

WWW を利用した貴重資料データベースの作成について

坂本 江見

富山大学総合情報処理センター

富山市五福 3190

(0764)-45-6948

(0764)-45-6949

sakamoto@cns.toyama-u.ac.jp

富田 正弘

富山大学人文学部

富山市五福 3190

(0764)-45-6157

tomita@hmt.toyama-u.ac.jp

概要

歴史学・人文学的に貴重な資料である古文書を、劣化や汚損から守るためにデジタル化し一般公開および後世に残すために役立てようという動きが高まっている。

ここでは富山大学附属図書館に所蔵されている川合文書に着目し、WWW ブラウザから目録と画像の検索／表示を行うことのできるシステムの作成について述べる。

キーワード

データベース, 貴重資料, 川合文書, WWW

1. はじめに

ハードウェアの性能が向上し、ATMなどの高速ネットワークを敷設する機関や企業が増えてくると、テキストの文章以外にも画像、映像、音声などを含むマルチメディアデータをネットワーク上で取り扱うことができるようになり、これらデータの収集／蓄積／配布をコンピュータ及びインターネット上で行いたいという要求が自然と持ち上がってきた。今やコンピュータは理工系研究者や技術者のための計算機の役割をするだけにとどまらず、日常生活でのコミュニケーションの手段としてその地位を固めつつある。

文学や古典または歴史などの人文社会系の研究者の研究生活においても、研究者同士の情報交換や資料の入手には、ネットワーク（主にインターネット）を介して行うことができるということが必修になってきた。

特に図書館は人文社会系の研究者にとって非常に重要な情報源であるとされているが、早くから情報処理及びネットワーク設備を介した利用に着目しており、中でもOPACの導入は図書館電子化の基盤整備の先駆けとなり、図書館利用の手段の1つとして定着した。

アメリカでも、6大学を中心とする電子図書館プロジェクトが国家情報基盤(National Information Infrastructure)上に展開するプロジェクトとして、1994年から4年計画で総額2,440万ドルの助成を受け進められている。[2][3]

日本では、国立国会図書館、情報処理振興事業協会[4] 文部省学術情報センター[5]、京都大学[6]、奈良先端科学技術大学院大学[7]などの機関でデジタル図書館の開発が進められているが、その他地方の各大学においても、附属図書館のデジタル化の一環として、古文書や絵図などの貴重資料をデジタル化していこうという動きがある。

富山大学には、ラフカディオ・ハーン(帰化名:小泉八雲 1850~1904)の旧蔵書である”ヘルン文庫”、加賀藩砺波地方の農政の記録である川合文書と菊地文書、旧砺波郡鷹栖村の幕末以降の村政記録である鷹栖文庫などが附属図書館に所蔵されている。過去に研究者等によって旧文書をテキストのコードに置き換えられ出版されたものもあるが、歴史や文学の研究や調査においては、他人とは違った解釈をする可能性があり、これらの2次資料では十分な研究・調査が行えない。

資料の現物の閲覧や撮影(もしくはそれに代わるもの)は研究者にとって必須であり、大学側はこの需要に答え広く資料を公開すべきであるが、状態維持という点から見ると、閲覧時の汚損の他、保管状況や気温・湿度の影響が原因の汚損によって状態が悪化した資料も多く、なるべく現物を厳重保管しておきたいというのが実情である。

ここで注目すべきデジタルデータは、劣化することがないので現状を保存するのに大変都合がよい。またインターネット上でのやり取りができるため、遠隔地からも即時に参照することができる。またどのような内容の資料があるかを知るために最も効果的な方法として文書目録のデータベースが挙げられる。

以上の背景を踏まえ、現物資料の保護と現状の記録、大学の持つ貴重資料の公開による研究者および一般閲覧者の時間と労力の節約を目的として、目録の検索とデジタル画像データベース検索をWWWブラウザから同時に行えるシステムを作成し、題材として富山大学附属図書館の所蔵する川合文書を取り上げた。

2. 川合文書について

川合文書は、加賀藩砺波地方の郷村吏員である十村と呼ばれる役職にいた高岡市戸出の川合家[8]が所蔵していた古文書で、一紙もの2,187点、冊子1,214冊、絵図類40点から成る[9]。藩政初期より藩政末期まで(天正から明治にかけて)の御用留、御触書、仕方書、願書その他、藩の農政を伺うための貴重な資

料が含まれている。過去3年間で学内外の大学関係者の他、郷土資料館、文化センター、市史編纂室、教育委員会などを中心に、約1,050点の資料に対して閲覧と撮影などの利用依頼があった。



Fig.1 川合文書

3. システムの概要と構成

本システムは、できるだけ多くのユーザに気軽に大学の持つ資産を公開し利用してもらうという目的から、WWWブラウザによるインターネットからのアクセスについて認証を行うことなく、無料で検索を行うことができるようになっている。

ハードウェアには IBM RS6000 F30 を用い、155Mbps ATM 幹線につながる Ethernet SwitchingHUB の10Mbps のイーサネットポートに接続している。DBMS の部分にはフリーウェアのリレーショナルデータベースシステム Msq12.0.3 を用い、WWWブラウザからのデータベースへのアクセスには CGI を採用している。

データベースは検索に使われる項目のほか、文書の管理・整理を行うのに必要な58の項目が格納されており、そのうち検索画面からは

- 文書番号
- 年号
- 月日
- 西暦
- 文書名
- 形態
- 差出人
- 宛所

- 内容の要約
- 文書中で主となる言葉

とこれら全項目をキーに検索を行うことができる。キーワードをもとに検索し得られた結果の一覧から詳細ボタンを押すとその文書について、詳細の表示がされる。このとき同時に画像データも読み込まれ表示される。画像データは100dpiの精度でフラットベッドスキャナにより撮影が行われており、ブラウザのウィンドウの横幅にあわせて画像が納まるよう表示されるようにした。

4. 人文社会系研究者のデータベース利用に関する調査

デジタル情報と図書館所蔵品の利用に関して、富山大学の人文社会系研究者27人に対し事前に意識調査を行った。これは紙のアンケート用紙を配布し、回答も用意した用紙に記入してもらった。

まずはコンピュータおよびネットワークへの利用価値については、すべての回答者から「利用したいと思う。」「利用すべきである。」もしくは「すでに利用している。」との答えが得られ、コンピュータに関して嫌悪感を持っている研究者が学内にはほとんどいなくなったのではないかと判断できた。すでに利用している研究者の主な使い道は電子メールによる情報交換やWWWによる情報公開、図書館OPACの利用、学術論文の検索などが中心になっている。

現在の資料の取り寄せ方法や利用については、「図書館、文書館などで直接見る。もしくは複写する。」「現物を見に行き行って写真撮影をする。」というように直接現物のあるところにまで出張していくという方法のほか、「マイクロフィルムを取り寄せる。」「複写の依頼をし郵送もしくはFAXで送ってもらう。」など複製を取り寄せる方法が主流であった。なかには「CD-ROMで取り寄せる。」「インターネットで見る。」など既にデジタルデータの入手をしている研究者もみられた。

次には電子メールによるアンケートで、インターネットの利用頻度と電子図書館やデータベースへの関心度を調査した。対象者は富山大学の人文学部と教育学部の文科系学科の研究者に対し行った。24人の研究者から回答が寄せられたが、紙のアンケート用紙で調査を行ったときには、約3週間で27人の回答が得られたのに対し、電子メールによるアンケートの場合は、アンケートを送信した当日のうちに17人からの返事を得ることができ、24人分の回答は4日で得ることができた。

アンケートは、電子メールの送受信の頻度（1週間に送受信する数）、WWWを見る頻度、OPACを利用する頻度と、電子図書館及び貴重資料のデータベース等のシステムに対する期待度、関心度を自身で10段階評価してもらうという形式をとった。

電子メールによるアンケート項目の概要は以下の通りである。

- 電子メールを1週間に送受信する数
- WWWブラウザで1週間にWWWページを見る回数
- 1週間にOPACを利用する回数
- 貴重資料データベースの利点はどのような点か [川合文書を例として見てもらう]
- 貴重資料データベースの欠点はどのような点か
- 電子図書館（特に貴重資料データベース）に対する期待は何点か

(点数の目安)

10点.....すべての蔵書に対してデジタル化に賛成。著作権などの規則についてもデジタル化に対応していくべきだと思う。

5点.....デジタル化に賛成。データベースがあれば利用したい。

1点.....デジタル化には反対。図書館には、先にもっとすべきことがある。

Fig.5で分かる通り、メールを頻繁に使うユーザはシステムへの関心度の低い領域には現れなかった。WWWやOPACに対する利用頻度を横軸にとったものもほぼ同等の結果が得られている。インターネットを良く利用するユーザほど、どのようなことがネットワークによって可能になるかを理解しているので、電子図書館やデータベースに対する期待も大きくなるのだと考えられる。

一方で、著作権や出版権などを気にする研究者も多く、上記の10段階評価で5を付けた回答者のほとんどは、法規に触れるのではないかという点が最も気になるとしており、古文書や古美術の写真のうちこれらの権利についてパスできるものをデジタル化していけばよいという意見が多く寄せられた。

更に研究資料の入手が比較的困難であると想定される4人の研究者に対してシステムの説明をし意見を求めたところ、

- 自分の研究素材で本システムと同じ方法で検索／表示することができれば、実物を見るために出張する回数が1/10程度に減らせる可能性がある。
- 内容を読む場合や筆跡を見るケースが多いので、このシステムで十分まにあわせることができ、原物を傷めなくてもすむので有効である。
- WWWブラウザを使っているので、新しい操作等を覚える必要がなく、印刷が簡単にできるのがよい。
- 画像のみのデータベースと違い、必要事項が検索できるので能率が上がると予想される。

等好意的な感想が得られた。

またデジタルデータおよび本システムの欠点として、

- 原稿が贋作かどうかを見極めるのが難しい。
- 紙質や厚さなどの手触り、実際の墨の色や濃淡などが判断しにくい。
- 画像の部分拡大機能の追加が必要である。
- 紙のようにパラパラとめくって探すことができない。(全体的、瞬間的、曖昧な表示ができない)
- 大学の持つ資料の全てをデジタル写真で公開してしまうと、遠隔地からも見るようになるようになってしまうので、地元研究者の利点がなくなってしまう。

などの点が挙げられた。

100dpiの精度については、端末の能力によっては読み込みに時間がかかって利用者のストレスになる場合もみられた。端末によって読み込む画像のdpiが変わるようにしてほしいという意見もあった。

一次資料を画像データとしてインターネットで見れるようになることで、遠隔地の研究者には大変便利になると思われる一方、場合によっては東京などの中央にいたほうが一次資料以外の情報を集めやすいという事情もあり、研究者同士の研究テーマの重複が頻繁に起こってしまうケースが増え、研究者の地方離れが進んでしまうのではないかという懸念もあった。

考察

人文社会系研究者にとって、コンピュータは直接研究の題材になるものではなく、研究を進めるための道具でしかない。よってシステムに要求するものは、高価で多機能なものではなく、いかに簡単に自分に必要な機能を安価に使えるかというところにある。調査の結果をみると、ここ数年で人文社会系研究者のコンピュータやネットワークに対する拒絶反応がほぼなくなってきたのではないかといえる。インターネットブームの影響もあるが、実際にコンピュータに触れる機会が増え、その便利さや簡易さが認められた結果であろう。

そのような背景で、なじみやすい GUI である Netscape や Internet Explorer 等の WWW ブラウザをデータベース利用のためのインターフェイスに選び、必要最小限の項目を検索項目として選んだことは現段階で最良の方法であった。また目録のみを扱うデータベースを用いるとき、検索の後に閲覧や文献の複写（もしくは取り寄せ）をしなければならないが、本システムでは文書に関する情報と同時に画像を表示させることができるので、一度の手間で実物の様子を伺うことができるという点がかかなり有効であった。

問題点としては、手触りや微妙な色使いなどが分からない点が挙げられたが、本データベースの画像は 100dpi でスキニングを行ったので、どの程度の精度で撮影した場合に、どのような研究者の用途に見合った画像になるかという調査が今後必要であると思われる。

しかしデジタルデータ表示が実物の閲覧に比べて優れている点がある一方、やはり触覚や人間の目でしか判断できない場合には、実物の閲覧も必要である。デジタルデータを 100% 実物の閲覧に置き換えることは不可能であるが、いかに歩み寄れるか今後の課題となる。

またこれと矛盾した問題となるが、地元研究者の利点を重視しなければならないという意見と、競争による分野全体のレベルアップをはかるためにも公開するべきという意見とのやり取りの今後の展開も注目する点である。

技術的な面では、パラパラと繰って「全体性」をつかんだり、思わぬ発見をしたりする「掘り出し物的発見」が OPAC などの目録データベースが紙のカード目録よりも未熟な点であると言われているが [11]、この機能に変わるものをデータベースやブラウザに備えてゆく必要がある。またマウスで囲った部分の画像の拡大なども研究材料として利用する場合に必要な機能である。拡大に耐えうる画質のデータを用意すること、ハードディスクの容量確保のための画像圧縮などの技術向上も課題として挙げられる。

まとめ

富山大学の所蔵する貴重古文書である「川合文書」を題材とした目録と画像を検索するデータベースを作成し、人文社会科学研究者への影響と利用効果を調査した。

目録カードと比較してブラウジングの点で劣る点も見られたが、現物の状態保護に有効であること、遠隔地からの利用が可能なこと、研究者の資料探索に要する時間が短縮されることなどが期待できるようになった。

今後更に利用者のニーズにあったシステムに更新していく必要があると考えられる。

参考文献

[1] The Informedia Project URL: <http://www.informedia.cs.cmu.edu/>

[2] 杉本雅則, 片山紀生, 越塚美加, 神門典子, 高須淳宏, 安達淳, 世界の電子図書館の研究動向について, 学術情報センター紀要, 1996, pp.221-pp231

- [3] 竹内秀樹, 米国図書館界におけるデジタル情報の保存をめぐる動向, 月刊IM'97年8月号別冊,1997
- [4] パイロット電子図書館, URL: <http://www.cii.ipa.go.jp/el/index.html>
- [5] 学術情報センター電子図書館サービス NACSIS-ELS, URL: <http://www.nacsis.ac.jp/els/els-j.html>
- [6] 京都大学 Electronic Library URL: <http://ariadne0.kuee.kyoto-u.ac.jp/>
- [7] 奈良先端科学技術大学院大学 電子図書館システム URL: <http://dlw3.aist-nara.ac.jp/>
- [8] 田川捷一, 加越能近世史研究必携, 北国新聞社,1995
- [9] 富山県大百科辞典, 富山新聞社,1976
- [10] 名和小太郎, デジタル図書館と著作権, 情報処理,1996,pp.857-860
- [11] 越塚美加, 文献のブラウジングが研究課程に与える影響学術情報センター記要,1996,pp.131-142
- [12] 波多野宏之, 国立西洋美術館における西洋美術研究支援アプリケーションツールの評価, 情報処理学会研究報告 Vol.97,No.48,1997,pp.31-36
- [13] Huges Technologies Web Site, URL: <http://www.Huges.com.au/>

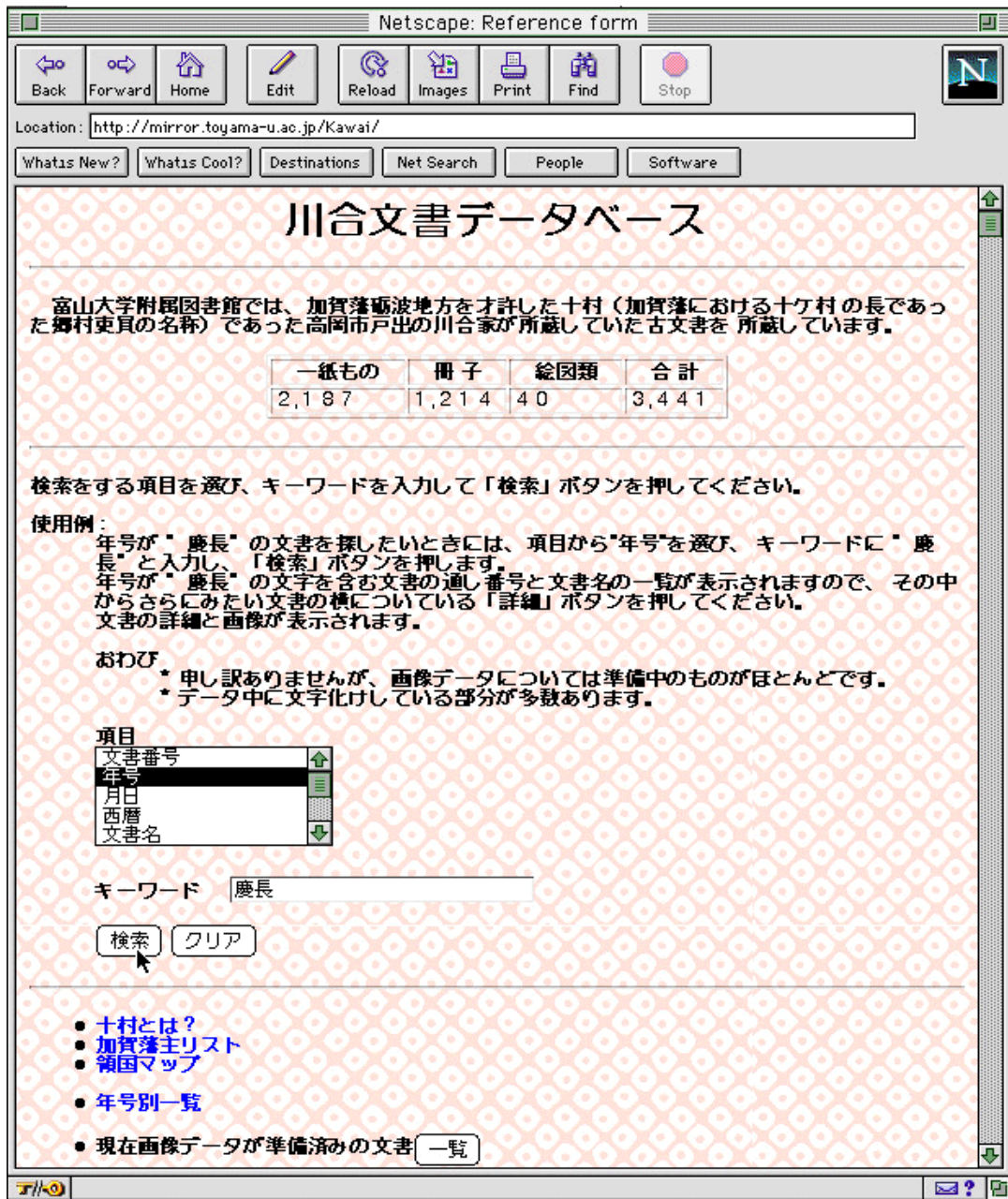


Fig.2 URL: http://mirror.toyama-u.ac.jp/Kawai/



Fig.3 キーワードをもとに検索した結果の例

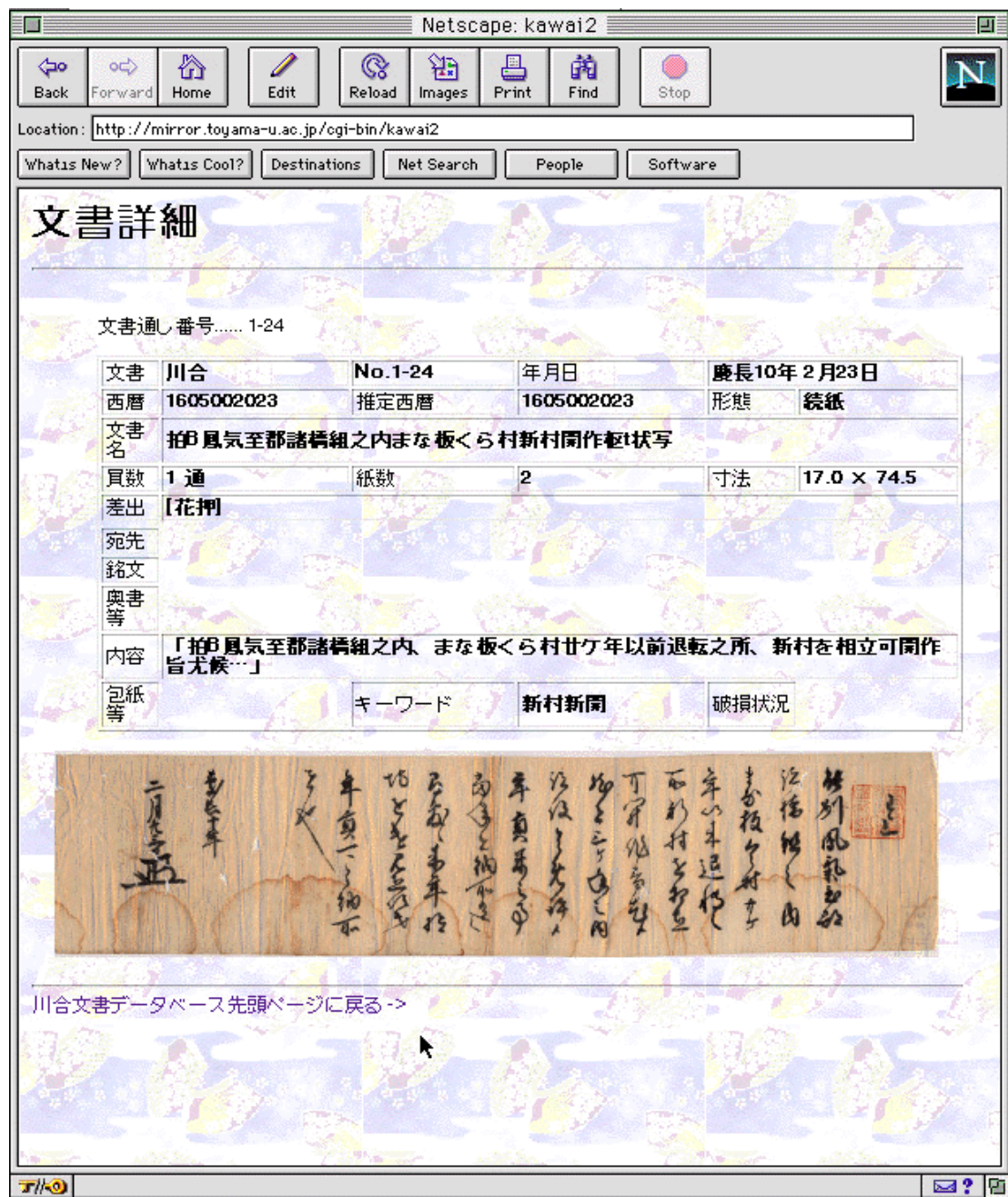


Fig.4 文書の詳細表示の例

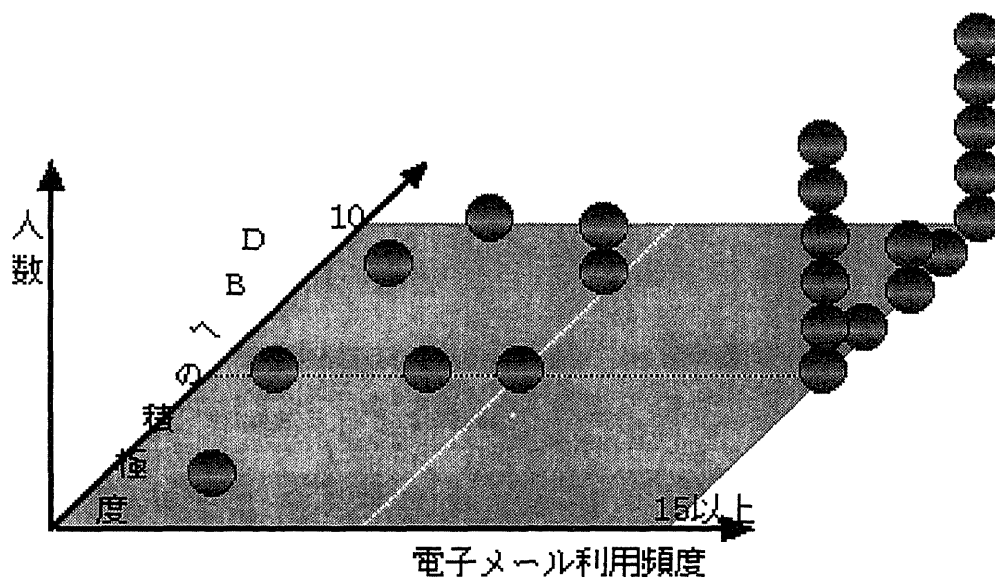


Fig.5 電子メール利用頻度と電子図書館への関心度