

解 説

人文・芸術系のデータベース—今そしてこれから—

2. 歴史データベース—日本史を中心に—

History Researches and Database by Takehiko TERUI (National Museum of Japanese History).

照 井 武 彦¹¹ 国立歴史民俗博物館

1. はじめに

歴史研究を支援する情報処理の発達と、それにひとりでも多くの人々に参加していただくことを、日頃願っている者である。現在、それにたずさわっている人口はたいへん少ない。このような仕事をはじめたころは、「面白そうな仕事ですね」といってくれる人も多く、すぐにでも人口が増えそうな気がしていたのだが、しばらくたってみて、その予想は全然あたっていないことに気がついた。その理由はさまざま考えられるが、ここでは触れる余裕はない。この解説がその打開に少しでも役立つことを願い、これまでに体験したことを交えて述べてみたい。

さて、我が国の歴史研究・教育は長い間、日本史、東洋史、西洋史の三者が並立し、こんにちでもその骨格は変わらない。そのうち、筆者はほぼ100%日本史にかかわってきたので、ほかへあまり言及できないことを理解しておいていただきたい。

日本史では、こんにちでこそ、その自由な研究環境が確保されているといえるが、それは第2次大戦後のことで、それ以前は皇室にかかわることをはじめとして研究対象にできない事柄が数多くあった¹⁾。戦後においても、たえず、さまざまな政治問題、国際問題が、とくに歴史教育の面に、大きなかげをなげかけていることは周知のところである。研究者はそのような研究環境が損なわれないように不断の努力を続けなければならない。その結果、ほかの分野の者にはだれであっても警戒を怠らない、との態度にできるものやむを得ないことであろう。支援を惜しまないつもりだから、われわれの分野だけ例外にしてくれ、との態度は通じない。初期に体験したのは、こんなことだったのではなからうか。

この解説では、歴史データベースの源である文献資料と「もの」資料について述べ、そのあと、データベースの現状と今後に関する話題を幅広く紹介するつもりである。

2. 文献資料、史料

歴史研究の対象物はいうまでもなく文字資料であり、それを文献資料とか史料と呼ぶ。まず、日本書紀、続日本紀(しよくにほんぎ)などの六国史(りっこくし)、吾妻鏡(あづまかがみ)、徳川実紀など編纂された史書がある。古く成立したもののほど原本は失われたり、細部の異なる筆写本(異本)が存在する。史料批判とは、異本同士や同時代のほかの史料とつきあわせ、はたして歴史的事実をあらわしているか、の論であり、史書はとくに、時の権力の手で編纂されたことによる疑いの目にさらされる。

古代の法律である律令(りつりょう)(大宝令など)、中世の御成敗式目(ごせいはいしきもく)なども重要である。

古文書(こもんじょ)は、差し出し名、宛て名、日付などを記し、特定の相手に自己の用件や意志を伝えるためのもので、単に「もんじょ」ともいう。古記録は、公家などの忘備録・日記であり、もともと朝廷の行事・儀式を子孫に伝えるためのものであったが、そのときどきの世のできごとが記されている。

考古学上の出土資料(金石文、木簡、漆紙文書など文字が記されている)も重要な史料である。

ところで、ここまでは古代・中世の、つまり戦国時代までのことを多く述べる結果となったが、それらの時代から残されているものは量的にきわめて少ない。その大部分は上層社会の記録であって、民衆の日常生活・文化を読み取ることができ

ない。最近さかんになった歴史考古学は考古資料を文献資料と組み合わせようとするもので、古代木簡、漆紙文書などの研究がその例である。

ところが、江戸時代から後になると様相は一変して大量に存在するようになる。マイクロフィルムは記録、活用に重用されている。江戸時代の古文書は地方にも多く保存されているので、史料探訪が行われ、その結果は東京大学史料編纂所、国文学研究資料館・史料館などに全国的に集積されている。大規模なマイクロ化もよく行われる(たとえば大名家文書)。明治以降では、とくに官公庁の文書が、重要な情報が含まれているものまでが広範囲に廃棄されたという²⁾。残念なことである。

3. もの資料

我が国の博物館数は、現在3000館を少し超えたぐらいであるが、そのうち歴史博物館は、歴史分野をもつ総合博物館を含めると、6割以上になる。博物館が収蔵・展示するものの多くは(文字資料に対する)「もの」資料である。もの資料は博物館以外に、古いものは社寺、近代以降は企業・団体が所蔵しているものも多く、企業自身が博物館を設立する例も多い。

ところで、どこの歴史博物館であっても、必ずある広がり土地を対象として成り立っている。その広がりさまざまである。それぞれの地域は独自の文化を伝えてきた。各館が所蔵する資料は、その地域の独自性を示す証拠である。一般に、もの資料はその土地から離れば価値が半減するといわれる。「現地保存主義」はここから生まれた主張である。歴史に限っては歴博のような中央博物館が不利なのである。それを解決するものとしてレプリカ(模型、模造)の利用が考えられた。

こうして始まったレプリカの多用は、歴史展示の近年の傾向として全国的にひろまった。テーマによる系統的展示には、所蔵品や借用品の実物だけでは十分でないことが、もう1つの理由である。こんにち模型の技術はたいへん進歩しており、実物とまったく変わらない外観を実現することができる。ただ、模型の製作には実物の価格以上に費用のかかることもあり、また、大型模型では展示後の収納場所に困るなどの問題も起こる。

4. 歴史データベースの現状

まず、オンライン公開サービスされているものをあげてみよう。以下はすべて「何々データベース」と称するが、「データベース」は省略する。文部省学術情報センターでは、「維新綱要(こうよう)」(東京大学史料編纂所提供)、「古文書目録」(同)、「木簡」(奈良国立文化財研究所提供)、京都大学大型計算機センターでは、「室町幕府奉行人奉書(ぶぎょうにんほうしょ)」、「続日本紀」、「令集解(りょうのしゅうげ)」、国際日本文化研究センターでは、「古写真」(画像・書誌)、「外像」(同)(外国文献中の日本画像)、国立歴史民俗博物館では、「旧高田領取調帳(きゅうだかきゅうりょうとりしらべちょう)」、「東大寺文書目録」、「日本荘園」を公開している。博物館資料目録を公開しているのは、国立民族学博物館、東京国立博物館(資料写真)、歴博などがある。また、図書館が貴重図書として史料を所蔵する例も多く、図書目録のオンライン公開は進んでいるようである。

公開にはいたらないが、機関として入力などを進めているデータも機関も数多くある。以下そのようなものの例をあげる。国文研で入力された「日本古典文学大系」には、日本書紀などの史書が含まれている。日文研では、「御触書(おふれがき)」など。歴博では、左経記(さけいき)、吾妻鏡など5件を「記録類全文データベース」として館内サービスしている。

表-1 旧高田領取調帳の年次経過

年度	事項
1983	入力(関東、中部)
1984	同(近畿)
1985	同(中国・四国、九州、東北)
1986	校正(第1次)
1987	同
1988	校正(第2次)
1989	同、データベース搭載準備
1990	オンライン公開(4月)
1991	(企画展示に出品・パソコンデータベース)
1992	(同 5月まで)
1993	増強(地名読み、市町村コード追加)
1994	同
1995	第2版公開

1つのデータベースの作成に着手してから公開にいたるまで、早くても数年を要する。その一例を表-1に示す。

機関としてでなく、研究者個人として自分だけのためにつくるデータベースは、それらの対極にあると考えられる。この「個人データベース」は、やがてほかとの共同利用データベースに成長していくものもあり、はじめから数人の研究グループを組んで推進をはかるものもあるが、知られずに終わってしまうものの方がはるかに多いようである。それは、個人の都合・方法でつくったためほかの利用に耐えるようにするには、結局作り直しになるのだという。技術的にだけみれば理解し難い(データ設計がわるいと決めつけるのは簡単である)ものの、考えさせられることの多い問題である。

次に、資料と情報の系列を考えよう。また旧高田領取調帳を引き合いに出すが、その資料と情報は表-2のように並べることができる。

これらは上ほど情報量が多く、下ほど利用の便がよいように並んでいる。系列の要素が豊富に揃っていれば利用者は下からアクセスしてみて、途中で目的を達する場合が多いだろうから、貴重な実物資料などの保存にも都合がよい。

これまで文献資料の場合を述べたが、もの資料でも同様である。その上、もの資料ではその情報を使いやすく引き出すのが不便ことが多い。それだけに、このような系列は充実されなければならない。それが立体物の場合、表-3のようになるであろう。

このような資料の情報をつくることのできるものは、ものの種類によって限られた人々であるから充実といってもそう簡単に進むものではない。人文系の場合、内容は書く人によって左右されるのが普通で、客観性には自ずと限界がある。ここでは、むしろ作者を明記すべきだといわれている。

歴史データベースの推進、情報交換の場として定期的な研究集會が開かれるようになった。我が情報処理学会に「人文科学とコンピュータ研究会」があるが、そのほかにも

京都大学大型計算機センター研究セミナー

東洋学へのコンピュータ利用

国文学研究資料館シンポジウム

コンピュータ国文学

表-2 資料と情報の系列事例—旧高田領取調帳の場合

1. 実物	関東大震災で焼失
2. 筆写本	明治大学図書館蔵
3. 筆写本	東京大学図書館蔵
4. マイクロフィルム	明治大学
5. 活字本	木村健校訂, 近藤出版社
6. データベース	歴博オンラインシステム

表-3 もの資料に対する系列

1.	実物
2.	複製品(実物が失われている)
3.	複製品(実物は現存する)
4.	写真(4×5などの大型ネガ)
5.	写真(35mmネガ)
6.	資料調査カード
7.	調査報告書
8.	図録
9.	館蔵資料目録(多項目, 館内用)
10.	館蔵資料目録(数項目, 公開用)

統計数理研究所合同研究会

文献情報のデータベースとその利用

などがあり、文学に次いで歴史の報告が多い。

5. 今後の課題

最大の問題は、歴史データベースの作成、公開(共同利用)が遅々として進まない(少なくとも情報技術の立場からはそうみえる)ことであろう。それに関することを2, 3あげてみよう。

短時間に大規模なデータベースをつくるには共同作業が不可欠であるが、歴史研究はきわめて個人的色彩が強く、人数を増やしても分担が円滑にいかない結果になりがちである。ひとりで時間をかけて、となる。それに加えてデータの内容的な信頼度が問題にされる。文献資料を活字本にするときは、表-2にあったように校訂を行う。歴史研究者はデータベースにも同じように臨む。実際、表-1のデータベースにも事実上の校訂者がいる。4章の末尾近くに書いた「作者明記」の考え方もこれと符合するといえる。このような、印刷物とデータベースの(過度の)アナロジーはなんとかならないものかと思ふ。

通信ネットワークの利用が、既成データベースの検索のためにあるとの認識は、歴史研究者にもかなり以前からあったが、最近、ようやくそれを



図-1 江戸図屏風—日本橋近辺
国立歴史民俗博物館蔵

超える動きが出てきた。パソコン出現のころ高校生・大学生で、それに親しんだ人が、ごく少数ながら歴史研究者になっていて、現在、その中心をなしているようである。具体的には、参考文献3)のような活動である。パソコン通信 NIFTY Serve の PATIO を利用した、「資料情報の電子化と標準化に関する会議室(主催者:田良島哲)」をあげることができる。これは、特定少数者の会議室であるが、NIFTY では大規模な「歴史フォーラム」が運営されていて、ここはほかのパソコン通信に比較して人文学研究者やその卵の割合が高いように見受けられる。

ここに述べたような人々はおそらく独学で情報処理をマスタしたのであろう。大学の史学科などでの組織的な教育はこれからである。参考文献4)は、歴史研究者が著した初のパソコン入門書であり、特定の機種ではあるが、史料からデータベースをつくる方法が詳細に説明されている。この本をみると、ほとんどの説明が文章でなされ図がないなど、歴史と情報処理でプレゼンテーションがこうもちがうのかと驚かされる。

日本史では、国文学、東洋学とならんで漢字コードが大きな問題である。JIS 第1・第2水準は周知のようにパソコンの初期から便利に使われているが、この世界では、いわゆる外字が頻発する。近年 JIS では補助漢字が追加されたが、期待したほど普及していないようである。改めて第3、第4水準を制定する動きがあり、準備の一環として第1、第2水準のときの字種出典の調査が行われた⁵⁾。今後制定する際にも、出典の明示を前提とするそうなので、現用歴史データベースの外字な

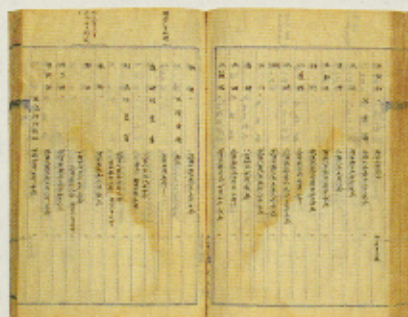


図-2 旧高旧領取調帳—下総国佐倉近辺
明治大学図書館蔵
国立歴史民俗博物館写真提供

どは有力候補になりうる。しかし、個々のデータベースの外字や異体字の取扱いは、その作者により漢字親が異なるため、字を置き換えてしまったり、かな書きにしたものもあり、調査は不可能ではないが、楽な仕事ではない。国際規格の Unicode が普及する勢いであるが、歴史関係者の間では Unification などに対して批判が強い。

なお、国文学研究資料館をはじめ、2、3の機関で外字(電子字形)の製作が継続的に行われており、その活用が期待される。また、漢字の内部構造に着目した研究が、地味ながら進行している。偏(へん)と旁(つくり)に分解するような手法を使う。その代表的な労作を参考文献6)にあげる。

これまで、文字データの話が多くてだが、文献資料、もの資料を問わず、マルチメディアの利用により格段に利便性が向上することは、容易に理解されよう。文献資料マイクロフィルムのデジタル化・オンライン転送が試行されている⁷⁾。直接デジタル化も始まっている(京都大学図書館など)。もの資料を収集・展示する博物館は、さらにマルチメディアに向いている。東京国立博物館では、古くはガラス乾板などの写真十数万枚を所蔵し、以前から 30cm 光ディスク利用を研究している。歴博では、PhotoCD(ProPhotoCD の割合が大きい予想)を標準媒体とする方針にした。このように高精細な静止画像の蓄積が中心的課題である。他方、民俗学では民具製作とか祭礼のような無形文化財の記録映像を収集・製作している。また民謡の音声記録が大規模に行われている。

これらは実演が廃れてしまう運命にあるから収録を急がなければならない。

問題なのは、デジタル媒体の記録寿命が不明であり、媒体の種類が短期間に移り変わることである。和紙は千年以上、マイクロフィルムは百年といわれるが、コンパクトディスクなどはどんなものだろうか。寿命研究の急速な進展を望む。

6. おわりに

歴史の話ではどうしても難読語が多くなってしまふ。できるだけよみがなをつけてみたが、全体として難しい雰囲気の説明になってしまった。すこしでもこれを補う意味で、付録に簡単な説明をつけたので参照していただきたい。

この解説が歴史研究と情報処理を少しでも近づけることになれば幸いである。

参 考 文 献

- 1) 井上光貞：私の古代史学，318p.，文藝春秋社（1982）。
- 2) 安藤正人・青山英幸：記録資料の管理と文書館，585p.，北大図書刊行会（1996）。
- 3) 多良島哲：史料情報の電子化とオンライン流通に関する諸問題，古文書研究，Vol.40，pp.57-66（1995）。
- 4) 中野栄夫：コンピュータ歴史学のすすめ，209p.，名著出版（1994）。
- 5) 家辺勝文（司会）：国境をこえる日本語の条件，GLOCOM NEWSLETTER，Vol.5，No.1，pp.2-13，国際大学グローバルコミュニケーションセンター（1996）。
- 6) 勝村，丹羽：漢字典，京都漢字研究会（1994）。
- 7) 並木，柴山：WWWによるマイクロフィルム画像検索システムー構想と初版の開発について，京大計センター研究セミナー報告，Vol.52，pp.25-42（1996）。

- 8) 照井武彦：歴史データの作成経験ー日本の歴史考古民俗の場合ー信学技報，DE93-59，pp.33-40（1994）。

（平成9年3月3日受付）



照井 武彦（正会員）

1933年生。1956年東北大学工学部通信工学科卒業。同年より（株）北辰電機，日本電気（株）勤務。1966年より岩手大学工学部電子工学科，情報工学科勤務。1980年文化庁，1981年より国立歴史民俗博物館勤務，現在，同館情報資料研究部教授。日本の歴史・考古・民俗学を中心に人文学の情報処理研究に従事。電子情報通信学会，情報通信学会，情報考古学会各会員。

付 録

解説本文のなかで，なじみのないと思われる歴史上の名称のいくつかを簡単に説明する。

六国史：古代の史書は，日本書紀，続日本紀，日本後紀，続日本後紀，日本文徳天皇実録，日本三代実録の6書からなるので，この名がある。

律令：古代の法律で，大化の改新直後の近江令にはじまる。つづいて浄御原令，大宝令，養老律令ができた。

令集解：律令の注釈は古代のうちにはじまるが，令集解は平安時代初期に集成された注釈書である。

吾妻鏡：鎌倉時代，幕府が編纂した正史。

御成敗式目：鎌倉時代の法令。

徳川実紀：江戸時代，幕府が編纂した正史。

御触書：江戸幕府の法令。

旧高旧領取調帳：明治時代初期に，幕末の旧領主，石高を村単位に調査した記録。

訂 正

本誌第 38 卷 5 号 (1997) pp.383 ~ 387 に掲載されました「歴史データベース—日本史を中心に—」の著者照井武彦氏の申し出により、pp.387 の参考文献 3) を次のとおり訂正いたします。

(誤) 3) 多良島哲：史料情報の電子化とオンライン流通に関する諸問題，古文书研究，Vol.40, pp.57-66(1995)。

(正) 3) 田良島哲：史料情報の電子化とオンライン流通に関する諸問題，古文书研究，Vol.40, pp.57-66(1995)。