

## FDIP プロジェクトについて

柴山 守：大阪市立大学学術総合情報センター

重点領域研究「沖縄の歴史情報研究」プロジェクト（領域代表者筑波大学岩崎宏之教授）では、琉球・沖縄に関する研究と共に、特に一次史料の蓄積・提供方式について検討してきた。諸史料の収集を進め、史料所在情報を整備することや、時代区分、分野等の多岐にわたる膨大な量のテキストや画像情報をいかに提供するかにあつてである。沖縄研究そのものは別として、具体的には(1)全文テキスト情報、(2)文書・絵図の画像情報、(3)テキスト・画像情報の目録作成と電子化、(4)テキスト、画像情報、目録等のデータベース構築、(5)提供方式の実験と確立が重要な課題であつた。

以下では、これらの諸課題うち、マイクロフィルムに関する画像データベースの提供方式を中心に報告する。

### [1] 古文書画像提供上の諸問題とその解決

「沖縄の歴史情報研究」プロジェクトでは、古文書画像を如何に蓄積し、提供するか多くの解決しなければならない問題点があつた。(1)過去から蓄積されている膨大な資料を有効に利用する。特にマイクロフィルムがそれに該当する。(2)新たに収集した画像を今後長期にわたり有効利用できるように保存する。(3)古文書や絵図の形態は、多様である。(4)歴史研究者が使える利用形態を念頭において、多様な媒体や提供方式を考える。現在の古文書画像の保存・提供では、殆どがマイクロフィルムである。即ち<アナログ>で、モノクロの世界である。デジタル技術における記録メディアの技術的動向やネットワーク環境の進展を眺めながら、改めて<アナログ>か<デジタル>かの判断が求められた。

<デジタル>は劣化がなく、計算(画像)処理に向くが、電子化の際の品質に依存する。<アナログ>は、従来の蓄積が豊富で、解像度に依存しない。古文書のような大量の情報である場合に、完全なデジタル化は、電子化の際の品質に固定されることや日常的に管理し、提供する際の人的、物的な問題を考えると容易な話ではない。

結論的には、「デジタルは万能ではない」、「歴史研究では高解像度の画質が必要になる場合がある」等から、つぎのような方式を採用した。(a)量的に多くデジタル化の採算に見合わない古文書画像は、アナログ、即ちマイクロフィルムで保存・提供する。(b)標準的な画像や頻繁に参照される画像はデジタルで扱う。このため、(1)マイクロフィルムによる保存と提供を考え、特に検索時にダイナミックにデジタル化する方式、(2)CD-ROMによるデジタル化の方式、(3)高精細な画像のための Photo-CD による3つのデジタル化方式を選択した。

特に前述の(a)におけるマイクロフィルムは、撮影時点の状態を保存でき、閲覧できる。複写が容易で分散的な保存と閲覧が可能、且つ省スペースの特徴をもつなど、過去から蓄積されてきた記録媒体としての利便

性、安全性に富む。また、過去からの蓄積が多く、膨大なデータ量がすでに存在しており、デジタル化を考慮しても作業は、時間的・経済的に大変な負担になる。さらに一旦デジタル化すると以降、分解能を上げることは出来ない。

以上のような結論から、マイクロフィルムでは、16mm カートリッジ・フィルムからデジタル画像に変換・入力できるイメージプリンタ（以下、FDIP 装置という）を採用することとした。また、ネットワークを介して諸資料を提供、発信するシステムが適当であるとの判断に基づき、WWW(World Wide Web)によるシステム開発を行うこととした。

## [2] WWW とマイクロフィルム画像 DB の結合

スキャナ機能をもつ FDIP 装置には、カートリッジ 200 本（最大 5,000 コマ/カートリッジ、全体で約 100 万画像）が格納できるオートストッカ（自動保管庫）を接続できる。このシステムを WWW と結合すれば、最大 100 万コマがインターネットを介して提供できる。さらにマイクロフィルムの目録データベースに、マイクロフィルム画像を読み込み、デジタル化する FDIP 装置を結合すれば、ネットワーク上の WWW により古文書を中心にした大規模な史料集の検索システムが実現される。

WWWによる画像検索システムの設計は

### （1）目録情報とマイクロフィルム画像の結合。

目録情報をマイクロフィルム画像データと結合する。この結果、キーワード検索による画像の特定と一次史料としての画像情報が提供できる。

### （2）マイクロフィルム資料の長期的提供。

大量の蓄積があるマイクロフィルムをそのままの形態で提供でき、さらに検索時のデジタル化によって拡大・縮小・回転等のデジタル画像処理を行う各々の長所・短所を補える。

### （3）簡便なインターフェースの指向。

計算機の非専門家が利用者であることから、検索システムやデータベース構築時に簡便でなければならない。

### （4）多様な提供手段の一つである位置づけ。

CD-ROM や PhotoCD など、他のメディアの提供にも拡張可能なように考える。

### （5）プラットフォームに独立した検索機能。

構築されたデータベースや各々の機能が機種や特定の OS に依存しない設計が重要である。

以上の考え方に基づいて、インターネットを介してWWWブラウザにより資料を検索・閲覧できるようにする。

## [3] マイクロフィルム画像検索システムの実現

本システムのベースになる FDIP 装置は、ワークステーションから SCSI インターフェースで制御される。これを図 1 に示す。また、この図 1 に示したシステムをネットワークに接続し、実現した検索システムのハードウェア構成を図 2 に示す。

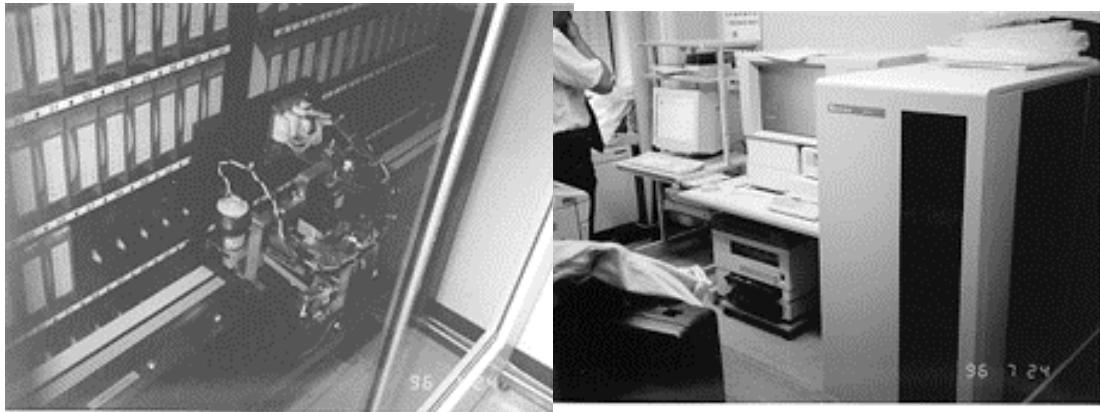


図1 WSとFDIP装置

100万コマの大規模な画像データベースが可能となるが検索時にはカートリッジ識別番号に対応するセル番号（オートストッカ内でカートリッジが格納される棚番号を意味する）とコマ情報、即ち物理的情報が特定されねば、当該マイクロフィルムを検索することはできない。一方、格納する歴史資料は主題別になり、古文書であれば標題、概要、年月日、差し出し、宛先等の情報、即ち論理的情報が存在する。

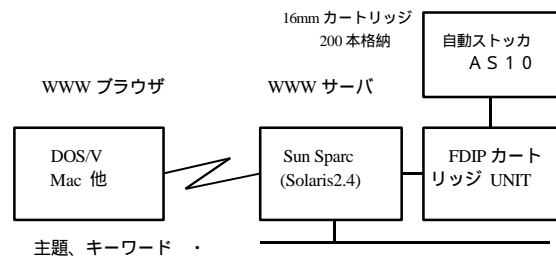


図2 検索システムのハードウェア構成

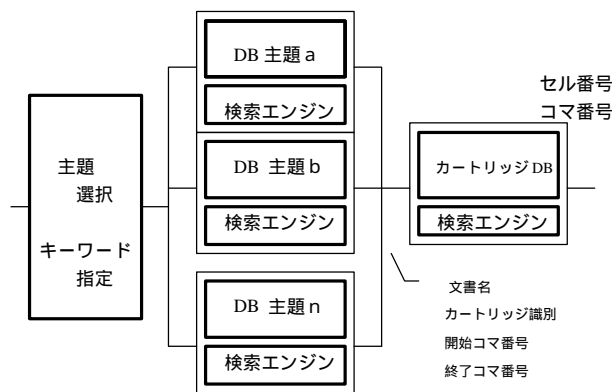


図3 キーワードからセル、コマ番号への写像

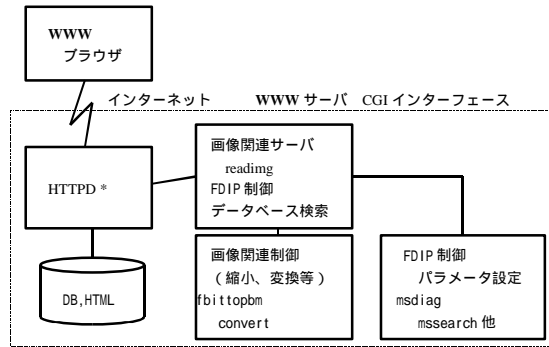


図4 ソフトウェア構成

検索時の論理的な情報から、物理的な情報へ一意な写像を行うために、図3に示すように2段階(種類)のデータベースを作成している。図3において、第1段階の特徴は、応答時間を短縮するために、主題別または副主題毎に独立したデータベースを構築する。また、第2段階でのデータベースではオートストックに格納されているカートリッジ毎に少なくとも(1)フィルム極性(ネガ/ポジ)、(2)プリップ極性(ネガ/ポジ)、(3)プリップ停止位置オフセット、(4)像方向(上下左右)等の情報に加えて、現在どの物理カートリッジがマウントされているかの状態を識別している。即ち、ここではマイクロフィルムの属性を確定する。これら2種類のデータベースは、WWWサーバ上のCGI(Common Gateway Interface)を介して、管理・操作される。

WWWブラウザからマイクロフィルム画像の検索を実現するために、図4に示すようなソフトウェア構成を実現している。

#### [4] WWWによる検索システムの今後

マイクロフィルム画像を検索時に動的にデジタル化する本検索システムは、基本的に<メカ>に依存している。したがって、<メカ>やマイクロフィルムそれ自身での障害発生が気になるが、現在のところ順調に稼働している。表3に大阪市立大学学術情報総合センターにおけるマイクロフィルム画像検索システム(伏見屋善兵衛文書、森文庫)のWWWによる利用状況を示す。利用が増加すれば当然、障害を予測しなければならない。稼働状況を眺めながら、前述したデジタル化への比重を高めることは必須であると考え、つぎのデジタル化システムを開発した。

表3 画像検索システムの利用状況

月別	伏見屋文書	森文庫
Oct/1996	- 715(260)	-( -)
Nov/1996	- 898(276)	-( -)
Dec/1996	- 822(332)	-( -)
Jan/1997	- 189(127)	200(125)
Feb/1997	- 164( 98)	319(117)
Mar/1997	- 150(111)	359(165)
Apr/1997	- 228( 61)	625(106)
May/1997	- 137( 86)	188( 90)
Jun/1997	- 261(157)	237(142)
Jul/1997	- 303(232)	224(113)
Aug/1997	- 118(105)	114( 74)
Sep/1997	- 128( 66)	140( 80)
Oct/1997	- 110( 82)	221(111)

注) (1) ( )内は外部からのアクセスで内数。(2) 単位は、ページ数。

### (1) マイクロフィルム画像の自動デジタル化

WWWによりFDIP装置に対してカートリッジ番号、及び開始コマ番号から連続したコマ番号を指示する。この結果、一括してデジタル変換が行えるマイクロフィルム画像自動デジタル化システムを開発した。本システムでは、連続したデジタル化の作業中に他の利用者からアクセス要求があっても中断することなく進められるよう考慮している。デジタル化の際のWebページを図5、6に示す。

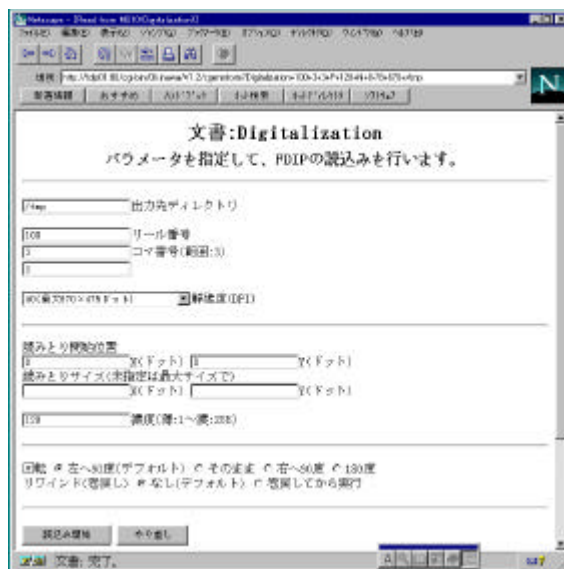


図5 デジタル化指示ページ

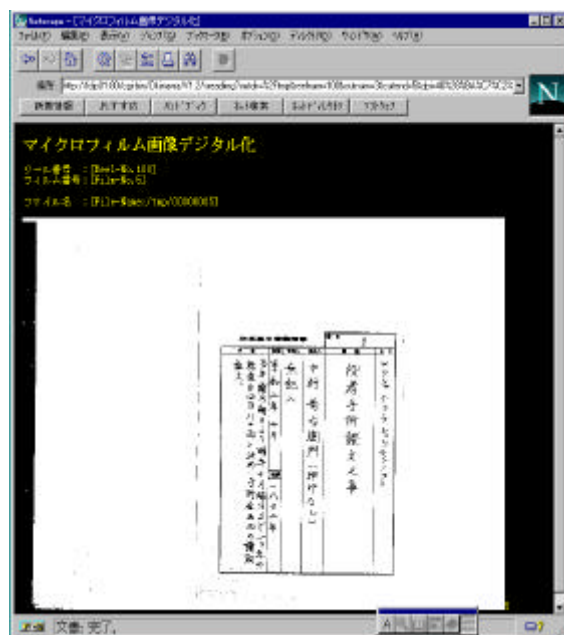


図6 デジタル化完了通知ページ

### (2) CD-ROMの利用

デジタル化された古文書画像の代表的記録メディアとして、CD-ROMが挙げられる。現在、マイクロフィルムからデジタル化した画像の保存に用いられることが多い。国文学研究資料館や東京大学史料編纂所では、CD-ROMチェンジャーを用いたシステムが実現されている。また、今年4月からカラーマイクロフィルムを中間媒体としたCD-ROMへの

デジタル化技術も登場している。しかし、大量の古文書画像に対しては多量の CD-ROM が必要になり、ネットワークでの提供では新たな問題が生ずる。また、メディアの技術革新に対しての対応も念頭におかねばならない。

### (3) データベース検索エンジンの利用

ハードディスクの利用や検索効率の視点からは、汎用のデータベース管理システム Oracle7 や Informix、Illustra に WWW を結合した方式の古文書画像データベースが考えられる。〈メカ〉を伴わない点で信頼性に優れている。しかし、それなりの専門的知識や永続的な管理体制、技術革新に対応する人的・物的諸条件が備わっている必要があり、「沖縄の歴史情報研究」プロジェクトには馴染まないと考えている。

### [5] マイクロフィルム画像検索システムの利用方法

WWW を介してマイクロフィルム版画像データベース「琉球史料集成」の利用法は、以下の節で取り扱う。