

「沖縄の歴史情報」情報化資料の概要

領域代表者 岩崎 宏之

1. 歴史研究でのコンピュータの利用

歴史研究に情報科学を援用する利点としては、

- (1)大量のデータを、迅速に処理することが可能である
- (2)網羅的かつ精密な検索・解析が可能である
- (3)文章、語彙、文字についての厳密な異同の確認が可能である
- (4)画像・映像の処理が容易である
- (5)コンピュータを介在させて遠隔地との通信が可能になった

ことなどを列挙することができる。しかし網羅的な処理、あるいは大量データの迅速な処理が可能となるのは、総てのデータが機械可読型の情報としてコンピュータの中に存在することが前提となる。コンピュータはわれわれの指示に従って答えを返してくるだけであって、入っていないデータは出てこないからである。コンピュータと付き合ううえで大切なことは、コンピュータと対話するためのプログラム言語の獲得と、大量かつ多様なデータをいかにして機械可読型のデータとして情報化できるか、の問題である。そこでは歴史資料を情報化するうえでの各種のソフトウェアの開発が必要であることは言うまでもないが、同時に、情報化すべき歴史資料そのものについての研究、すなわち電子化に対応した史料の構造や機能についての検討、原史料の画像のイメージ情報の持つ階層的な構造、史料そのものの文脈や文章の構造の解析、さらには一つひとつの語彙・文字にいたるまでのさまざまな次元での問題に対処しなければならない。いわば歴史資料の情報化のためには、これまでの歴史研究では経験しなかった史料研究の新たな領域があり、そのための歴史情報資源研究が不可欠なのである。われわれはなによりも「コンテンツ」を重視し、歴史研究に役立つ歴史情報の集積につとめた。それには多様な方向からの研究が共同して行なわれることが求められ、またその共同研究は、研究のプロセスの公開を含めた公開性が要求されると考えている。大量の史料の情報化には共同作業による能率向上が効果的であり、その成果は学界に広く公開されることが期待されている。「情報化」の必然的な帰結として、共同研究の取組み方と情報化された資料の公開と共有に関する新しい原則の確立が求められているといってもよい。私たちのこの重点領域研究「沖縄の歴史情報研究」は、このような観点にたったの実験的な試みなのである。

2. 「沖縄の歴史情報」データベース

さて、琉球・沖縄史と環東シナ海世界の地域間交流史に関する多種多様な歴史資料を、いかにして機械可読のデータとして情報化することができるか、これを本領域研究では、

- (1)各種研究文献の統合的把握のための研究文献情報データベースの作成

- (2) 各種歴史資料の所在情報に関するデータベースの作成
- (3) 琉球王朝期の外交文書集「歴代宝案」や琉球家譜など、琉球・沖縄史研究にとっての基本的文献の全文テキスト・データベースとその検索システムの作成
- (4) 琉球・沖縄史と環東シナ海世界の地域間交流史に関する基本的史料をイメージ情報として利用することが出来る画像データベース「琉球史料集成」の構築（マイクロフィルム、CD-ROM などを使用）とネットワークを介しての公開・提供

に集約して研究開発と情報化作業を進めてきた。

3．各種研究文献の統合的把握のための研究文献情報データベース

琉球・沖縄史に関する過去の書誌類・研究文献情報を総合する試みである。これまでに生み出された琉球・沖縄に関する歴史研究の論文はいったいどれほどの数になるものであろうか。たとえば沖縄県立図書館史料編集室編集『史料編集室紀要』巻末に毎号掲載されている「沖縄歴史関係研究文献目録」・「沖縄歴史関係主要論文目録」には、第20号(1995年3月刊行、1994年分まで)までの20年分を合わせると約2万タイトルあまりが掲載されている。この目録は単行本・論文(雑誌・紀要等)の他に新聞等に掲載された歴史関係評論記事までが収められた詳細なもので、いわば最近20余年の琉球・沖縄に関する歴史研究文献の総体と考えてもよい。沖縄の歴史研究が本格的に展開した最近20余年間の研究成果を約2万タイトルと考えれば、これを明治期にまで遡ったとしても、この倍にはならないと考えてよいであろう。本領域研究ではまず、この琉球・沖縄に関する歴史研究の総体を把握するための研究文献情報データベースの作成を試みた。しかし、新規に研究論文等のすべてを確認することは困難であるため、本領域研究ではこれまでに作成・刊行された各種の研究文献目録類に依拠し、これら公刊された文献目録類を集成することによって過去における琉球・沖縄の歴史研究に関する研究文献目録の全体を把握しようと試みた。入力した研究文献情報は6万レコードを超えるが、これはそれぞれの目録で重複する分を含むため、単一のファイルに統合して全体を検索するシステムの構築を意図しているが、残念ながら時間的な制約から研究期間の終了までに完成できなかった。難点の一つはそれぞれの目録ごとに収載している情報のレベルが異なり、これを最も少ない情報のレベルで統一したのでは切り捨ててしまわなければならない情報があまりにも多いことである。単一化を図るよりは、むしろ個々の目録の特性を活かして個別に必要なデータを検索することのほうが歴史研究者の使い勝手に適していると考えられないこともない。そのため、まず素材として統合する以前のものをそのまま提供することにし、できるだけ利用者が各自の使い勝手に応じて加工しやすいような形で提供することにした。

4．各種歴史資料の所在情報に関するデータベース

琉球・沖縄に関する歴史資料が、どこに、どのようなものがあるかの史料所在情報データベースである。琉球・沖縄に関する史料の所在情報としては、各種資料所蔵機関等が収蔵する文献目録や収集資料目録などがある。史料の存在に制約されることが多い歴史研究にあっては、史料の所在情報を把握することが重要な要素となっている。われわれは史料の所在に関する各種の情報を研究者それぞれの努力の結果として経験的に蓄積してきているが、これはなかなか公開されにくい性格のものである。文献情報データベースが期待される所以でもある。この点で沖縄の場合は、新城安善氏が編纂された

『沖縄書誌総覧 - 沖縄書誌の書誌 - 』（ひるぎ社、1991年1月）の存在がきわめて有効である。本書には前項の研究論文等の文献目録情報を含めた各種の書誌情報が取りまとめられているが、新城氏のご好意でこれを電算化することができた。

本領域研究の一環としておこなわれた史料調査・収集によって作成された目録類がある。東京大学史料編纂所では本領域研究の一環として同所所蔵の島津家文書・島津家本の写真撮影と新規の目録作成が行なわれた。また研究分担者であった中村質氏（九州大学名誉教授、平成10年3月1日に急逝）のご尽力で、長崎県立図書館所蔵環シナ海域交流関係史料や平戸松浦史料博物館所蔵「新出奉書・御内書目録」が作成されたことも本領域研究の成果として注目されることである。また既存の各種史料群のなかから必要な情報を抽出し、特定主題によるデータベースとして統合する作業も進められた。

5. 琉球・沖縄史研究に関する基本的文献の検索システム

本領域研究では、「歴代宝案」や清代中琉関係档案統編、琉球家譜、「中山世譜」、仲吉朝助編「琉球産業制度資料」、島津家本「琉球外国関係文書」、オランダ海軍省琉球関係文書、「大島筆記」等々、琉球・沖縄史研究に関する基本的文献の全文テキスト・データベースとその検索システムを作成した。沖縄県立図書館歴代宝案編集室は校訂本『歴代宝案』刊行がすすめられているが、本領域研究はその既刊分について校訂本の使用と編集原稿ファイルの提供を受け、これをもとに「歴代宝案検索システム」を作成した。また台湾大学本全巻6千数百ページのページイメージのCD-ROMを星野聰氏が作成するなど、歴代宝案関係資料の情報化が進展した。

6. 画像データベース「琉球史料集成」の構築

文献研究にとって必要なことのひとつは文献資料本来の姿を知ることである。活字化される以前の、本来の手書き文字の原資料のイメージ情報を伝えるための画像データベースの構築を本領域研究の大きな課題として設定してきたが、最近の画像処理関係情報機器の普及とネットワーク通信環境の整備によって具体化してきた。

沖縄の場合は、失われた歴史史料の回復のために、マイクロフィルムによる史料収集が行なわれてきた事情があり、沖縄県内に膨大な量のマイクロフィルムが蓄積されてきた。それらは個人の研究活動によって収集・集積されたものだけではなく、県史編纂事業や市町村レベルでの自治体による誌・史類の編纂事業を通じて集められたものも多い。これらのマイクロフィルムの多くは昭和40～50年代に撮影されたものがあり、困難な条件のもとで調査・収集されたことを反映して画質はかならずしも良いとは言えないものがあるが、今日ではかつての所蔵者を離れたもの、あるいは原本がすでに失われたものも少なくない。貴重なフィルムであるとともに戦後の沖縄研究の史料的基礎を示すものとして重要な意義を有している。今回の領域研究では各種マイクロフィルムの所在や撮影経緯の調査を行ない、研究分担者我部政男氏から多数のマイクロフィルムの提供をうけるなど収集に務めた。また沖縄県立博物館、八重山博物館、筑波大学附属図書館、建仁寺、長崎県立図書館などの所蔵する琉球・沖縄史関係資料、環シナ海域交流関係資料の新規写真撮影を行ない、多数の資料を収集することができた。新規の写真撮影は主として16mmマイクロフィルムによって行ない、また収集した35mmの既存マイクロフィルムも同様に16mmマイクロフィルムに変換した。

本領域研究では琉球・沖縄史と環東シナ海世界の地域間交流史に関する基本的史料をイメージ情報として利用することが出来る画像データベース「琉球史料集成」の構築を目指している。これは富士写真フィルム(株)のイメージ・ネットワークシステム FDIP6200 を使用して琉球・沖縄史に関する基本的文献史料をイメージ情報として提供しようとするもので、基本的にはマイクロフィルムによって収集された資料を画像情報として電子化してネットワークで公開する。マイクロフィルムを素材とするのは、将来に向けての史料保存にとって、目下のところマイクロフィルムによる方式が最適と考えられるからである。FDIP6200 によるイメージ・ネットワークシステムは、オートストッカーを併用することで1ユニット200本の16mmマイクロフィルム・カートリッジを制御することが可能である。ここに収められるマイクロフィルムはおおよそ40万齣となり、これを検索するためのインデックス・データベースを作成した。

7. 外字の処理について

多様な言語・文字の資料をひとつのテーブルで検討しうるための情報化の試みがある前提となる。また漢字文化圏を対象とする東洋学研究へのコンピュータ利用にとって大きな制約となっているのが漢字処理をめぐる問題であるが、多様な、また時代的に変化を重ねている言語資料を含む歴史資料の情報化は、そのための試みの場としてまことに恰好な領域といえよう。

本領域研究における研究成果の取りまとめにとって、外字の処理は最初から大きな課題となっていた。歴史資料の情報化が、JIS 第1水準・第2水準という条件によって制約されている状態からどのようにして解放されるか、また JISX0221(ユニコード)制定をめぐる問題が決着を見るまでの経過措置をどのようにすべきか、これまで種々の論議を重ねてきたが、与えられた本領域研究の研究期間の間に抜本的な解決が見いだせない以上、将来への布石として取り敢えずの措置を構ぜざるを得ない。詳しくは本報告の第4部で述べるが、例えばその方法として以下のような対応を考えた。

(1) 琉球・沖縄史関係文献資料にはたしてどれほどの外字を必要とするか。計画研究金城班の赤嶺守氏が「歴代宝案」「清代中琉関係档案続編」情報化の作業を通じてこれまでに確認した外字は1433字であるが、この外字を1バイトの数字と記号との組み合わせたもので置き換えたものが「赤嶺コード」である。文献史料本文の全文テキストデータベースの入力には、当面この赤嶺コード(「外字・記号対象一覧表」参照)によって外字を入力しておき、将来において文字コードの問題が解決した段階で変換テーブルによって処理することにした。

(2) 全文テキストデータベースの場合と異なって、研究文献情報データベースや史料所在情報データベースの場合はデータ自体に1バイト文字を含み、(1)の方法をとることが出来ない。本領域研究ではこれらのデータベースの多くはデータベースソフト「桐」(管理工学研究所)で作成されている。ネットワーク上で公開・提供する際には「桐」形式とともにCSV形式トファイルに変換したものを合わせて提供できるようにし、利用者は各自の使いやすい方法で利用することを考えている。そのためデータベースのフィールドのひとつに予め外字コード番号を記入する欄を設けておき、当該の外字を■で表示してこの外字のコード番号を「外字コード」欄に注記する方法をとることにした。外字のコード番号としては勝村哲也・丹羽正之編『漢字典』(京都漢字研究会発行、1994年)など参照し、『漢字典』コード番号のほかに諸橋『大漢和辞典』コード、康熙字典コード(京都大学大型計算機センターが作成した康熙字典文字フォントのコード番号)等の対照表を作成するなどの工夫を試みた。

8 . 分散型情報ネットワーク

このようにして得られた本領域研究の膨大な情報化資料を、いかにして公開・共同利用に供し得るかの体制整備が緊急な課題となった。領域研究の成果であるデータベース等の電子化情報はCD-R、MOなどさまざまな方式での提供を行なうが、われわれは最近の情報ネットワーク、とりわけWWWサーバーを利用したインターネットへの情報発信が大きな可能性をもつ方法として考えており、旧来のピラミッド型、中央集権的ではない、情報発信者の特性と状況に応じた「分散型情報ネットワーク」編成の可能性を模索してきた。領域研究の成果の多くの部分は筑波大学附属図書館の電子図書館構想と連携して同図書館のサーバから発信することになるが、島津家文書などの公開はそれを所蔵する東京大学史料編纂所の日常活動のなかで公開されるなど、多様な方式が試みられる。