視化や解析を通じ、データの価値を引き出す取り組みが求められる。オープンサイエンスの推進に伴い、他機関や他研究者が提供するデータも積極的に活用されるようになりつつある。これにより、データの二次利用が促進され、単なる論文や書籍からの引用に留まらず、データからデータへの引用が可能となる環境が整備されている。

このようなウェブ上でのデータ利用環境を実践していくため、FAIR原則（Findable：発見可能、Accessible：アクセス可能、Interoperable：相互運用可能、Reusable：再利用可能）を実現していくことが求められる。例えば、永続的な識別子であるDOI（Digital Object Identifier）の活用がその1つとしてあげられる。

また、データの二次利用促進のためには、データ利用条件の設定が不可欠である。例えば、クリエイティブ・コモンズ (cc) ライセンスを用いて利用条件を示す方法がある。特に、CC BY、CC BY-SA であれば、データ提供者へ問い合わせることなく、オープンデータとして利用できることができる。

■ 4. データ作成者と利用者間のギャップ

デジタルデータの利用を巡っては、データ作成者と利用者の間に期待のギャップが生じることがある。例えば、どのデータを残し、どのデータを省略するかといった判断は、利用者のニーズを考慮しつつも、データ作成者の提供したい内容をベースとして決定されている。例えば、史料編纂所は『大日本史料』などの史料集を出版しているが、現在は紙媒体をベースとして編纂している。研究 DX を推進する中で、ウェブをベースとした編纂データ提供方法を解決していく必要がある。さらに、AI や機械学習の進展に伴い、オープンサイエンスの一環として、機械可読形式（例えば TEI/XML によるデータ提供）でのデータ提供環境整備やメタデータの標準化が必須であると考えられる。

■ 5. データ公開と持続可能性

デジタル化を推進する上で、データ公開の持続可能性は避けて通れない課題である。データを公開・維持するためには、ハードウェアやストレージの確保、運用経費、人件費、そしてセキュリティ対策など、多岐にわたる課題が存在する。例えば、史料編纂所では約 300TB のストレージを運用しているが、その運用を支えるための経費や技術更新は継続的に必要である。しかし、これらの経費を運営費交付金でまかなうことは困難な状況にある。結果として、科研費などの外部資金への依存が増加しているが、多くの外部資金は維持費を目的とした申請に対応しておらず、新たな研究課題を提示し、採択されると、それを解決していく必要がある。

■ 6. 結論と展望

デジタル化は、人文学における新たな研究分野として「デジタル・ヒューマニティーズ」を形成しつつある。一方で、その発展には技術的・運用的な課題が多い。これらを克服するためには、データ作成者と利用者間の協力、持続可能な運用体制の構築、そして新技術や概念の導入が不可欠である。これらの取り組みを通じて、人文学研究 DX の実現に近づき、人文学におけるデータのみならず、研究そのものの重要性が高まると信じている。

コメント コメント

中村 雄祐（東京大学大学院人文社会系研究科）

いずれも、歴史研究が分厚く重い伝統を背負いつつ新しい視点を取り入れて前に進む現場を垣間見させてくれる充実した報告であった。

何よりもまず印象的なのは、筆記や印刷だけでは実現困難な様々な分析や可視化が次々と試みられており、いわば、プロの歴史研究者のための「理想の文房具」の開発が情報学研究者との連携を通じて活発に進められていることである。報告で示されるダッシュボードを見ていると精緻な印刷アトラスが思い出されるが、現代の「デジタル・アトラス」のカラフルでインタラクティブなデザインには目を見張るものがあり、すでに歴史研究のあり方に少なからぬ影響を及ぼし始めていることが実感できる。

デジタル化の進展は止めようもなく、今後も「歴史研究者のための理想の文房具」の開発が続いていくと思うが、同時に歴史研究の未来に関して新たな問いも浮かんでくる。

一つ目の問いは、長いモノグラフの未来である。従来より、理工系諸学が論文中心に研究を進めてきたのに対して、歴史学を含めて人文系諸学では長く分厚いモノグラフが中心的な位置を占めてきた。長いモノグラフを書き上げるには一次資料の精査を含めて長い修練が必要だが、それらを読みそして書く経験を通じてこそ体得できる歴史観や世界観が大切にされてきた。他方、近年のデジタル技術を駆使した研究には設計から環境構築、データ整備、分析、デザイン、維持管理まで膨大な時間、労力、そして資金が必要である。独りでできることは限られており、おのずと資料の精査も分業の一つとなる。また、出版ビジネスのレベルでも（印刷）学術書は苦境にあり、研究者キャリアにおいても人文系でも査読付き（電子）論文の比重が高まりつつある。デジタル・ダッシュボードを使ったマルチディメンショナル、マルチモーダル、カラフル、インタラクティブな思考は新しい歴史研究の可能性を垣間見せてくれるが、よりデジタル化された環境で育つ次の世代の歴史研究者たちは「語り続ける声」を軸とする長いモノグラフの伝統を今後どのように継承していくのか？そんな問いを考えさせられるシンポジウムであった。

心に浮かぶもう一つの問いは、ちょうど出版産業の停滞と対をなすように活況を呈する「歴史もの」ドラマ、映画、ゲームなどのエンタテインメントとの関係である。周知のごとく、エンタテインメント業界ではデジタル技術の活用が進み、その規模も拡大し続けている。もちろん過去の出来事の再現や演出をめぐる解釈の不一致や対立は避けられないが、その一方で、エンタテインメント業界の人々も公開された研究を精査し専門家に時代考証を依頼するなど、両者の関係も変わり続けている。プロの歴史研究者であれば従来より一般社会の「俗流」歴史観との付き合い方は心得ているであろうが、分厚いモノグラフに入れ替わるかのように存在感を増し続ける「歴史もの」デジタル・エンタテインメントとどんな関係を築いていくのだろうか？相互の対話が深まることを期待しつつ、今後の展開を興味深く見守っている。

ラウンドテーブル・質疑応答

記録：菊地 智博

■ ラウンドテーブル

山田氏・中村氏のコメントに対する各報告者の応答という形でラウンドテーブルを行った。

菊地は山田氏のコメントに対し、近世史編纂支援データベースでも編纂時の原稿形態（本文・標出・目次）に制約されてデータ形態が決まっているという点で、「データ作成者の提供したい内容」になってしまっていると述べた。また責任ある史料解釈を示すことが「編纂知」の意義と述べ、データ構造が逆に版面の表現に制約されることはあるが、それも含め DX による解決を進めたいと述べた。山田氏はウェブ上では紙と同じ意味の注記を紙史料集とは異なるウェブに適した表現で提供することも必要だと応答した。

山崎氏は経済学における再現性とは同じデータを同じ分析方法にかけることで同じ結果を得られる意味だと述べ、生データへのアクセスを可能とすることが再現性を担保する方法だとした。ハーバード大学が提供する Dataverse などのサイトや Nature のデータ専門ジャーナルの存在を紹介しつつも、利用者へのデータカタログの提供に大きな課題があることを指摘した。中村雄祐氏の提起するモノグラフの問題については、経済学においても分野を俯瞰するために必要な存在だと述べ、共著により専門家が協業することが重要だと指摘した。また経済学においても歴史学と同じく一般社会とのコミュニケーションが社会還元のために重要だと述べた。

高槻氏はこれに関連して『社会経済史学』における電子附録の試みを紹介した。典拠データを論文とは別に J-Stage で公開するもので、日本史全般においても史料編を電子附録として提供できる可能性に触れた。一方で、データの持続的提供には多額の費用が必要であり、国家的な支援が必要だとも指摘した。またデータ提供のギャップについては、エンタメ業界から求められるのは日常の物価情報であるのに対し、研究者は米価情報しか持っていない、といった例を紹介した。一方で、山崎氏の侍プロジェクトはエンタメ業界にもインパクトのある成果だと述べ、両者の歩み寄りに期待した。

中村覚氏は高槻氏に同意しデータやツールの持続的提供は個別の機関や個人には担いきれるものではなく、ここに規格の標準化が大きな役割を果たすだろうと指摘した。また人文系のデータに特有の多様性を標準規格にどう載せるかが、人文情報学の研究者の役割であると述べた。

■ 質疑応答

初めに、横山伊徳氏から報告を踏まえてのコメントがあった。菊地報告に関しては『井伊家史料』の索引データベースを作る過程で直面した課題として、人名の多様さに対して見出し語を一つに定める必要がある点を挙げた。特に女性名についてはジェンダーバイアス等の問題もあり、利便性のため一般に通用する名前を採用することには問題があるとした。また VIAF のような人名に関するデータセット・ツールの整備が必要だとも述べた。山崎氏・高槻氏の報告に関連しても、国立歴史民俗博物館の江戸商人名データベースや名古屋大学の人事興信録データベースなどのような商人名鑑の整備が必要であろうとした。中村覚氏の報告に関しては史料編纂所における構造化テキスト制作の重要性を指摘した上で、API の整備と活用、他機関との連携が今後の課題にな

ると指摘した。

高槻氏は横山氏への応答として、現在制作中の大坂の商人名データベースにおいても、屋号や通り名、当主名の扱いが難しいと例を挙げて述べた。

高島正憲氏はまずコメントとして自ら『平安遺文』などにあたった経験を挙げ、かつては限られた研究者しかアクセスできなかった原史料に学生レベルが触れられる現在、デジタル化時代の『遺文』のようなものの構築について分野をまたいで議論する必要があると述べた。続けて高槻氏報告への補足として、「近世経済データベース」の収録元の一つである三井文庫編『近世後期における主要物価の動態』初版と増補版においてデータ系列に大きな差異がある例を挙げ、利用者からはやや不便なものの、その変更の背景に同一資料・同一系列の原則という「編集知」のような哲学があったという気づきを述べた。そしてデータベース等でのデータ提供においてそうした哲学をいかに検証可能にするかが課題であるとした。さらに高島氏はデータベースの提供が進む一方で、歴史学の成果がデータの提供元としかみなされない現状、またデータベースが業績とみなされない問題について各報告者の意見を求めた。

高島氏の質問に対し山崎氏は、データベース提供は研究のインフラ整備のようなものとした上で、業績として評価はされずとも、データや史料に真面目に向き合うならばインフラに対してのアプリシエイトは生まれるはずだと述べた。高槻氏は、経済史においても実証経済学の「下請け」と反発する向きがあるとする一方、データの制作自体に高度なスキルが必要であり、自分はそれを貴重な作業と考えて進めていると答えた。菊地は高槻氏を受けて歴史学者の立場から、データベースは土管のようなもので、自分が作ったものを他人に提供し自分も他人の作った土管を使う相互主義だと考えていると述べた。また業績という観点は別にしても、データの正確性を示すための責任のクレジット明示は必要だとした。

藤野晋氏は高槻氏の報告に関して、シェイクスピアの全集において明らかな誤りを修正せず利用者に委任した例を挙げ、物価賃金データベースにおいて原史料の明らかな誤記をあえて残すことの意味を問うた。高槻氏はデータベースではデジタルの利点を活かし原史料通りの系列と編纂者の責任において修正した系列を同時に提供していると述べた。基本的には後者を利用する想定であるが、万一のため原系列や原史料へのアクセスも可能にしている、と高槻氏は答えた。

以上にて質疑応答は終了した。

最後にシンポジウム企画の杉本氏から挨拶があり、シンポジウムは閉会となった。



2

デジタライジング研究会

■ 研究会の概要 ■

本科研では、史料集編纂を紙媒体・デジタルデータの両面から捉えなおすことを目的として「デジタライジング研究会」を開催してきた。科研メンバーを中心としたメンバーにおいてデジタル化に関わる情報を共有するとともに、外部から講師を招き、最新動向へのキャッチアップを図った。

本研究会の成果を元にして、2024 年度のシンポジウム「[デジタル化と歴史研究の未来](#)」が開催されたほか、システム開発なども行われている。

■ コア・メンバー

本科研メンバー（前掲参照）

高槻泰郎（神戸大学経済経営研究所）

中村雄祐（東京大学大学院人文社会系研究科文化資源学研究専攻）

山崎潤一（京都大学大学院経済学研究科）

■ 報告の記録 ■

林 大樹（第 1～4 回）・黒滝 香奈（第 6～8 回）・菊地 智博（第 5・9～11 回）

■ 第 1 回 NDL OCR の紹介と応用—OCR 技術の最前線

■ 日時：2022 年 6 月 28 日（火）15:00～16:00 Zoom

■ 講演者：中村覚（東京大学史料編纂所助教）

■ 参加者数：11 名

■ 内容概略

各自の自己紹介ののち、30 分間の報告、および 30 分間の質疑を行なった。報告では、国立国会図書館による NDL OCR の機能、および編纂所 DB と連携させた試験的システム「史料集版面検索」の概要が紹介された。質疑では、レイアウト認識の互換性や、誤認文字の学習機能、今後の刊行物への影響などが議論された。また杉本氏から、傍注内容も認識対象とするよう提案がなされた。

【付記】箱石氏の提案を受け、本研究会の成果に基づき、中村覚氏「幕末維新史料・横断検索システム」が構築された（[後述](#)）。

■ 参考 URL

» NDLOCR

https://github.com/ndl-lab/ndlocr_cli

» 幕末維新史料・横断検索システム

<https://ishin.lab.hi.u-tokyo.ac.jp/>

■ 第 2 回 暗黙知の可視化とデジタル・ヒストリー

■ 日時：2022 年 11 月 7 日（月）10:00～11:00 Zoom

■ 講演者：小風尚樹（千葉大学人文社会科学系教育研究機構助教）

■ 参加者数：10 名

■ 内容概略

約 30 分間の報告に続き、約 40 分間の質疑を行った。報告では経験に基づく史料のコンテキストへの「暗黙知」を可視化するための史料検索・構造化の方法について、伝統的なツリー構造と様々なグラフ構造の試みを紹介、両者を両立させる UI の開発が必要だとまとめた。

司会から、紙媒体を前提とした作成者・編纂者（史料集に編纂）の営為を、デジタルデータとして表現することの意味を問題としている本科研にとって示唆に富む報告であったとの位置づけが示され、大要以下の議論がなされた。本科研質疑ではツリー構造に関する生物学・情報学分野での議論が紹介されたほか、史料編纂所刊行物の構造化のあり方について議論がなされた。

■ 参考文献

» マニュエル・リマ（三中信宏訳）『THE BOOK OF TREES: 系統樹大全』ビー・エヌ・エヌ新社、2015 年

» マニュエル・リマ、（翻訳）奥いずみ、（監修）久保田晃弘『ビジュアル・コンプレキシティ—情報パターンのマッピング—』ビー・エヌ・エヌ新社、2012 年

» 西村陽子・北本朝展「ディジタル史料批判と歴史学における新発見」『人工知能』31 巻、2016 年、
https://doi.org/10.11517/jjsai.31.6_769

■ 第 3 回 日本史資料デジタル公開・共有の可能性と課題—制度・「コミュニケーション」の議論を中心に—

■ 日時：2022 年 12 月 6 日（火）13:30～14:30 Zoom

■ 講演者：後藤真（大学共同利用機関法人 人間文化研究機構 国立歴史民俗博物館 研究部 准教授）

■ 参加者数：13 名

■ 内容概略：

約 30 分間の報告に続き、約 30 分間の質疑を行った。報告ではデジタルヒューマニティーズ(DH)における史料のデジタル「公開」を「共有」としていくための議論の材料を示され、人文学や情報学の対話、「ルール」作りの必要性を提起された。

質疑ではまず、デジタル化された史料画像の欠落・改変の恐れが指摘され、所蔵機関による現物公開が担保されていることが大事である旨の回答があった。また史料の有するセンシティブな情報をめぐり、公開範囲やアクセス権限の深度についても意見が交換された。大量分析のためのマス・データの提供は魅力的であるが、個人情報の保護とのせめぎあいも重要であり、今後は「公開」・「共有」をめぐる共通「ルール」作りが必要であることが確認された。

■ 第 4 回 歴史資料等の公開における二次利用条件表示について

■ 日時：2023 年 1 月 10 日（火）19:00 ～ 20:00 Zoom

■ 講演者：南亮一（一橋大学大学院法学研究科法学・国際関係専攻博士後期課程（応用社会コース））

■ 参加者数：11 名

■ 内容概略：

約 30 分の報告に続き、約 40 分間の質疑を行った。報告では二次利用条件表示の前提をフローチャートで整理したうえで、二次利用条件表示のパブリック・ドメイン・ツール、CC（クリエイティブ・コモンズ）ライセンスの特徴について主にまとめられた。

質疑では、所蔵者と権利所有者が異なることの多い歴史資料（および複製物）のデジタルデータの公開について議論となった。権利の効力や表示方法、周知をめぐっては、なお議論を進めていく必要があることが確認された。

■ 第 5 回 幕臣職歴のデータセット構築とその数量的把握

■ 日時：2023 年 4 月 10 日（月）13:00 ～ 14:00 Zoom

■ 講演者：高槻泰郎（神戸大学経済経営研究所准教授）・

山崎潤一（神戸大学大学院経済学研究科講師）

■ 参加者数：13 名

■ 内容概略：

35 分間の報告、および 25 分間の質疑を行なった。

報告では『柳営補任』・『寛政重修諸家譜』のデータ化による幕府人事の包括的な検討について、まずその方法や設計が整理され、続いて刊本をデータ化するにあたっての課題が示された。最後には成果として幕府人事の時代変化グラフが紹介された。

質疑では、「同断」等の判定にあたって参考となる史料や、『柳営補任』刊本編纂時に行われた修正について参加者から紹介があった。またデータ化の方式やグラフの内容、今後のデータ活用について議論がなされた。

■ 参考文献

» コウゾウ・ヤマムラ『日本経済史の新しい方法 徳川・明治初期の数量分析』ミネルヴァ書房、1976 年

» 小川恭一『徳川幕府の昇進制度－寛政十年末旗本昇進表－』岩田書院、2006 年

■ 第 6 回 Rights Statements の意義と権利表記をめぐる動向

■ 日時：2023 年 6 月 16 日（金）13:30 ～ 14:30 Zoom

■ 講演者：数藤雅彦（五常総合法律事務所）

■ 参加者数：11 名

■ 内容概略

35 分間の報告、25 分間の質疑を行った。

報告では、まずクリエイティブ・コモンズ（CC）と Right Statements（RS）の違いと日本における使用事例が示された。その上で、デジタルアーカイブにおける権利表記の動向と、望ましい権利表記が紹介された。

質疑では、CC や RS マークが示す意味や具体的な使用方法について質問が集まった。また、CC と RS の併存に対する理解の仕方について、活用する館として、さらには資料使用者の側として議論がなされた。

■ 参考 URL

» クリエイティブ・コモンズ・ジャパン

<https://creativecommons.jp/>

» Right Statements

<https://rightsstatements.org/en/>

■ 参考文献

» 数藤雅彦「Rights Statements と日本における権利表記の動向」『カレントアウェアネス』

No.343、2020 年、<https://current.ndl.go.jp/ca1973>

» 大井将生・渡邊英徳「デジタルアーカイブ資料の活用を促進する二次利用条件のあり方」『デジタルアーカイブ学会誌』7 巻 3 号、2023 年、https://doi.org/10.24506/jsda.7.3_e24

■ 第 7 回 「「デジタルアーカイブ活動」のためのガイドライン及びデジタルアーカイブアセスメントツール（ver3.0）をふまえた権利処理・権利表記の留意点

■ 日時：2023 年 11 月 28 日（金）13:00 ～ 14:00 Zoom

■ 講演者：数藤雅彦（五常総合法律事務所）

■ 参加者数：19 名

■ 内容概略

40 分間の報告、20 分間の質疑を行った。

報告では、まず著作権の基礎知識やクリエイティブ・コモンズ（CC）と Right Statements（RS）の概略が紹介された。その上で、「「デジタルアーカイブ活動」のためのガイドライン」及び「デジタルアーカイブアセスメントツール（ver3.0）」のうち、著作権等の権利処理と権利表記に関する部分が解説された。

質疑では、著作権の解釈に関する具体的なケースについての質問があった。また、ガイドラインにおけるメタデータの解釈や、RS の位置付けについても確認があった。最後に学術機関としてのデジタルコンテンツオープン化への対応の仕方についての質疑がなされた。

■ 参考 URL

» 「デジタルアーカイブ活動」のためのガイドライン

https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/jitumusya/dai16/siryoku1-1.pdf

» デジタルアーカイブアセスメントツール (ver 3.0)

https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/files/assessment_tool_2023.xlsx

» デジタルアーカイブ戦略懇談会及びデジタルアーカイブ推進に関する検討会

https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/index.html

■ 第 8 回 武鑑全集：空間的・時間的な構造化データを構築する効率的なワークフローの探求と江戸長期時系列データの構築

■ 日時：2024 年 1 月 25 日（金）16:00 ～ 17:00 Zoom

■ 講演者：北本朝展（国立情報学研究所）

■ 参加者数：10 名

■ 内容概略

40 分間の報告、20 分間の質疑を行った。

報告では、江戸時代の武鑑を対象に、空間的・時間的な構造化データ構築のプロセスが提示された。具体的には、IIIF Curation Platform や vdiff.js などのツールの活用の実態や、版比較アルゴリズムを用いた「差読」によるデータ構造化の実践例が紹介された。最後に、江戸長期時系列データ構築の展望が述べられた。

質疑では、経時的差分翻刻をめぐって、データの活用方法や手順の妥当性について確認があった。また、ワークフローの過程で用いるインフラとその活用方法についても質問がなされた。最後に「藩 ID」の妥当性について質疑がなされた。

■ 参考文献

» 北本 朝展・藤實 久美子・本間 淳「ブックバーコーディング法：版本の差読に基づく『武鑑全集』の網羅的な解析に向けて」『人文科学とコンピュータシンポジウム じんもんこん 2021 論文集』2021 年、pp. 268-275、<http://id.nii.ac.jp/1001/00215692/>

» 北本 朝展・堀井 洋・堀井 美里・鈴木 親彦・山本 和明「時系列史料の人機分担構造化：古典籍『武鑑』を参照する江戸情報基盤の構築に向けて」『人文科学とコンピュータシンポジウム じんもんこん 2017 論文集』、2017 年、pp. 273-280、<http://id.nii.ac.jp/1001/00184666/>

■ 第 9 回 近世史料のデジタル化と藩・大名家 ID 付与の試み

■ 日時：2024 年 5 月 29 日（水）13:30 ～ 14:40 Zoom

■ 講演者：北本朝展（国立情報学研究所）・

山崎潤一（神戸大学大学院経済学研究科講師）

■ 参加者数：18 名

■ 内容概略

両報告者から 15 分ずつ、合計 30 分間の報告ののち、40 分間の質疑を行った。

北本氏の報告「江戸全藩を対象とした藩 ID の構築」では、武鑑全集の作業に際して付与した「藩 ID」について報告され、地理的エンティティ（実体）である藩に対し社会的エンティティである大名家 ID を別に定義し紐づける必要があること、また人物 ID の定義や歴史地理的な活用など歴史ビッグデータの将来像まで展望した。

山崎氏の報告「藩・大名家データベースについて」では大名の転封研究に際して付与した「藩 ID」について報告された。石高の変遷など経済社会的データと接合する観点からは藩 * 年レベルのテーブルが必要であること、また藩 ID の定義が困難ないくつかの例が紹介された。

質疑では ID を付与すべき「藩」の定義について活発な意見が交換され、一律の定義では割り切れない例が近世初期・明治期に存在することが紹介された。辞典を元にした機械的処理の課題が指摘される一方、何らかの書籍をもとにした基準の必要性で議論が一致した。

■ 第 10 回 グラフによる史料の構造化

■ 日時：2024 年 10 月 15 日（火）13:30 ～ 14:40 Zoom

■ 講演者：小風尚樹（千葉大学人文社会科学系教育研究機構助教）

■ 参加者数：11 名

■ 講演概要

45 分の報告ののち、25 分の質疑を行った。

報告では、まず TEI でも用いられるツリー構造ではテキストの構造化が難しい事例について、三つの類型に分けて紹介された。TEI では近年ツリー構造外の情報を記述する構文が導入されており、事例として会計史料をグラフ構造で構造化するために開発された DEPCHA が紹介された。このほか、学籍簿史料の TEI による構造化において、外部情報の参照をグラフ表現も用いて記述した事例も取り上げられた。

質疑では、Excel 等の表形式に対する DEPCHA の利点についてまず質問があり、記述は複雑だが計算の利便性とテキストの文脈を両立が可能だと小風氏は答えた。また両立することにより判明する史料の興味深さや、グラフ構造の可能性についても言及がなされた。さらに構造化の際に曖昧な解釈を許容するかという問いに小風氏は、TEI では解釈を記録することを重視しており、解釈のアップデートを共有するプラットフォームともなると答えた。

参考文献

- » 小風尚樹「会計史料の構造化：ツリーかグラフか」『情報知識学会誌』32 巻 4 号、2022 年、
<https://doi.org/10.2964/jsik.2022.029>
- » Elli Bleeker, et al., “Texts as Hypergraphs: An Intuitive Representation of Interpretations of Text”, Journal of the Text Encoding Initiative [Online], Issue 14, 2022,
<https://doi.org/10.4000/jtei.3919>
- » DEPCHA - Digital Edition Publishing Cooperative for Historical Accounts
<https://gams.uni-graz.at/context/depcha>

第 11 回 西洋史から見た日本史の研究環境：DH 資源を個人研究に活用する可能性と限界

■ 日時：2024 年 12 月 17 日（火）13:30 ～ 14:40 Zoom

■ 講演者：小風綾乃（東京大学史料編纂所 前近代日本史情報国際センター 特任研究員）

■ 参加者数：9 名

講演概要

30 分の報告ののち、40 分間の質疑を行った。

報告では、歴史学の方法論に基づくミクロストーリー的な DH について報告者の西洋史における実践例が紹介された。Transkribus による AI 翻刻や翻刻困難な手稿の例、また史料から作成したデータベースを探索し、総体把握と同時に個別事例への理解を深める手法が技術的背景とともに紹介された。一方で DH ではデータ整備に関する議論が中心で個人レベルでの活用法が深められていないこと、日本史においては AI 翻刻の難しさやオープンデータの整備が不十分だという課題が指摘された。小さなデータセットの活用例やデータ作りのノウハウ共有が必要だと小風氏は締めくくった。

質疑では、Digital Humanities と Digital History の間のギャップが指摘されつつ、異分野間の研究者が DH を通じて相互作用する可能性にも言及された。小風氏が取り組む『大日本史料』における重層化された時間や TEI による分析について、編纂者の立場からのコメントもあった。情報学の立場からは OCR の精度向上のほか、DB にナレッジグラフを応用する手法などについてコメントがあった。

参考文献

- » 小風綾乃・中村覚・山田太造「TEI による編纂史料の構造化」『人文科学とコンピュータシンポジウム論文集』Vol.2024 No.1、2024 年、<https://ipsj.ixsq.nii.ac.jp/records/241540>

3

公開データ

■ 1. 近世史編纂支援データベースの拡充 ■

■ 近世史編纂支援データベースとは

菊地 智博

» <https://wwwap.hi.u-tokyo.ac.jp/ships/w30/search>

近世史編纂支援データベースは、東京大学史料編纂所歴史情報処理システム（SHIPS）上で公開するデータベースの一つで、近世史料部門が出版している史料集の編纂支援を目的とする。本所出版物を構成する目次・本文・索引・標出の4要素を検索可能であり、2025年3月現在8つの書目についてデータを公開している。編纂支援用ではあるが、研究用としても十分に利用できる内容・機能を備えており、「版面ギャラリー」で公開されている刊本の画像（最新巻を除く）、Hi-CATで公開されている原本の画像（一部のみ）へのアクセスも可能である。

以下、各検索型について説明する。

【目次型】各書目の目次部分を検索できる。文書名に加え、和暦での年代や差出・宛所、出典等の情報を収録している。（※2024年度、本科研において「目録型」から改称する改修を実施）

【索引型】各書目に出現する人名情報と関連情報を検索できる。簡易的な人名データベースとしても機能する。

【標出型】各書目の版面に頭注として掲出している見出し・要約等（標出）を検索できる。

【本文型（目次）】電子組版データをもとに、各書目の目次部分を検索できる。

【本文型（本文）】電子組版データをもとに、各書目の本文を検索できる。

各書目・各検索型について、2025年3月現在のデータ搭載状況は下表の通りである。

書目	既刊	目次型	索引型	標出型	本文型 （目次）	本文型 （本文）
井伊家史料	30巻（完）	1～30	1～23	-	24～30	24～30
市中取締類集	32巻	1～2	-	-	-	-
広橋兼胤公武御用日記	15冊	-	-	1～6	-	8～15
細川家史料	28冊	-	1～16	-	17～28	17～28
幕末外国関係文書	54巻	1～54	47～49	-	-	-
斉藤月岑日記	10巻（完）	-	1	1	-	-
諸問屋再興調	15巻（完）	1～15	-	-	-	-
松平昭休往復書翰留	1巻	1	1	-	-	-

本科研では、『井伊家史料』『市中取締類集』『細川家史料』『松平昭休往復書翰留』を対象として、データの大幅な拡充を実施した。以下、拡充の成果について書目ごとに詳述する。

■ 井伊家史料

菊地 智博

『大日本維新史料 類纂之部 井伊家史料』は、彦根藩井伊家に伝来した井伊家文書（現在、彦根城博物館所蔵）のうち、井伊直弼関係文書を中心に編纂したものである。幕末期の政治に関する基幹史料集として、『大日本維新史料 類纂之部』の一書目として1958年度から刊行している。

近世史編纂支援データベースでは目次型において全30巻の目次、索引型では人名、本文型では全文テキストからの検索が可能である。本科研開始以前の2019年度までに目次型（全30巻）・索引型（1～6・19巻）・本文型（24～30巻）のデータを搭載済であった。

本科研では索引型データの拡充を進め、表に示す通り7～18・20～23巻のデータを新規登録、また3・5・19巻について情報の追加を行った。以上により、電子組版を採用した24巻より前の全巻について索引型データを完備した。作成にあたっては、既存の人名傍注以外の人名出現箇所や未比定だった人名についても新たに採録を行い、維新期の人名データセットとしても利用できるよう考慮している。なお、24巻以降の索引データについても、23巻以前と形式や基準に違いはあるものの、本文型（本文）検索から利用可能である。

データ作成は、学術専門職員の林大樹氏（8～12・14巻）・黒滝香奈氏（15～17巻）・世川祐多氏（7・13・18・20・22・23巻）・高橋直大氏（13・17巻）・菊地智博（20・21巻）が担当し、全般的な監修と整形を菊地が行った。採録基準や比定については、杉本史子氏・箱石大氏が監修した。なお本科研以前のデータ整備に吉川紗里矢氏・横山伊徳氏の貢献があったこともここで特記しておく。

巻数	データアップ	担当	レコード数
7	2024年度	世川	1,961
8	2022年度	林	2,201
9	2022年度	林	2,967
10	2022年度	林	2,628
11	2023年度	林	3,228
12	2023年度	林	2,489
13	2024年度	世川・高橋	3,103
14	2022年度	林	2,436
15	2023年度	黒滝	2,095
16	2023年度	黒滝	1,871
17	2024年度	黒滝・高橋	4,949
18	2023年度	世川	2,424
20	2023年度	菊地・世川	2,799
21	2023年度	菊地	2,145
22	2024年度	世川	2,302
23	2024年度	世川	2,015

■ 松平昭休往復書翰留

菊地 智博

『大日本維新史料 類纂之部 松平昭休往復書翰留』は、本所所蔵「松平昭休往復書翰留」（請求記号4344-1）と関係史料を編纂するもので、2023年度、新たに刊行を開始した。主に徳川斉昭九男松平昭休が忍藩世子松平忠矩であった安政5年（1858）から岡山藩主池田茂政となる直前の文久3年（1863）までに各藩主や藩主夫人となった兄弟姉妹らと交わした書状を収録している。

本書目は電子組版による編纂過程において近世史編纂支援データベースへのデータ搭載を前提とした作業フローをとっており、2024年度に第1巻の組版データを元に作成した目次型データ（417件）・索引型データ（2,805件）を新規に登録した。以後も新刊時に当該巻の目次・索引型データ、前巻の本文型・標出型データを登録する予定である。なお、原稿作成に用いたタグ記法については中村覚氏が監修し、データの抽出・整形は菊地が担当した。本書目に関わる技術的な部分については、2024年度シンポジウムの報告記事も参照されたい。

■ 細川家史料

林 晃弘

『大日本近世史料 細川家史料』は、肥後熊本藩細川家に伝来し、現在公益財団法人永青文庫が所蔵する史料を底本とする江戸時代前期の藩政・幕藩関係に関する基幹史料であり、1969年より刊行している。本所のデータベースのうち、「近世史編纂支援データベース」において「索引型」で人名、「本文型」で全文テキストから検索でき、「本文型」分は「古文書フルテキストデータベース」でも利用可能である。

本科研の開始段階では、既刊27巻のうち、1巻～16巻を「索引型」、19巻以降を「本文型」でデータを公開しており、17巻・18巻については検索用のデータを欠く状況にあった。そこで、この2巻分を対象に全文テキストデータを作成した。

2022年度に18巻（本文456頁、書状案406通）、2023年度に17巻（本文484頁、書状案492通）を対象とし、タグ付け指示等の事前作業を行ったうえで、株式会社カンテックにテキスト入力とXMLデータ・管理データの作成を依頼した。納品後にデータを確認し、情報支援室に作業を依頼して、データベースに搭載した。これにより既刊分の全てについて検索可能なデータが揃うことになった。

また、本所では1960年代から永青文庫所蔵の古文書類の撮影をマイクロフィルムにより進め、現在もデジタルカメラにより継続している（マイクロフィルムはスキャニングを行い、デジタル化を行っている）。これらの史料画像について1点ごとの目録情報を付与し、「所蔵史料目録データベース」等で検索するためのデータを蓄積した。これらは順次公開を進める予定であり、また、将来的に「近世史編纂支援データベース」等とも連携できるものと考えている。なお、本データ作成作業は高垣亜矢氏（学術専門職員）が担当した。

■ 市中取締類集

立石 了

『大日本近世史料 市中取締類集』は、国立国会図書館に架蔵されている旧幕府引継書類の一部である市中取締類集を収めた、市政・社会・経済・風俗など各方面における幕末期の江戸の基幹的史料集である。市中取締類集は、江戸の町奉行所において行政の参考となる書類を部類別に整理・編纂したもので、その年代は天保12（1841）年～文久3（1863）年にわたり、先例や関連史料としてこれ以前の年代の書類も採録されている。『大日本近世史料』シリーズとして、既に32巻が刊行されている。

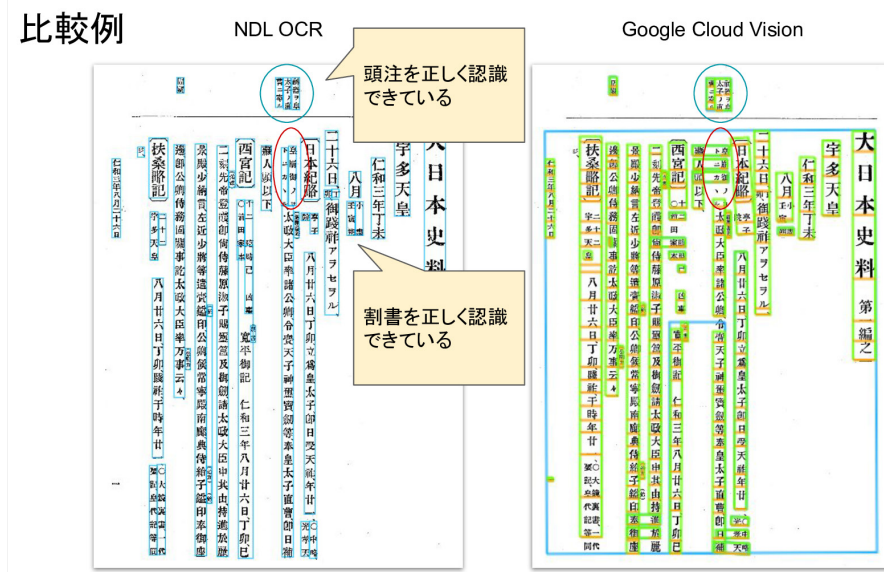
本所のデータベースにおいては、「近世史編纂支援データベース」の目次型から、第1巻と第2巻の目次データが公開済みとなっている。本科研では、これに引き続き第3巻以降のデータを「近世史編纂支援データベース」から公開していくために、目次データの作成を行うこととし、2023年度に、第3巻～第6巻の目次データ合計896件を作成した。公開済みのデータや各巻毎のデータの整合性を確認したうえで、今後「近世史編纂支援データベース」に登載する予定となっている。なお、本データの作成作業は竹村到氏（学術専門職員）が担当した。

2. 幕末維新史料・横断検索システム

中村 覚

幕末維新史料・横断検索システムは、本科研のデジタルライジング研究会をきっかけに開発が開始され、東京大学史料編纂所 HII Lab から公開されている、OCRを用いた全文検索システムである。

システムは「東京大学史料編纂所・史料集版面検索」を元になっている。史料編纂所では版面ギャラリーにて刊行物の版面画像を公開しているが、画像内の文字列に対する検索が行えない点が課題であった。同システムでは、国立国会図書館が公開している版面向け OCR 処理プログラム DLOCR を用いることで、日本語刊行物の構造を踏まえた高精度の OCR を実現した。



幕末維新史料・横断検索システムは、幕末維新期の史料に特化したシステムとして、幕末維新期に関する史料編纂所刊行物に加え、国立国会図書館が公開する『孝明天皇紀』および日本史籍協会叢書の OCR 結果を検索することができる。追加された資料は史料編纂所の出版物と共通する性格を持つほか、いずれも活字印刷で、史料編纂所の刊行物とレイアウトに共通点が存在する。

2025 年 3 月現在、146,261 ページ分の OCR データを搭載済みである。

幕末維新史料・横断検索システム

東京大学史料編纂所が公開する『大日本維新史料 編年之部』『大日本維新史料 類纂之部』『大日本古文書 幕末外国関係文書』『復古記』『明治史要』『保古飛呂比』『明治維新史料選集』『維新史』などに加えて、『孝明天皇紀』と日本史籍協会叢書の OCR テキストを検索



本システムは、デジタルライジング研究会（科学研究費補助金・基盤研究（B）「日本近世史料学の再構築——基幹史料集の多角的利用環境形成と社会連携を通じて（研究代表者：杉本史子）」（課題番号 22H00692）、2022 年 6 月 28 日オンライン開催、を契機として開発がスタートしました。

※本稿は中村覚「OCR テキストを用いた幕末維新史料・横断検索システムの開発」（明治維新史学会例会、2022 年 10 月 29 日）報告資料を元に菊地が構成した。

4

論文・口頭報告等

■ 2022 年度 ■

■ 論文

- » 杉本史子「書評と紹介 米家志乃布著『近世蝦夷地の地域情報：日本北方地図史再考』『日本歴史』897号、2023年、pp.98-100

■ 口頭報告

- » 中村覚「OCR テキストを用いた幕末維新史料・横断検索システムの開発」明治維新史学会例会、2022年10月29日、オンライン
- » 林晃弘「近世前期における熊本藩の真宗統制」第72回佛教史學會學術大會、2022年11月6日、龍谷大学大宮学舎
- » 中村覚「東京大学史料編纂所出版物を用いた NDLOC の応用事例紹介」2022年度 NDL デジタルライブラリーカフェ、2023年1月18日、オンライン
- » 箱石大「OCR テキスト横断検索システムへの期待と今後の幕末維新史研究」2022年度 NDL デジタルライブラリーカフェ、2023年1月18日、オンライン

■ 図書

- » 杉本史子『絵図の史学：「国土」・海洋認識と近世社会』名古屋大学出版会、2022年
- » 稲葉継陽・清水克行共編『村と民衆の戦国時代史—藤木久志の歴史学—』勉誠出版、2022年
- » 熊本大学永青文庫研究センター編『永青文庫叢書 細川家文書 災害史料編』吉川弘文館、2023年

■ 2023 年度 ■

■ 論文

- » 林晃弘「近世前期における熊本藩の真宗統制」『古文書研究』96号、2023年、pp.57-74
- » 林晃弘「書評 北陸中世近世移行期研究会編『地域統合の多様と複合』『日本史研究』728号、2023年、pp.64-72

■ 口頭報告

- » 林晃弘「細川忠利の『公儀御書案文』と近世政治史研究」シンポジウム「近世初期における「御国」と「公儀」—永青文庫細川家文書から—」、2023年9月16日、熊本大学工学部百周年記念館
- » 稲葉継陽「元和～寛永期における「御国」統治と文書・記録 —ポスト戦国世代の領国支配—」

シンポジウム「近世初期における「御国」と「公儀」—永青文庫細川家文書から—」、2023 年 9 月 16 日、熊本大学工学部百周年記念館

■ 2024 年度 ■

■ 論文

- » 林晃弘「近世前期の寺院相論と幕藩領主—防長地域の曹洞宗の事例から—」『歴史評論』894 号、2024 年
- » 林晃弘「細川忠利の花押の変遷—元和・寛永年間における—」『東京大学史料編纂所画像史料解析センター通信』106 号、2025 年

■ 口頭報告

- » 林晃弘「「越中真宗の江戸時代」塩崎報告・袖吉報告へのコメント」越中史壇会 令和 6 年度特別研究発表会「越中真宗の江戸時代」、2024 年 6 月 9 日、富山県民会館 701 号
- » 菊地智博「編纂知の DX—『井伊家史料』『松平昭休往復書翰留』編纂とデータベース構築の経験から」シンポジウム「デジタル化と歴史研究の未来 —人文学・社会科学の協働—」、2024 年 9 月 7 日、東京大学史料編纂所大会議室
- » 山崎潤一「社会科学から見た歴史データ」シンポジウム「デジタル化と歴史研究の未来 —人文学・社会科学の協働—」、2024 年 9 月 7 日、東京大学史料編纂所大会議室
- » 高槻泰郎「経済史研究における基盤的数量データの公開・活用」シンポジウム「デジタル化と歴史研究の未来 —人文学・社会科学の協働—」、2024 年 9 月 7 日、東京大学史料編纂所大会議室
- » 中村覚「歴史資料のデジタル化を支える技術」シンポジウム「デジタル化と歴史研究の未来 —人文学・社会科学の協働—」、2024 年 9 月 7 日、東京大学史料編纂所大会議室
- » NAKAMURA Satoru “Technologies Supporting the Digitalization of Historical Materials”, International Symposium Commemorating the 100th Anniversary of the Toyo Bunko: Accumulation of Asian Knowledge and East-West Exchange, November 17, 2024, Toyo Bunko
- » KIKUCHI Chihiro ““Historiographical knowledge” at the Historiographical Institute, The University of Tokyo, and Digital Transformation”, International Symposium Commemorating the 100th Anniversary of the Toyo Bunko: Accumulation of Asian Knowledge and East-West Exchange, November 17, 2024, Toyo Bunko
- » Fumiko SUGIMOTO “(Comment)“Compilatory Knowledge” and Digital Data”, International Symposium Commemorating the 100th Anniversary of the Toyo Bunko: Accumulation of Asian Knowledge and East-West Exchange, November 17, 2024, Toyo Bunko
- » 杉本史子「近世絵図の特質と近世社会」基盤研究 (A)「大型絵図類のデータ構造化と関連史料の連携による南西諸島「海上の道」の復元的研究」(23H00012、代表黒嶋敏) 例会、2025 年 2 月 21 日、東京大学

東京大学史料編纂所研究報告 2024-1

日本近世史料学の再構築

—基幹史料集の多角的利用環境形成と社会連携を通じて—

編・発行：

2022-24 年度 JSPS 科学研究費補助金 基盤研究 (B) 「日本近世史料学の再構築—基幹史料集の多角的利用環境形成と社会連携を通じて—」 22H00692/23K21964 (研究代表者：杉本史子)

初版：2025 年 3 月 31 日

組版・レイアウト：菊地智博