

歴史系支援多機能システムとデータの課題

八重樫純樹

国立歴史民俗博物館

各種分野において情報処理システムの活用は、いまや常識化しており、歴史系の人文科学諸分野においてもパソコンの普及にともない、非常に活発化しつつある。しかし、情報システムのこれら分野への応用研究や多くの試みはすでに十年以上になろうとしているが未だ試みの状態であることも現実である。これらは、各分野における情報あるいはデータの問題が基本的に情報システムの上で生かされていない可能性もある。ここではこれらの観点から、データの問題と、このデータをハンドリングするための情報システム機能について、従来から進めてきているシステム開発の経緯やこれにともなった問題を整理し、まとめ、報告するものである。

Some Problems of
Multifunctional System
and Data Aiding for
Historical Studies

Junki YAGASHI

National Museum of Japanese History

Jyounai-chou-117, Sakura-si, Tiba285, Japan

Now, it is a matter of common method to use of information system for academical object in various branches of sciences. Although many trials and experiments using of information system had been performing in many branches of Japanese cultural sciences for ten years, it is not to say that the methodologies of using of information system is established in these branches. Some reasons of this state are possibly to become that essential informations or data are not made a good use for system. So, this paper describes the characteristics of information or data in these branches, and some problems of system based on multifunctional system that we have been developing for some years.

1. はじめに

情報システムを中心とした場合の、歴史系人文科学研究支援を目的とし、ここ数年来、大学、文化財関係諸機関関係者の協力のもとで、共同研究としてこれら歴史系データの処理実験、システム開発等の研究を進めてきた。ここではこれらの経緯とその過程における、特に基本であるデータと、それらの処理要求に対するシステム機能上の各種問題について報告する。

歴史系研究の素材となる学術資料はきわめて多く存在しているが、研究分野の歴史や方法・体質等により、分野毎の学術資料の取り扱いと整備の状況は非常に異なっている⁽⁹⁾。さらにきわめて多く存在するとはいえ、研究の対象や分野によってもその学術資料の精度と信頼性に対する認識や評価に多くのへだたりがあり、かつ同一分野においても様には評価しにくいほど精度にバラツキが多い。さらにこれらに対する一般的な基準も曖昧であるが、本質的には歴史的な事象の結果生成された情報媒体としての歴史的史・資料（文化財ともいう、以下歴史的資料とする）そのものが研究上の本質的意味を有している。しかしこれらはさらにいくつかの問題を含んでいる⁽⁸⁾⁽⁹⁾⁽¹⁰⁾⁽¹¹⁾。

これら学術資料をもとに、解明すべき視点にそってそれぞれの知識とスキルで新たな情報生成活動をおこなうことが研究であるが、以上の性質のデータをもとにすることから、必然的に研究はよく整備されていないこれら断片的データの収集と、データに対する濾過操作と意味解釈をともしつつ、推論による考察が行われる。

したがって、研究支援の課題は基本的に

- (a) データベースをもとにした学術的基準データの体系的整備（学術データベース）。
- (b) 濾過操作と意味解釈結果としてのデータに対する推論過程支援の基準データ作成システム（多機能研究支援システム）。
- (c) 濾過操作、意味解釈、推論部分への応用（知識処理）。
- (d) その他

本研究においては（a）、（b）の段階を基本として進めている。特に（b）については、あつかうデータがイメージ画像であり、線図形であり、記述文であり非構造的な部分を多く含み、オブジェクト指向型システムが望まれるのは明らかである。ここでは歴史系人文科学研究の特質の問題を整理しながら、研究開発の経過とこれらデータ、処理の問題点を示すものである。

2. 歴史系学術資料と研究支援モデル

2.1 歴史系学術資料

（1）累積性と更新性

歴史系研究は基本的に過去の状況に関する事柄を明らかにしてゆくことであり、様々な学術資料も、すべて歴史的な事象の発生から現在までの間の時間と研究分野の歴史的時間の累積のなかで集成されてきたもの全てが意味をもってくる。このため、新たに発見・生成される学術資料は資料として新たな意味をもつが、従来からの累積の中ではじめて学術資料として意味をもつ。極端に言うならば、新たな資料も、過去の資料も同一価値であるか、あるいは場合により、過去に集成された資料がはるかに高い価値と現在性を有することも少なくない。

したがって、未来的要素が多く、かつ再実験可能な自然科学系研究の学術資料と比較し、実験は不能であるため資料の保存・管理が重要であり（資料館、博物館の存在）、過去への遡及調査が極めて重要な事柄となる。

（2）おおまかな種類

これら歴史系学術資料をおおまかに分けると以下の二つからなる、

（a）一次的資料

一つは往時の歴史的な事象の結果生成された、事象の情報媒体としての歴史的資料そのもの。しかし、離散的にしか存在せず、かつ再生不能な情報媒体（個別資料意外にも、遺跡、史跡等も含む）。いわゆる文化財としてあつかわれもの。

（b）二次的資料

一つは歴史的資料をもとに、専門家あるいは研究者により二次的に生成された調査報告書、論文、資料集、目録、図録等、現在的な二次資料。

2. 2 資料情報の性質と研究調査の関係

(1) 一次的資料の問題

ここには基本的に二つの問題がある。

一つは歴史的事象の結果として生成されたものであるが、これら全てが現時点に存在しているわけではなく、ほんの一部である。さらに意図して残るものは希であり、離散的あるいは確率的な過程を経て現在に至っているものが殆どである。したがって、情報媒体とはいえこれらが歴史的事象のどの部分の情報を包含し、かつ存在として一般性を有するかどうか（母集団となるかどうか）、確認しようがない部分がある。

ここは、研究調査員の認識の問題となる。

一つはこれら資料の発生から現時点までの時間経過のなかで人為的かつ自然的作用を経ている。つまり、資料の発生から現在までに様々な雑音が付加されてきており、その極端な場合が消滅である。

往時の情報と雑音を濾過する処理も研究調査員の認識によるしかない。

また、分野にもより、これら資料の管理・保存・利用の方法と状況は異なるが、一般的に、地方あるいは地域的な分散管理が原則であり、所在あるいは現状資料の状況情報の交換も充分には一般化されていない。

(2) 二次的資料の問題

上述のように二次資料の生成は研究調査員の認識と専門知識とスキルをベースとした広い意味における信号情報処理の結果である。

つまり、

- ・調査員の調査視点（問題意識）
- ・判断する場合の基準値（認識処理・知識処理）
- ・高度な熟練

これらは、調査員の知識レベル、スキル、経験、感性に多くは依存してくる。したがって、これらの結果である二次資料は精粗ひらきが大きく、かつ、調査員の視点から漏れた情報はこのなかから拾えない。

しかし、現実の学術資料そのものは文化財行政の中

意の資料を調査できるような状況ではなく、また、対象により再調査不可能な場合も多く、この二次資料生成の意味は極めて大きい。

また、歴史学、国文学のように、長い歴史と伝統のある分野では二次資料そのものも時間系列のなかで、雑音空間に埋没しているものも少なくなく、さらに、多くの人手の加わっている資料は一次的資料とすべきか、二次的資料とすべきか判断しきれぬものも多い。

これらの目録と所在情報の一元化（データベース化）も一部、開始されはじまったという状況である。

2. 3 研究支援の視点とモデル

(1) 支援の視点とアプローチ

研究は種々存在するが、資料そのものの同定・認識と歴史的事象に関する推論操作が基本である。2. 2で示し、かつ触れたように、基盤となる資料の一般性ある同定と認識を抜きに説得ある推論は成立しない。これは二次資料においても同様である。

また、分野にもよるがこれら離散的かつ不均質な素材によることから、研究方法論自体明白に確立した体系が存在しているわけでもない。このため、研究プロセスとその結果は個人差が非常に大きく、単なる個人的見解にすぎない場合も往々に存在する。したがって、個人研究の視点と方法に密着するか、あるいは意味解釈・推論等有機的部分の問題に関係する場合、極めて注意を要する。

これらの個人研究あるいは意味解釈にコンピュータを応用し、その結果を、データと方法の問題を抜きし説得性を高めるため客観的方法とした場合、逆にコンピュータ不信、ひいては情報処理一般の不信と人文科学研究における方法論の存在としての否定に結び付く場合がある。

したがって、無機的あるいは明白な部分に支援の視点を設定せざるを得ず、かつそれが一般論として成立してゆくのではないかと考える。

(2) 支援モデル

上記視点から、これら研究支援の位置付、研究過程（巨視的意味の情報処理）への意味と関係を明白にするため支援モデルについて考察する。

で管理されており、すべての研究者が任意の時に、任からの情報処理過程全般における関係として考える。つまり、歴史的事象そのものはヒトと自然あるいはヒトとヒトにおいて発生した情報処理過程である。この事象はN次元情報空間⁽⁴⁾で発生した新たな情報であり、結果としての歴史的資料はその情報媒体である。

研究調査員はこれら資料および二次的資料をデータとして新たな情報処理と情報生成活動（研究・調査）を行う。支援はこの情報生成活動のどの部分になにを提供するかということである。

2. 1、2. 2、2. 3（1）から、これらの情報空間を満たしてゆくための、部分的であれ、可能なデ

ータとシステムを形成してゆくことであろう。

ここでは巨視的観点から整理し、支援を考えると、
・基本としてよく整理・集成されていない一次あるいは二次的資料の整理・集成の方法と実現（索引データベース）

・歴史系個人情報処理の中間素材作成処理システムの実現（多機能システム）

である。

これらの全体モデルをFig. 1に示す。

ここにおいて、最も重要な問題は歴史系学術資料の諸問題から、入力するデータとコンピュータの応用視点がデータの性質に合致するか否かである。

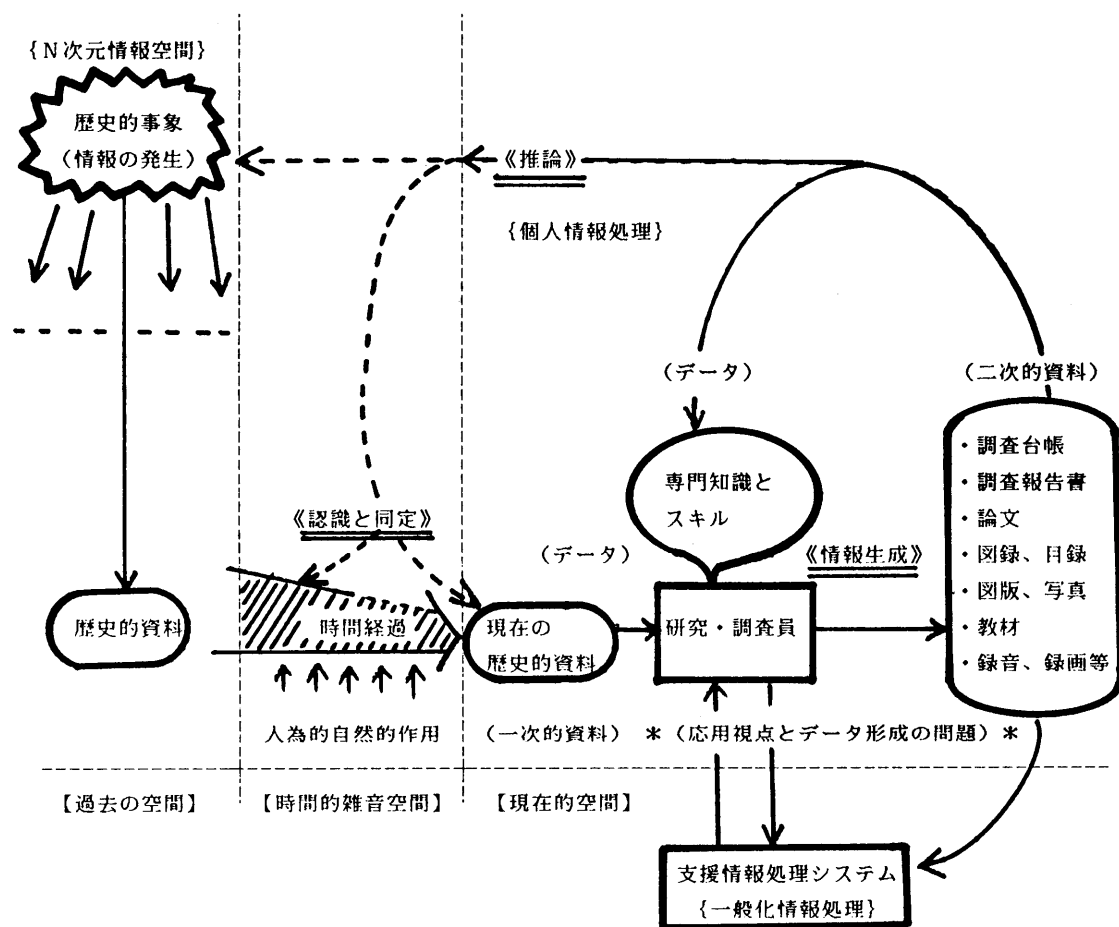


Fig. 1 研究支援の構造モデル

3. データ基本形と機能の関係

資料とその情報生成方法自体に2章で示したような問題が存在する。これら性質をさらに整理し、データ形成の観点から考察を進めると以下のように集約されてくる。

3. 1 巨視的情報要素

研究における主な推論対象はF i g. 1における歴史的事象（以下事象とも略す）であり、これはN次元情報空間⁽⁹⁾で発生した事柄である。これを一つの事象要素に着目し分解して、全体関係をみると、時間情報と地理空間情報要素の上に事象が成立するところがわかる⁽⁴⁾。事象要素（あるいは結果として一つの歴史的資料（以下、資料とも略す））をデータとして置換すると、二次的資料情報の関係から、以下の4項情報が基本要素となる。

《E》：＝＜T＞＜G＞＜O＞＜C＞……………①

E；歴史的な事象あるいは歴史的資料データ

T；時間情報要素

G；地理的空間情報要素

O；事象あるいは資料内容情報要素

C；文献、所在等現在の二次情報要素

これら4項情報要素にはまた、それぞれ問題と展開の方法が種々存在する⁽¹¹⁾。特に地理的空間要素情報は時間情報要素と密接に相関する部分があり、かつ事象とも密接に相関している。歴史系データ全体の辞書機能、あるいは事象、資料等のバックグラウンドデータとして重要であるが、歴史的な部分はコンピュータ処理データとしてまだ体系的に整理されていない。

しかし、支援システムにおいては、歴史地図作成等の機能要素は必要となり、地理空間データにこれに対応するデータ要素を設定する必要がある。

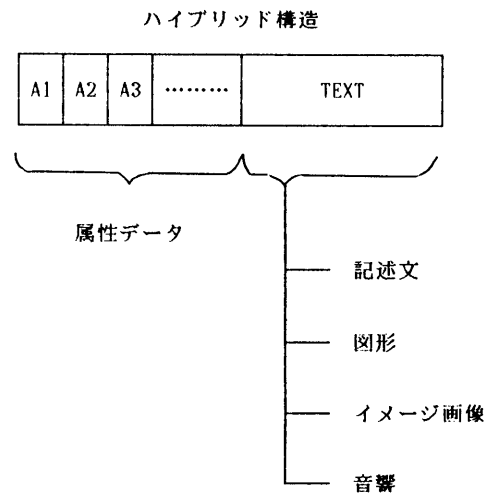
3. 2 基本形と要求機能

これらは、F i g. 1に示したように、各種データ媒体が存在する。これらは記述文、図形、イメージ画像等各種あり、研究者はこれらを総合的に活用し、新

たな情報生成活動をおこなう。

しかし、いかに二次的資料とはいえ、2章でも述べたように、内容の基準や形式は多様であり、これらすべてを要素化し構造を設定することは極めて困難である。したがって、これらをデータ化する場合、基本的に要素化可能部分（属性データ）と要素化不能部分（テキストデータ）のハイブリッド構造として捉えるしかない。

この関係をF i g. 2に示す。



F i g. 2 基本データ形式

これらにおける問題は

- ・属性データの用語や構造の基準化と体系化
 - ・およびテキストデータの実現基準、アクセスの実現方法
- である。

2章および以上で示したように、これらハイブリッドデータのハンドリングは現在の課題であるマルチメディアデータベースシステム、あるいは事実型データベースの実現そのものである。さらに、事象あるいは資料データの検索結果を操作処理する機能も支援として必要であり、事実型データベースの上に複合処理機能の付加された多機能事実型データベースシステムとして実現してゆかねばならない。

4. 多機能支援システムとその研究開発

4. 1 システムの設定と開発指針

(1) 目的と設定基本指針

本システムの開発研究は事象および資料データをN次元情報空間上⁽⁹⁾に再現し、情報処理の方法論で何等かの歴史的意味の問題にアプローチし、かつ新たな方法論として実証してゆくことが基本目的である。また、データのテキスト部処理問題が歴史系研究の情報処理と意味解釈の問題、あるいは歴史系研究のコンピュータ応用視点と支援の方法問題(機能と方法のプロトタイプ)解決にアプローチする基点ではあるまいかと考えた。

したがって、システムの設定方向はFig. 2のデータハンドリング機能の必要性、さらにN個に分解された事象部分⁽⁹⁾のデータ群の相互関連能力を満たす機能が前提として必要であった。このためには関係型に近い情報検索システムを中心に図形、画像データ処理機能の統合的構築である。

(2) システム開発の方法

開発は基本的に(1)の指針に添うが、2章で示したようにデータと要求機能は分野や対象とする課題、また目的により問題が異なってくる。さらに、それが歴史系研究者毎の視点と興味により、異ってくる。

さらに、ともすればとてつもなく高度な処理要求となり、その実現に要した時間と知的労力の結果がその分野や対象課題一般に一般性を有さず、徒労に終わる結果となりかねない。さらに、システムの仕様決定にいたる論理的分析思考がこれら分野の研究方法に定着しておらず、完全にボトムアップ指向であり、仕様の変更は日常と思ってさしつかえない。

したがって、仕様と機能の一般性を見通すため、歴史系対象分野あるいは課題毎の研究背景と位置付、問題点を把握する必要(研究概要の把握)があり、機能として現在の技術の実現範囲で解決してゆかなければ集束しない。このため類似の機能的仕様を有する複数の研究者の意見を調査し、機能単位にはボトムアップ的にいくつかの実験の上でこれらを統合する方法を進めた。

4. 2 経緯概要とシステム

(1) 経緯概要

これらは昭和59年度から以下のシステムをもとに開始した。

[第一期] (59年度以降)

- ・ VAX 11/780 (VMS; 1MB);
ファイル; 256MB×2、64MB×1
主な言語 ; FORTRAN
情報検索システム; 日本語 DATATRIEVE
図形、画像ソフト; UNIRAS (RASPAK)
図形、画像端末 ; VS 11、V80
- ・ NEXUS 6400 画像プロセッサと周辺機器
- ・ PC9801E
- ・ 一部 UNIVAC 1100

本システムにおいて、以下の基礎実験をおこなった。

- ◎各種分野のデータ構造的性質の分析と実験、
 - ◎ユーザインターフェイスの実験、
 - ◎線図形データベースとした分布図作成システム
 - ◎イメージ画像の処理実験
 - ◎システム機能の統合化実験
 - ◎その他
- をおこなった。

[第二期] (62年度以降)

- ・ VAX 11/785 (VMS; 8MB)
ファイル; 456MB×2、205MB×1
主な言語 ; 上記の他、C、LISP
情報検索システム; 上記の他、Rdb、FMS
図形、画像ソフト; 上記の他、(GEOPAK、
GEIMAGE)
図形画像端末 ; D-SCAN、CJ5700
- ・ NEXUS 6800 画像プロセッサと光りディスク
- ・ MAC II

第一期システムの移行、拡充とともに

- ◎画像データベースシステム
- ◎線図形データ作成システム
- ◎土偶データベースの構築
- ◎その他

基本は第一期システムの移行であり、システムのバージョンアップにともなった移行の問題と、これらの

拡張、統合化を現在すすめている。

(2) 支援システムの現状

Fig. 3に第二期における全体システム論理構成を示す。4. 1で述べたように基本は情報検索システム

を中心に据え、図形、イメージ画像処理とそのI/O、データ蓄積媒体を拡張しこれらの各種実験を行いつつ統合を計って来ている。この点はボトムアップ的であり、いくつかの問題を残している。

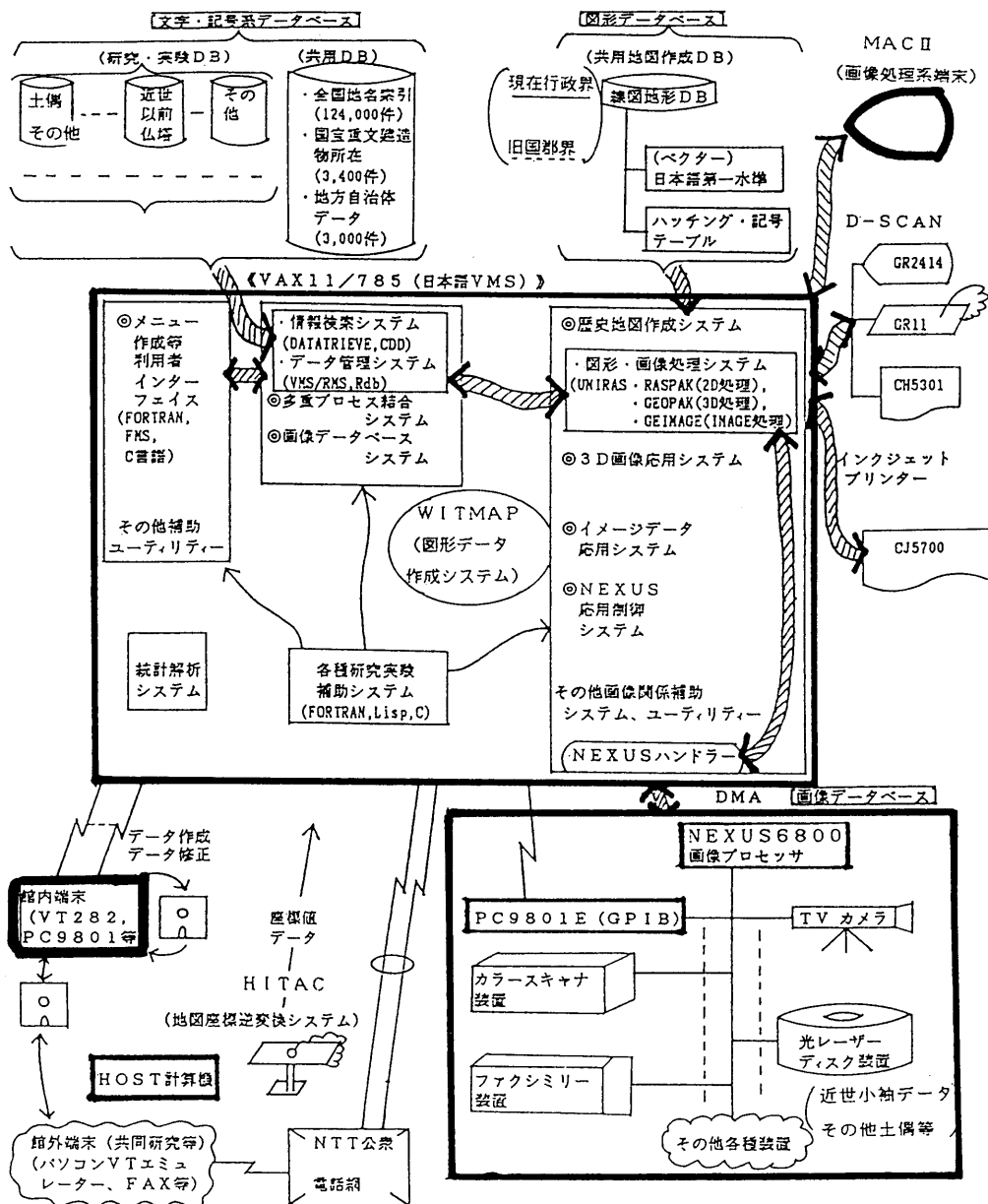


Fig. 3 支援システム論理構成図

5. まとめ

本報告は主として

- (a) 歴史系研究および学術資料の諸性質と問題、
 - (b) これらから生成される諸データの支援処理目的システムの試み、
- について示した。

(a) については、分野や研究対象により一概に言えない面があるが、一般化すると2、3章のような性質としてとらえられる。これらは具体的事例処理を積み上げさらに整理してゆく必要がある。しかし、歴史的事象や資料はきわめて多面的な情報を有し、それらが全国的に散在している状況では、人文系の体質にそぐわないとはいえ、ある程度のモデル化の上で、データ基準や標準を計らないと、データベースも次第や方法により収集つかなくなってしまう。

(b) は開始当初、システム全体のおおよその方向性は把握できたが、細部的な機能の必要条件と、データの意味的事柄に付随するアクセスメソッド、ユーザインターフェイスの問題が充分把握できず、かつ現実のデータハンドリング機能の整備が問題であり、機能とその集合としてシステムをとらえてゆかざるをえなかった。

また、開発の基になるシステムは、条件として、同一OSの管理下でデータファイルとプログラムオブジェクトが均一性を有し、かつ並列プロセス処理が可能であることは明白である。UNIXは当時日本語処理に問題があり、かつアプリケーションソフトが貧弱であった。したがって、VMSの環境下ですすめた。機能としては、これらはFig. 2におけるデータハンドリングとして、必要条件であるが、細部的な問題をさらに一般化し、考えてゆかねばならない。

これら経緯において以下の事柄が今後問題である。

◎事実データ系においては、内容と構造にシステム機能自体が密接に関係してくる。概念設計時に同時に解決をしておく必要がある。したがって、事実データ内容構造とハンドリング、処理機能の問題は外部スキーマの問題かどうか。

また、事実データ処理系の機能的一般性の問題。

◎アクセスメソッドやユーザインターフェイスはさらに専門家の研究に密着する必要があるが、ともすれば個

の都合に振り回され、すべてが局所的問題に終わりがねない。「土偶」等のデータをもとに、組織的解決の方向で行いたい。

◎研究開発の段階とレベルはすでに個人研究の容量を超えている。システムの運営、管理さらに研究経費もさることながら、知的水準においても、上記同様幅広く組織的にこなう必要がある。

なお、本研究開発は主に千葉大学工学部倉田研究室との共同研究であり、協力いただいた研究室の諸氏に感謝する。

《参考文献》

- 1) CODASYL Development Committee; 「An Information Algebra」 ACM (1962)
日本UNIVAC総研紀要5. 1 (1975)
- 2) 小林功武; 「データとデータ処理の構造」 情報処理学会誌 Vol.17 No.10 (1976)
- 3) 植村俊亮; 『データベースシステムの基礎』 オーム社 (1979)
- 4) 有澤 博; 『データベース理論』 情報処理学会 (1981)
- 5) J.D.Ullman; 『Principles of Database Systems』 Computer Science Press (1980)
- 6) D.C.Tsichritzis, F.H.Lochoovsky; 『Data Models』 Prentice-Hall (1982)
- 7) 八重樫純樹、小林達雄、野口正一; 「縄文時代土偶の情報構造に関する基礎的考察」 国立歴史民俗博物館研究報告 第3集 (1984)
- 8) 八重樫純樹、小林達雄、野口正一; 「歴史的”もの” 資料データベース構築への基礎的考察」 情報処理学会 データベース研資 41-3 (1984)
- 9) 八重樫純樹; 「歴史系研究支援情報処理の研究」 情報処理学会 情学基礎研資 1-4 (1986)
- 10) 八重樫純樹; 「歴史的資料画像データの問題と研究支援システム」 情報処理学会 情学基礎研資 8-2 (1988)
- 11) 八重樫純樹; 「歴史的データの基本構造に関する研究」 1989年情報学シンポジウム(1989)