

# 日本史学と情報技術 —30年が変わったこと、変わらなかったこと—

後藤 真†

**概要**：本報告では、日本史学にたいして、どのような場面で情報技術が用いられるようになったのか、それにより変わったことは何であり、変わっていないことは何なのかについて、概観を試みる。日本史学におけるデータベースの歴史は古く、1980年代にはすでに大型計算機によるデータベース構築が行われ、その後もさまざまな場面での活用が試みられてきた。その活用のあり方は、大きくはデータベース目録検索の時代、辞書と史料集のCD-ROMの時代、データベース「乱立」の時代、遍在するコンピュータの時代と大きく4つ程度に区分することが可能であるという暫定的な見通しを持っている。このそれぞれの状況が「デジタルアーカイブ」などどのように切り結ばれてきたのか、今後の未来のありようなどについて、現状での見通しを述べたい。

**キーワード**：日本史学への応用 デジタルアーカイブ 研究史

## Information Science Applied to Studies in Japanese History: Overview from Changes or No-changes in the Last 30 Years

MAKOTO GOTO†<sup>1</sup>

This paper offers an overview of information science applied to studies in Japanese history, and will examine changes and no-changes using this information science. Databases of studies in Japanese history has a long history. Constructing databases by large scale computers was already developed in 1980's, and then the use of databases has been applied to various issues. Their uses may be divided into four categories: the time to search on database catalogues, time to use CD-ROM of dictionaries and historical sourcebooks, time to "upsurge" of database, and time for ubiquitous computers. This paper shows the current perspectives in terms of how to compete these situations with projects like "digital archives" and the further future methodologies.

**Keywords**: Applied to Studies in Japanese History, DigitalArchives, history of study

### 1. はじめに

本報告は、日本史学に対して情報技術がどのような影響を与えたのかについて述べるものである。日本史学における情報技術の適用については、古くは東京大学史料編纂所や国立歴史民俗博物館がデータベース構築を行っていた1980年代にまでさかのぼる。しかし、実際に多くの研究者の行動に影響を与えようようになるのは、1990年代末以降であると理解してよいと考える。そこで、本稿では、日本史学における基本的な「所作」を再確認しつつ、その中でデータベースがどのようにかわり、どのような影響を与えたのか、もしくは与えなかったのかについて、検討を加えるものである。

### 2. 日本史学研究を進める際の基本的な所作と情報技術

日本史学における研究論文執筆の際のプロセスはおおむね下記のとおりである。なお、研究者や分野、時代等によって違いがあるため、必ずしもすべての研究者が同一のプロセスを経るものではないことは注意されたい。

#### (1) a 問題意識の設定

当然であるが、まずはどのようなことを明らかにするためのスコープを設定する。その際には、多くの史料・資料や先行研究の中の課題から抽出されることになる。この部分では、史料の総覧や学術論文を多数読むという、ある意味では基礎的な動作であり、具体的な情報技術の関与があるわけではない。

#### (2) a 先行研究の網羅

いわゆる歴史系の学術論文では、対象となる先行研究には、原則としてその論文のスコープとなるすべてについて言及するという原則がある。主要なもののみではなく、すべての研究論文という点がポイントとなる。

そのため、研究者は関連文献を原則としてすべて「知っておく」必要がある。そのため、財団法人史学会が毎年6月号に出している「回顧と展望」は先行研究サーベイの際に最も基礎的な文献となっている。

#### (3) a 関連資料の探索

テーマと関連する史料・資料を原則としてすべて探索する。この資料の適切な発見と関連付けが論文の全体の手続き的成否を決めるといってもよい。この資料の探索範囲に関しては、古代史であれば基礎文献であるいくつかの編纂

†1 国立歴史民俗博物館  
National Museum of Japanese History

史料から始まり、木簡や正倉院文書、漢籍や仏典、関連する説話史料などにおよぶ。中世史であれば、編纂史料をはじめとする基礎資料に加え、寺院や博物館などにある文書類や絵図類などが対象となり、近世史のうち幕藩関係であれば各地域にある大名関係資料などや編纂史料類など、また村方関係であれば、地域の村落にある古文書などを用い、これに加えて既存の博物館の文書類などを活用する。近代史となると、公文書館などのアーカイブズ史料と地域資料などが対象となるが、その対象範囲は日本国内のみならず海外にも及ぶ。この範囲の設定と、適切な情報抽出が、日本史論文の実証性のカギを握っているといえるであろう。

#### (4) a 関連資料のうち、必要な資料の抽出

これらの資料のうち、必要な資料については、より詳細な読解を加え、行論に必要なものを抽出する。

#### (5) a 執筆

(2) と (4) を中心に論文の執筆を行う。そのうえで、自らの思考と歴史像を適切に表現することで一つの論文として完成する。なお、厳密にはこの完成のプロセスの際に査読等による修正を経ることとなる。

では、これらの作業のうち、どこに情報技術が入る余地があるかをここで確かめておきたい。

#### (2) b 先行研究の網羅

先行研究を探索する際には、論文データベースが大きな役割を示す。Cinii[1]による全般的な探索は言うまでもなく、個別の分野ごとの論文データベース[2]も有益である場合が多い。また、近年は史料から関連する論文を探索する試みが行われており、先行研究を発見する際に、これらを活用する時代が来たといえるであろう。

#### (3) b 関連資料の探索

関連資料の探索についても有益な部分があるのは言うまでもない。例えば木簡であれば木簡データベース[3]にあるものを探索のスタートに据えることは言うまでもない。また、東京大学史料編纂所のデータベース[4]は中世史に関連する編纂史料を発見するためには非常に有益なものであるといえるし、近代に関しても国立公文書館[5]・アジア歴史資料センター[6]・国立国会図書館などのデータベースについては、最初に確認すべきデータベースとして位置付けられているであろう。

#### (5) b 執筆

最終的な執筆に際しては、辞書類のデータベースなどが有益になる場合も多い。無論、史料の読解にも辞書を外すことはできないが、現時点において前近代の資料読解に役立つ漢和辞典などはほとんど入っていない。したがって、辞書類で有益なのは、Japan Knowledge に収録されている国史大辞典などが有益なツールとしてあげられるであろう。

上記のような過程の中において、情報技術が用いられていることがわかる。それでは、これらの情報技術の活用

において、何が可能となっており、何が達成されていないかを整理してみたい。

### 3. 情報技術の活用によって可能になったこと

本章では、2 章との関連で、情報技術の活用によって、可能になったことを整理する。

#### 3.1 基礎的なデータベースの活用

いうまでもなく、論文データベースである cinii の存在は大きい。また、東京大学史料編纂所をはじめとする、多くの史料所蔵機関がデータベースを公開したことによって、それらのデータベースの活用が可能となった。このことにより、研究者が探索にかかる時間が軽減されることになった。

#### 3.2 アクセス困難な資料へのアクセス

とりわけ、木簡データベースや正倉院文書データベースのような、日本古代史に関する資料の場合、そもそも現物にアクセスすることが難しい。すべての画像の公開という観点からは東寺百合文書へのアクセス[7]が容易になったことも大きな進展であるといえるであろう。

また、そのような誰もがアクセス困難であったような資料だけではなく、空間的に制約の大きかった、国立公文書館や国会図書館の資料へのアクセス改善も、重要な機能であるといえるであろう。

#### 3.3 可視化手法の高度化

GIS での可視化を代表例として、さまざまなデータの可視化がより容易なものとなった。とりわけ、博物館に關係する部分では、その機能の進展は大変に大きく、デジタル展示などで歴史へのアウトリーチは確実に大きくなったといえるであろう。

では、これらの可能になったことは、CH もしくは広くデジタルアーカイブの潮流でどのように位置づけられているのかを見ていきたい。

## 4. 日本史学に関する情報技術の変遷

### 4.1 2000 年前後の状況

これらのデータベースが、どのような段階で入ってくることになるのかを、簡単に概観しておきたい。まず、2000 年前後における日本史学関連のデジタル化については、基本的には CD-ROM での配布という形態をとっていた。平安時代の基礎資料である『平安遺文』のデジタル化や『群書類従』のデジタル化などが行われるとともに、辞書類についても『平安時代史辞典』などの CD が発売されたのがおおむねこの時期である。これらの CD-ROM の位置づけは検索できる「本」であり、基本的には、今までの書籍の中から情報をより簡単に発見できることが最大のメリットであった。これらの CD-ROM は書籍のコストをそのまま引き継いだようなものであり、大変に高価なものであった。また、ソフトウェアの変遷に弱く、OS などの大幅なアップデー

トにともない、現在は見られなくなってしまったものが多い。一方で、CD という物理的な形態のものによる配布という特徴があったため、販売側はデータの「管理」を行いやすいという実態があった。

#### 4.2 2000 年代前半のデジタルアーカイブの時代

2000 年代前後には合わせて第一次「デジタルアーカイブ」ブームがあったことも指摘しておく必要があるであろう。政府の IT 戦略とともに、多くの優品がデジタル化され、公開された、また、人文科学とコンピュータ研究会でもそれらに関連する研究報告がなされたこともある。じんもんこんのメインテーマが長く「デジタルアーカイブ」であったことも、触れておく必要があるであろう。この際に作られたデータの多くはデジタルギャラリーとしての「アーカイブ」でしかなく、画像が多数デジタル化されたが、結果的に残るものが何であったのかという点では、大きな課題を残すものであった（なお、じんもんこんのなかではデジタルアーカイブという際には分析に使うためのデータ蓄積という側面もあったという理解ではある）。データ形式の変遷に多くは対応できず、現在では廃墟になったような Web サイトも散見されるのが実情である。その中では、東京大学史料編纂所が Japan Memory Project[8]として、史料編纂所のデータベースの再整理を行い、多数のデータを入れていた点については、注目に値する。このときのデータベースの基礎が日本史学の基礎的なデータベース探索に大きな影響を与えているといえるであろう。

#### 4.3 2000 年代後半におけるデジタルアーカイブの鎮静化と DH

2000 年代も後半に入ると、政府による補助がなくなったことも影響してか。一挙にデジタルアーカイブブームは沈静化へと向かう。

それと並行して現れたのが Digital Humanities (DH) と GIS 技術の台頭であるという認識である。デジタルアーカイブブームの中でも蓄積されたデータ類の活用研究の活性化が起こっている。このころに起こったさまざまな解析系の成果は、現在の DH や地域情報学のような研究として位置付けられる一方、日本史学にはそのような動きが位置付けられることが多くはなかった。

#### 4.4 2010 年代の DH の系統の深化と Linked Data、Open Data へ

2010 年代に入り大きな傾向をもたらしたのはオープンデータと Linked Data であるといえよう。オープンデータ・オープンサイエンスの可能性は、文化資源を新たな形で活用可能にする方向へと進みつつある。それは人文学であっても同様であり、資料や成果をオープンにすることで機械が読めるようにすることも含めた、新たな活用段階に至ったといえるであろう。

この潮流を日本史学の中でいち早く取り入れたのは当時の京都府立総合資料館、現在の京都府立歴史館にある東

寺百合文書 Web である。東寺百合文書 Web は、いち早く CC BY 2.0(Ja)を宣言し、歴史資料のオープン化に一つの道を開いた。一方でこの Web サイトでは、さまざまな活用へとつながる仕組みがシステムとしては不足であるとともに、運用のしかけとしてオープンな事例の活用などの展開が行われなかったこともあり、その後、具体的に活用できているとはいいがたい。その点では、国文学研究資料館と CODH によるデータのオープン化とその後の活動は成功を収めており、展開が可能になっているといえるであろう。このような点でも、日本史学のデータが遅れを伴ってしまっている部分は否めない。

### 5. 30 年で変わったこととは

#### 5.1 2010 年の研究動向

本稿執筆の7年半前にあたる2010年1月に、『日本歴史』が「日本史研究とデータベース」という特集を組んだ[9]。そこでは、東京大学史料編纂所や歴博をはじめ、多くの歴史研究者が提供者の視点と利用者の視点の両方から、さまざまな議論を展開した。その特集は時代も地域も非常に幅広く、多数の展開が行われていた。

その目次を文末に参考表として記す。

その中には大変に示唆に富む言及があり、それを紹介してみたい。

例えば、岡本真はこの2010年段階で、データベースへのパーマリンクに必要性について述べている。「データごとにパーマリンクを持たないということは、そのデータにインターネット上でリンクを張れないことを意味する。リンクはインターネットの最大の特性であり、リンクできるかどうか、つまり情報のありかを指し示すことができるかどうかは、そのデータや情報のインターネットにおけるプレゼンスに直結するのだ」(同書 58p) この指摘は、図書館等では早くに受け入れられ、国会図書館が永続的識別子を付与するなどの動きがあった。また、この研究動向が、DOI の付与などの研究に結び付いていったことは言うまでもない。しかし、実際には日本史のデータベースでこのような識別子が付与できているデータベースはごくわずかしかない。

この例が端的に示すように、日本史のデータベースやそれをとりまく状況は2010年代以降、停滞の状況にあるのでは、という懸念をもつのである。

#### 5.2 2010 年代の日本史学関連のデータ化

実際、2010 年段階で執筆された本書において、照会されたデータベースから、実質的に「増えているのか」といことについては、疑問が残る。無論、史料編纂所をはじめ、多くの組織がリプレースをおこない、その機能を充実させてきたことは言うまでもない。しかし、2010 年代の新たな動向である DH のように研究課程そのもので、デジタルデータを活用するような事例は、いまだにごく少数であるといえる。GIS も可視化が中心であり、研究へと使ったもの

はいまだ少数派にとどまっている。

つまり、研究者の基本的な所作は変わらず、部分的な最適化が行われただけであるともいえる。このこと自体には両者の面があり、日本史学の研究手法そのものが、強固であるがために、デジタル化においてもその周辺の数値（辞書検索等）を上げることで対応できるとも言えるし、一方では必要な電子化なり前提条件なりの不足で研究手法の変革が起こらず、さまざまな分野から後れを取っているともいえる。

### 5.3 データ化の後退??

また、論文等の電子化に至っては、2010年の段階で岡本によると6学会誌が電子化されていると指摘されているが、決して状況は改善していない。機関リポジトリのデータのみが充実しているが、それは学界による自発的行為によらないことには厳しい目を向ける必要がある。

さらに言えば、テキストマイニング等に用いることのできる資料データについては、2010年段階では（ある意味グレーのまま）ダウンロードできる民間サイトが堂々と紹介されているのだが、一方で研究機関がこのようなデータの公開を行いにくなっている状況があり、停滞どころか後退に近いのではないかと感じる向きもある。

結果的にグレーなデータが残り続けることは、解析に用いるデータがいつまでたっても流通しないということでもあり、根本的な改善が望まれる。

『日本歴史』前掲書の中で、横山氏は下記のように述べて危機感を表明している「コンピュータがネットワーク化することは社会的現実だから、その現実を踏まえれば史料情報のネット利用という可能性は存在する。可能性が現実に変化するためには「歴史学に関する学術情報をどう共有するか」という問題を歴史学研究者として考え」（前掲書7p）るべきであると。そのうえで「日本史研究がその（ネットワーク）社会との関係を形づくっていかうとしているのか、いろいろ試みてはいるものの、私にはなかなか計りかねている」と。実際データベースの活用においても「すべてを見る」という部分に関しての「信仰」は取り切れているといい難い。そのようなある意味では因習に近い所作とどのように向き合うかも課題であろう。

## 6. 「次」の日本史研究の可能性

では、このまま日本史学については、何らの進展もないのであろうか。無論データへの見方など根本的な部分でも改善を求めるともいえる。一方で「みんなで翻刻」[10]のような歴史資料そのものの理解や位置への変更を促すような事例が生まれてきているのも確かである。歴博としても、「総合資料学の創成」の中で、機械可読可能なデータやさまざまなひらかれたデータ作成などの試みがある。

今後の日本史研究にとって重要なことは、いかに外に開いた、他者に開いたデータを作り、それをさまざまな形で

研究へと活用することであろうと考えている。そのための日本史学には情報学に加えた新たな学際性が求められているともいえるのではないだろうか。

**謝辞** 本研究の成果は人間文化研究機構機関研究プロジェクト「総合資料学の創成」およびJSPS科研費17H00773の助成を受けたものである。

## 参考文献

- [1] <http://ci.nii.ac.jp/>
- [2] 正倉院文書マルチ支援データベースの研究文献検索や歴博の荘園文献関係目録 ([https://www.rekihaku.ac.jp/up/cgi/login.pl?p=param/ronb/db\\_param](https://www.rekihaku.ac.jp/up/cgi/login.pl?p=param/ronb/db_param)) など
- [3] <https://www.nabunken.go.jp/Open/mokkan/mokkan.html>
- [4] <http://wwwap.hi.u-tokyo.ac.jp/ships/shipscontroller>
- [5] <https://www.digital.archives.go.jp/>
- [6] <https://www.jacar.go.jp/>
- [7] <http://hyakugo.kyoto.jp/>
- [8] <http://www.hi.u-tokyo.ac.jp/personal/yokoyama/jpmem/index.htm>
- [9] 『日本歴史』740号、2010年1月、吉川弘文館
- [10] <https://honkoku.org>

参考表

新年特集 日本史研究とデータベース	
日本史研究データベースはどこへ行こうとしているのか	横山 伊徳.
日本古代史研究のためのオンライン・データベース	小口 雅史.
分野別現況	
中世史研究資源としてのウェブデータベース	田良島 哲.
日本近世史研究とデータベース	鶴飼 政志.
インターネットで歴史研究	櫻井 良樹.
中国・台湾史に関するデータベース	江川 式部.
朝鮮前近代史に関するデータベース	長森 美信.
朝鮮近現代史に関するデータベース	河 かおる.
ぞんざいな検索、丁寧な検索--日本文学関連データベースの周辺	荒木 浩.
日本史研究におけるインターネットの学術利用--これまでの成果と、これからの課題	岡本 真.
提供者の立場から	
国立歴史民俗博物館のデータベースと統合検索	安達 文夫.
国文学研究資料館(アーカイブズ関係)のデータベースの紹介	五島 敏芳.
国文学研究資料館(文学系)データベース	入口 敦志.
国立公文書館・アジア歴史資料センターのデータベース	平野 宗明.
国立国会図書館・憲政資料室	堀内 寛雄.
東京大学史料編纂所データベース	遠藤 基郎.
東京都公文書館におけるデータベース利用	中元 幸二.
法政大学大原社会問題研究所のデータベース	若杉 隆志.
墨書土器データベース	吉村 武彦.
早稲田大学文学学術院学術情報データベース	海老澤 衷.
京都大学附属図書館「貴重資料画像」	西村 暁子.
大阪市立図書館商用データベース紹介	岡本 泰子.

奈良文化財研究所のデータベース--木簡を中心に	渡辺 晃宏.
奈良県立図書館地域史料目録のデータベース	大宮 守友.
九州大学デジタル・アーカイブ	宮崎 克則.
利用者の立場から	
日本史学におけるデジタル・アーカイブの現状と展望	後藤 真.
日本思想史系データベースの利用とその現状	桐原 健真.
四国における歴史教育・研究とデータベース--デジタルアーカイブとeラーニングの取り組みを中心に	鈴木 正信.
琉球史に関するデータベース紹介	渡辺 美季.
古代史研究とデータベース	加藤 友康.
日本古代史関係のデータベース	倉本 一宏.
古代史データベースと初歩的な使用方法	服部 一隆.
韓国木簡のデータベース	橋本 繁.
研究環境の最適化	野村 朋弘.
学会・研究会等データベースの現状	川戸 貴史.
日本禅宗史に関するデータベース	川本 慎自.
藩政史研究におけるデータベースの活用	野口 朋隆.
藩政史・地方史に関するデータベース	脇野 博.
学術空間としてのインターネット	梶田 明宏.
近代法制史に関するデータベース	天野 嘉子.
日本外交関係資料のデータベース	熊本 史雄.
大正・昭和の政治史に関するデータベース	清水 唯一朗.
地方公文書館等のデータベースの現状と課題	松沢 裕作.
国立国会図書館近代デジタルライブラリーおよび新聞・雑誌に関するデータベース	真辺 将之.
日本経済史に関するデータベース	宮地 英敏.

1 <http://ci.nii.ac.jp/>  
2 正倉院文書マルチ支援データベースの研究文献検索や歴博の荘園文献関係目録 (<https://www.rekihaku.ac.jp/up->

[cgi/login.pl?p=param/ronb/db\\_param](cgi/login.pl?p=param/ronb/db_param)) など  
3 <https://www.nabunken.go.jp/Open/mokkan/mokkan.html>  
4 <http://www.wap.hi.u-tokyo.ac.jp/ships/shipscontroller>