

人文科学におけるコンピュータ利用の現状と課題

及川 昭文

国立教育研究所

最近のコンピュータ関連の国際会議などで、コンピュータを利用した人文科学研究の報告が多くなされるようになってきている。日本においても博物館、美術館、図書館や大学などでこの種の研究が積極的に進められる傾向にあり、その重要性、必要性が広く認識されつつある。しかしながら、欧米のそれに比べて質、量ともまだ充分でなく、情報交換や研究会、あるいは学会活動などの積極的な推進が望まれている。人文科学におけるコンピュータ利用の現状と今後の課題について述べる。

COMPUTING IN THE HUMANITIES

Akifumi Oikawa

National Institute for Educational Research

Shimomeguro 6-5-22, Meguro-ku, Tokyo 153, Japan

In recent years, international conferences on computer applications into the humanities have been held so frequently that we face many papers treating this area of research. In Japan, the study of this area has been being continued steadily by some groups of museums, libraries, and universities, and some technical achievements have been shown to attract many people. The necessity and importance of this study has been strongly recognized. In this paper the present state of computing in humanities is described.

はじめに

最近のコンピュータ関連の国際会議などで、コンピュータを利用した人文科学研究の報告が多くなされるようになってきている。日本においても博物館、美術館、図書館や大学などでこの種の研究が積極的に進められる傾向にあり、その重要性、必要性が広く認識されつつある。しかしながら、欧米のそれに比べて質、量ともまだ充分でなく、情報交換や研究会、あるいは学会活動などの積極的な推進が望まれている。人文科学におけるコンピュータ利用の現状と今後の課題について述べる。

データベース

人文科学におけるコンピュータ利用は、データベースの構築およびその活用が中心となっている。最近テキスト情報のみでなく、イメージ情報のデータベース化の試みも盛んに行われるようになってきた。しかしながら、どのようなデータベースがあり、どのような利用が可能かといった情報を入手することははなはだ困難であり、人文科学分野のデータベース一覧などの作成が期待されている。

一般的にデータベースの構築には多くの時間と費用が必要であるが、1981年度より文部省科学研究費補助金（以下、科研費）による、データベース構築のための援助が開始されている。1988年度までに約170のデータベースについて科研費が交付されており、その内訳を見ると約3分の1が人文系のデータベースである。最近のものを以下に紹介する。

◇日本独文学会文献情報データベース（作成者：日本独文学会）

日本で発表されたドイツ語学、文学関係の文献のすべてを網羅することを目標としている。1988年度より作成を開始しており、初年度は日本独文学会の機関誌「ドイツ文学」に掲載されている寄贈文献目録に含まれている約1000件が入力されている。

◇地名データベース（作成者：日本地理学会）

日本の地名に関して、その位置座標、名称、読み、種別などがデータベース化されている。

◇連歌作品目録データベース（作成者：国文学研究資料館）

国文学研究資料館で収集した連歌資料をデータベース化したもので、構築されたデータベースを利用して、目録、各種索引（発句、人名、名称、事項など）などが作成されている。

◇人文・社会科学関係学術論文題目情報データベース（作成者：学会誌刊行センター）

いわゆる人文・社会諸科学の雑誌に掲載された文献の欧文による索引であるが、文学・法学、経済学などのほか、人類学・農業経済学・社会医学・情報学・家政学・科学技術史・体育学・地理学・科学教育などの分野も収録している。欧文の題目のないものでも、原則として原著者自身に欧文訳を依頼して収録するようにしている。

◇歴史天候データベース（作成者：歴史天候復元グループ及び山梨大学情報処理センター）

気象観測開始以前の日本の気候情報を収録したデータベースである。江戸時代の藩庁日記、社寺、家農、町会所などの日記に記載されている日々の天候記述を1次資料としている。1989年4月現在、28地点約2000年分のデータが入力されている。

◇インド学仏教学論文データベース（作成者：日本インド学仏教学会）

インド学仏教学関係の論文（日本出版のもの）の書誌データを収録している。現在までに2種の雑誌に収録されている約7000件のデータ入力完了している。本年6月より学会関係の30数大学で分担し、約17000件の入力を行う計画である。

◇日・中・韓法律用語データベース（作成者：九州大学）

日本、中国および韓国で用いられている主要な法律用語について、見出し、概念およびその意味内容をそれぞれの国の言語と共通語としての英語（あるいは独・仏語）表記が収録されたデータベースである。日本のある法律用語が同じ漢字圏の中国では、どのような表記で、あるいはどのような意味、概念上の異同を伴って用いられているかということを検索したりするのが目的である。

◇オノマトペ表現データベース（作成者：神戸大学オノマトペデータベース作成委員会）

日本語におけるオノマトペ表現を英語との対照研究に利用することを目的として作成されたデータベースである。各レコードは、まず擬音語、擬態語に大別され、それぞれ基本的意味、および種々の例文が含まれている。それぞれの例文には、ローマ字表記および英語訳が付与されている。現在までに約1500の見出しについての入力終了している。

◇維新史料綱要データベース（作成者：東京大学史料編纂所）

維新史料編纂会が1937年に刊行した『維新史料綱要』（全10巻）の綱文を全文入力し、単語をキーにした検索や、人名、官職名、地名・国名、こと項名などの項目をキーにして必要な綱文を検索することが可能となっている。また、その綱文を持つ稿本の出納番号が付与されているので、『大日本維新史料稿本』の迅速な検索ができる。現在は文久二年正月～慶応元年秋（1982～1865）の分（全体の約2割）のみが構築されている。

◇航空写真情報データベース（作成者：奈良国立文化財研究所）

地図を作成する際に撮影された航空写真に関するデータベースである。航空写真そのものはマイクロ化されており、その写真の位置情報、撮影年月日、マイクロフィルム番号などが収録されており、指示された地図上の任意の地域を含む航空写真を検索することができる。

◇埋蔵文化財文献情報データベース（作成者：奈良国立文化財研究所）

埋蔵文化財の発掘調査終了後作成される調査報告書の文献データベースである。単なる書誌情報だけでなく、遺跡名、時代などの項目が収録されており、研究者の多様な検索にも耐えられるように設計されている。

◇続日本紀データベース（作成者：京都大学・星野聰）

続日本紀の全文データベースで、一字索引などが作成されている。

◇教育実践研究報告データベース（作成者：国立教育研究所）

全国の小・中・高の教育現場で発行されている教育実践に関するレポート、実践記録、論文などの文献データベースである。データの収集は、全国の教育センター、教育研究所などを通じて行っており、

毎年冊子体の目録も作成している。

◇教育ソフトウェアデータベース（作成者：国立教育研究所）

全国の小・中・高の教育現場で作成されている教育用のソフトウェアに関する目録情報を収録したデータベースである。プログラムの適用機種、対象とする教科・科目、学年、具体的な内容などの項目が収録されている。

これらの他に、

◇地形学データベース（作成者：日本地形学連合）

◇太平御覧データベース（作成者：京都大学人文科学研究所）

◇国文学原資料データベース（作成者：国文学研究資料館）

◇東北大・松下単語音声データベース（作成者：単語音声データベース作成委員会）

などが科研費によるデータベースとして構築が進められている。また、科研費によらないデータベース構築も少なくないが、これらに関しては詳細な情報を入手することが困難であり、網羅的な調査が望まれている。以下、これらのいくつかを紹介する。

◇木簡データベース（作成者：奈良国立文化財研究所）

平城京の発掘調査から出土する木簡に書かれている文章の全文データベースである。任意の文字、出土地点などをキーにしての検索が可能である。

◇考古学遺跡データベース（作成者：筆者を中心としたグループ）

日本全国に30万以上あるといわれている考古学遺跡をすべて網羅したデータベースを目標としており、遺跡名、所在地、時代、出土遺構・遺物種別、文献、および地図上の座標位置などの項目から構成されている。現在までに佐賀県、姫路市の遺跡、約1万件が入力されている。

◇貝塚遺跡データベース（作成者：筆者および奈良国立文化財研究所を中心としたグループ）

日本全国で発見されている貝塚遺跡のデータベースで、遺跡名、所在地、時代、出土遺物、文献などの項目から構成されている。貝や動物遺存体など出土した遺物の具体的な種類が入力されている。現在までに約4000件のデータが入力済みで、最終的な校正処理を進めている段階である。

◇前方後円墳データベース（作成者：大阪電気通信大学・小沢一雅）

前方後円墳に関するデータベースで、古墳名、所在地、時代などの項目の他に、地図上の座標位置、古墳の形状データも入力されている。これらの情報を利用して、古墳を3次元画像として復元するなどの研究を行っている。

◇土偶データベース（作成者：国立歴史民俗博物館）

日本で出土している土偶に関するデータベースで、出土遺跡、所在地、地図座標などが入力されている。

これらの他に、大阪国際大学の柴山守を中心としたグループによるタイ語の三印法典データベースや、

千葉大学の加藤尚武を中心としたグループによる各種テキストの全文データベースなどがある。いずれにせよ、これらのデータベースをお互いに利用できるような態勢を整えることが、これからの課題であろう。

コンピュータ利用の現状

個人、グループあるいは機関でいろいろな研究が進められているが、ここでは機関を中心にその研究内容についてまず簡単に紹介する。

国立国語研究所では、日本語の語彙分析などに早くからコンピュータを活用している。国文学研究資料館では、収集した国文学資料のデータベース化を積極的に進めており、最近ではマイクロ化された原資料をそのままコンピュータに取り込み、イメージデータベースの構築を始めている。また、全文データベースの構築も試みている。東京大学の資料編纂所でも光ディスクを利用した原文データベースの構築を試みている。

東京外国語大学のアジア・アフリカ言語文化研究所では、各種言語によるデータベースを数多く構築している。このため、それぞれの言語のフォントの作成や、特殊な入力装置の開発を行っている。現在までに開発されているフォントとしては以下のようなものがある。

- | | | |
|--------------|-------------|-----------|
| ・ Arabic | ・ Khmer | ・ Thai |
| ・ Devanagari | ・ Burmese | ・ Korean |
| ・ Bengali | ・ Mongolian | ・ Tibetan |
| ・ Tamil | | |

国立民族学博物館では設立当初より、大型コンピュータを導入し、各種データベースの構築、民族学研究へのコンピュータの活用など、積極的にコンピュータを利用してきた。収集資料に関しては、テキスト情報のみではなく、それらの資料の写真をデータベース化しており、そのための特殊な画像入力装置なども開発している。また、最近では日本アイ・ビー・エムと共同研究で、画像データベースを利用した画像検索システムの開発なども行っている。国立歴史民俗博物館でもコンピュータを利用した画像処理システムの開発を進めている。

奈良国立文化財研究所では考古学研究にコンピュータを活用しており、ここでも多くのデータベースの構築が進められている。これらのデータベースは考古学研究に利用されるわけであるが、一般的な検索プログラムでは利用しにくいところが多くあり、それぞれのデータベースにあった検索プログラムを開発している。また、木製品の出土遺物を利用した年輪年代学の研究などでも、データ処理にコンピュータを活用している。

国立教育研究所では教育関係文献のデータベース化、およびその検索サービスを行っており、最近開始したプロジェクトとして、外国人のための日本語教育支援システムの開発がある。特に日本語教育に適した辞書および教材のデータベース化を中心に研究を進めている。

最近では研究を主目的としない博物館や美術館でのコンピュータ利用なども盛んになってきており、岐

岐阜県立美術館では、ハイ・ビジョンを利用した映像提示システムなども開発されている。

これからの課題

人文科学におけるコンピュータ利用を推進して行くための課題としては、以下のようなことが考えられる。

・文献リストの作成

これからコンピュータ利用を考えようとする人文科学研究者にとって、まず第一に必要なことは、先行研究について調べることである。しかしながら、現在のところこの種の研究に関連した文献リストは皆無に近い状況である。あるとしても、誰でもが簡単に入手できるようなかたちでは流通していない。また、この種の研究は科研費でサポートされているものが少なくなく、これらの報告書は一般的な流通ルートには乗らないため手に入れることが非常に困難である。このような報告書を含め、人文科学分野でコンピュータを利用した研究に関する文献リストの作成は早急にとりかかる必要があるだろう。この文献リストは、どこで誰が、どのような研究を行っているかという情報を提供することでもあり、それによって研究者間の交流などを促進することになる。

・データベースについて

人文科学におけるコンピュータ利用の対象は、多くの場合データベースである。研究者にとって、これまでカードやリストの形で収集、管理してきた膨大な研究資料を、データベースという形でコンピュータで管理できるようにすることが、コンピュータ利用の第一歩である。ところがどこにどのようなデータベースがあるかという情報も、文献と同じように入手することが非常に困難である。個人的に個人利用を目的としたものは別にして、グループや研究機関でいわば公的なデータベースとして作成されたものに関しては、何らかの形で流通できるようにすることが必要であろう。

これまでに構築されたデータベースのうち、公開あるいは流通されているものは、まだ数としては多くなく、また公開、流通されていても、その範囲は多くの場合狭い範囲に限定されているのが現状である。公開、流通を促進することは、今後の人文科学研究にとって重要であり、そのための方策を検討することが必要である。とくに、データベースの流通を促進するためには、できるだけ早い段階から流通のための標準フォーマットについて検討を始めるべきであろう。

・ソフトウェアの開発について

多くの人文系の研究者にとって、コンピュータを利用する場合の最大の問題はソフトウェアの開発であろう。既存の市販されているようなプログラムや、他の研究者が開発したものをそのまま利用するというのがほとんどであり、自らがプログラムを開発するのは非常にまれである。したがって、自分のやりたいことが充分に行えないというようなことも多々起こっている。専門家に開発を委託するといっても、予算上簡単にいかない場合が多い。このような状況を単純に解決するというのは困難であるが、共同でソフトウェアの開発に取り組んだり、プログラムの流通を促進するための組織づくりなどを積極的に進めていくことが必要である。

以上いくつかの課題について述べてきたが、人文科学におけるコンピュータ利用は、本格的な利用が始まったばかりというところである。これらの問題を解決して行くためには、まず研究会や学会活動などを通じて研究者間の交流を促進することから始めるべきであろう。