

## 電子図書館 | 将来像 |

2W-6

堤 泰治郎 諸橋 正幸 丸山 宏 野美山 浩 渡辺 日出雄

日本アイ・ビー・エム株式会社 東京基礎研究所

## 1. はじめに

近年、マルチメディアやネットワークに関する基礎・応用技術の発展が次第に加速されつつある。また、文書、静止画像、ビデオなどのマルチメディア・コンテンツの電子化（デジタル化）も徐々に進み始めている。これらの技術的・社会的背景を考えると本格的な電子図書館の実現も夢ではなくなってきた。従来より様々な研究開発が行われてきたが総合的に完成度の高い本格的な電子図書館の実現例はまだ見当たらない。本稿では、まず従来の研究開発について概観する。次に本格的な電子図書館の役割、メリット、機能について議論する。また本格的な電子図書館を実現するために必要な技術やアプローチについて、特に情報検索の観点から考察し、最後に日本アイ・ビー・エム株式会社東京基礎研究所での研究概要を紹介する。

## 2. 従来の研究開発（または実験、実施）の例

従来の例は、文書の電子化の程度、文書の構造化の程度、マルチメディア対応の程度、ユーザインタフェースでの工夫の度合、ネットワーク利用の程度など各種各様である。総合的に完成度の高い本格的な電子図書館の実現例はまだないが、いくつかの特徴ある例は報告されている。孫悟空 [1] や Virtual Library [2] は、仮想の図書館環境を実現するユーザインタフェースに重点を置き、CG開架図書検索やCG司書などを提案している。学術論文の頁を画像として蓄積し既存の二次情報と統合化し、広域ネットワーク上で実現しようとする学術情報センターの例 [3] もある。国内ではその他、東京工業大学でのLAN上のCD-ROMの複数端末からの同時検索の試みや、国会図書館での将来構想 [4] などが知られている。海外では特に米国のCOREアットやマカリブアットがよく知られており [5] 大学、国公立図書館を中心にネットワーク（キャンパス・ネットワークやInternet）に重点を置いた多くの実験が行われてきている。新しい試みとしては、パナソニック、IBM、Pontifical Catholic 大学の共同プロジェクトが1994年から始まり、パナソニックの貴重な所蔵品を高精細カラー画像で世界中の研究者・教育者に検索させる計画がある。また注目すべき例として、文書を単に電子化するだけでなくSGML形式とし、ハイパーテキストの柔軟な検索方式を享受しようとする試み [6] も出はじめている。

## 3. 電子図書館の役割・メリットおよび機能

基本的には、「だれでも、いつでも、どこでも、簡単に、好きな情報がすぐに手元へ」と言う環境を実現することである。電子化は、（1）書誌情報とインデックスだけを電子化、（2）テキスト本体も画像として電子化、（3）テキスト本体もすべてコード化情報とする、という三段階が考えられる。しかし、本格的な電子図書館では、（3）のように一次情報（文書・資料の本体）も電子化（デジタル化）しそれらの貴重な情報を最大限に有効活用する必要がある。

## 3.1 電子図書館のメリット

一次情報をも含めた電子化情報および、ネットワークの利用により次のようなメリットが生まれる。

**Electronic Library I - Future Prospects -**  
 Taijiro Tsutsumi, Masayuki Morohashi, Hiroshi Maruyama,  
 Hiroshi Nomiyama, Hideo Watanabe  
 IBM Research, Tokyo Research Laboratory

## (1) 内容の電子化（デジタル化）によるメリット

- ・ 唯一の文書・資料を同時に多人数に「貸し出し」可能
- ・ 保管スペースの削減
- ・ 長期保存が容易
- ・ 取り扱い注意の古文書なども高精細カラー画像で簡単に公開可能
- ・ 美術品や古文書のしみ、汚れ、変色などに画像処理が可能
- ・ コピーや伝送などの場合に劣化しない
- ・ バックアップの作成が簡単
- ・ 遠隔地でも電送が容易（物流コスト、リードタイムの大幅な改善）
- ・ 閲覧のための工夫が容易（拡大・縮小、自動読上げ、ブラウズなど）
- ・ 文書・資料全体または部分の再利用が容易
- ・ テキスト、動画像などの多様なメディアが同一の表示装置で参照可能
- ・ 内容の高度な加工が容易（機械翻訳、自動抄録など）
- ・ ユーザ自身の「ミニ図書館」の構築が容易
- ・ 相互参照が容易（同一文書内や文書間）
- ・ 出版が容易に早く、かつ安くなる
- ・ 文書・資料の改訂版を作成するのが容易
- ・ 語彙・表現、作者同定、語源調査などが容易
- ・ 文書・資料の分類やテーマごとの関連文書収集が容易
- ・ 著作権侵害調査が容易

## (2) ネットワーク化によるメリット

- ・ 遠隔地からの文書・資料検索、「借りだし」が容易
- ・ 分散された図書館でも集中管理が可能
- ・ 特定の場所に行く必要がない（オフィス、家庭でサービスを受信）
- ・ クライアント・サーバー方式で他業務との連携が容易
- ・ 文書・資料の定期配布、情報フィルタリングの実現が容易
- ・ ユーザ自身の「ミニ図書館」の開放が容易

またこれからは、ユーザは単に情報を受信するだけでなく、発信も容易に行えるようになり、双方向の情報のやり取りが創造的活動を促進させることが期待できる。

## 3.2 電子図書館の機能（現在および将来）

電子図書館が備えるべき基本機能としては、マルチメディア情報の入力・認識およびデジタル化（圧縮・伸長も含む）、オーサリング、インデクシング、データベースの構築・維持・管理・更新およびサーバー機能、情報検索・表示・閲覧および情報の再利用、アクセス管理、さらに「貸し出し」「受け入れ」などの従来型図書館業務機能などがあり、補助機能としては、遠隔対話・会議、機械翻訳などが考えられる。これらのうち、重要な基本技術を三つ挙げるとすれば、使い易いユーザインタフェースを備えた高速・高精度情報検索、高速広帯域ネットワーク、そしてオブジェクト指向マルチメディア・データベースであろう。ATMネットワークは近年急速に進展していて、156Mbpsの速度も実現されつつある。これにより例えば1年分の新聞データも数秒でユーザに転送できるし、MPEG1/2などによる圧縮動画も余裕をもって連続転送できる。大規模なマルチメディア・データベース・システムの開発、維持管理ではオブジェクト指向のアプローチが生産性向上の面で有効である。情報検索機能については以下に詳しく述べる。

## 4. 情報検索機能

テキストを対象とした情報検索では、キーワードによるインデックス方式と自由ワードによる全文サーチ方式がよく使われている。前者は、従来から親しまれた使い方、キーワードが想起しやすい場合には便利であるが、辞書の維持管理に手間がかかる欠点がある。後者は、とにかく言葉を入力すればヒット

したすべての結果を返すので、細かくしらみ潰し的に調べると都合がよい。著者らはユーザに同時に両方の機能を提供し、相互の長所を生かした使い方ができるようにしている[8]。

マルチメディア情報の検索では言語情報だけではなく、それぞれのメディアによる、より直接的な検索手法も重要になる。表1の右欄にアプローチの一例を示す。これらに関してはすでに幾つかの報告があるが、たとえばIBMでは、カラー静止画像に対して色・形・生地などによる類似検索技術(QBI C)をすでに製品化している。

	短期的アプローチ	中・長期的アプローチ
テキスト	キーワード、全文、あいまい検索など	自由な言葉による検索など
静止画	プロファイル記述(テキスト)に対し	色・形・生地、例、言語表現による検索など
動画 CG	キーワード、全文、あいまい検索など	静止画のアプローチに加えて、物の動きによる検索など
音声		発話、物音、例、言語表現による検索など
音楽		メロディー、リズム、ハミング、例、言語表現による検索など

表1. マルチメディア検索方式の例

このような直接的な検索方式は最終的には従来からのテキストによる検索方式と融合されなければならない。すなわちユーザから見れば、検索のために入力するメディアには制限を受けず、自由に指定したものがすべて有効な、組み合わされた検索条件になる必要がある。

またユーザにとっては、色々な種類のメディア(一般書、専門書、辞書、百科辞典、新聞、雑誌、地図、絵画、映画など)や各種フォーマット(テキストでは、SGML, TeXなど)を特に意識せずに、自分の欲しい情報が検索できる環境が望ましい。

本格的な電子図書館ではいわゆるマルチメディア・データベースは一ヶ所とは限らず各ユーザの個人的コレクション(個人の「ミニ図書館」)も一部開放されるであろうことを考えると、LANやWAN上に分散された数多くのデータベースをうまく検索するという分散型情報検索が重要になる。現状ではこの一例はInternet上のWWW/Mosaicである。実際には広範囲に分散されたデータベースにはその品質にばらつきがあり、ユーザにとって膨大な情報の概要を効率良く知ることが効率的な検索につながる[8]。また将来的には、このような環境では、いわゆるエージェント(またはノウボット[7])のようなネットワーク上の知的情報処理が重要になってくると考えられる。基本的な役割は、「ネットワーク上に分散された情報にアクセスし、ユーザの指示や嗜好に応じて必要なまたは関連する情報を収集する。また状況により収集情報の加工を行う。」であろう。個別機能には、エージェントの移動・ルーティング、エージェント同士の接触・情報交換、緊急度と情報品質のバランス、自然言語対話などによるユーザの指示・要求、好み・嗜好の理解・学習、高度情報加工(自動抄録、自動ハイパーテキスト/メディア化など)が考えられる。

従来の図書館では、利用者は求める文書や資料を探し出し閲覧することが主な目的であった。しかし将来の環境では次のような拡張された意味での知的情報検索が要求されると考える。

- ・ 欲しい文書を探すのではなく、文書の単位にとらわれない情報を探す
- ・ 関連情報を収集する(関連の種類、程度はある程度自動的に判断)
- ・ 欲しい情報を抽出し、さらに読み易いように加工してユーザに返す(自動抄録、テーブル化、グラフ化など)

- ・ ユーザは「質問」をしてその答えを受け取る
- ・ 意志決定や問題解決の支援を行う

また、ユーザインタフェースに関連して、自然な音声対話を含むCGによる対話型情報検索環境や動画を含むハイパーメディア、さらに膨大な情報に対処するための超高速並列検索技術などもこれからますます重要になってくるであろう。

## 5. 東京基礎研究所での研究開発

著者らは、本稿の3で述べた基本機能を満足する、高速ネットワーク上のマルチメディア分散DB検索システムとして電子図書館を位置付け、当面(1)使い易いナビゲーションつき高性能情報検索、(2)オブジェクト指向マルチメディア・データベース、(3)博物館等での実績に基づく高精細画像データベース、(4)先進的ネットワークおよび分散処理、(5)ビデオ・オン・デマンド技術などを中心に研究開発を進めている。(1)については、動画の一部を指示するとその説明が現れたりするようなハイパーメディア的な検索から、様々な切り口、整理方法を提供する、キーワード中心の高度な検索手法までを、初心者向けや熟練者向けのインタフェースと組み合わせ実現する予定である。(3)については、特に美術品や古文書、芸術品、絵画などを当面2K X 2K画素以上の高精細カラー画像で提供できる環境を考えている。(5)では、特に映画やニュースなどの非同期配信機能などを考えている。著者らは、第一ステップとして、特に、複数の観点から同時にしかも動的に検索範囲を絞りこんでゆるいナビゲーション機能を開発した。これらの詳細は他の報告[8]にゆずる。

## 6. まとめ

本稿では、まず電子図書館研究の現状を概観し、本格的な電子図書館に期待される役割・メリットおよび機能についてまとめた。次に情報検索について、より柔軟なテキスト検索、マルチメディア・データの直接的な検索手法、ネットワーク上の分散型情報検索の必要性などについて触れ、最後に著者らの研究のねらいや概要を紹介した。しかし、電子図書館の実現には幾つかの問題が残されている。まず、動画や音声などの連続データをネットワークで送る場合のQOS(サービス品質)の問題である。また、膨大なコンテンツの作りこみ(電子化)の問題、さらに、マルチメディア・データの著作権・知的所有権の保護、機密・プライバシー保護、サービスの課金ポリシーおよび方法、物流制度の改革、データ・フォーマット等の標準化の問題など広い範囲に渡っており、これらは今後の課題である。

## 7. 参考文献

- [1] 佐藤 衛 "CG画像による電子図書館「孫悟空」"、情報処理学会情報学基礎研究 24-2、1991年11月
- [2] 一色 他 "電子図書館システム「Virtual Library」の概要"、情報処理学会第48回全国大会予稿集、1994年3月
- [3] 安達 他 "学術文献を対象とした電子図書館システムの構成法"、情報処理学会情報学基礎研究 29-7、1993年5月
- [4] 長尾 貞 他 "電子化図書館の未来の姿"、電子化図書館研究会、1992年12月
- [5] 安達 他 "欧米における「電子図書館」プロジェクト"、情報処理 Vol. 33 No. 10、1992年10月
- [6] 石塚 他 "全文検索システムのリソースとしてのSGML方式データベース"、情報処理学会情報学基礎研究 33-6、1994年5月
- [7] C. Malamud 著、後藤 地 訳 "インターネット縦横無尽"、共立出版、1994年3月
- [8] 諸橋 他 "電子図書館I I-V I"、情報処理学会第49回全国大会予稿集、1994年9月