

4 NAIST 電子図書館の状況

4.1 電子図書館システムの現状と将来像

本学の電子図書館システムの調達は、1995年度より7ヵ年計画で導入が進行している。この調達において、各年度に調達されるシステムは4年間のレンタルを基本としており、1998年度(1999年1月)に導入されたシステムをもって、ひと区切りがついたこととなる。今後は、導入されたシステムの更新と追加によって調達が進められることになるため、1999年度の調達においては、これまでのシステムの設計を見直し、より快適なシステムの構築と新たな技術の導入を行っている。ここでは、1998年度までに導入されたシステムの現状を述べるとともに、次の4年間のシステム設計方針について概観する。

現行システムの設計方針と現状

大学附属図書館は、あらゆる情報源を利用して教育研究に必要不可欠な情報収集活動を支援するものでなければならない。また、学内で生成される学術情報を外部に対して積極的に提供するための情報発信機能も有していなければならない。このため、本学附属図書館では以下のような目標のもとに、先端的な教育研究活動を支える基盤環境の構築を目指している。

- 国際的な規模での情報流通のグローバル化への対応
- 従来の印刷物中心の図書から音声、画像、映像等を含むマルチメディア情報への対応
- 情報の収集・公開を加速するための情報源に関する情報の収集、一次情報(論文や記事そのもの)の収集、および一次情報保有機関への迅速なアクセスルートの確保
- 利用者に最も近いところでのサービスの提供

さらに、上の4つの要件を満たし続けるためには、高度情報化社会における図書館を取り巻く状況の急速な変化(新しいメディアや情報サービス形態の出現、出版形態の変化、研究発表・情報発信形態の変化など)に柔軟に対応する必要がある。このため、ある時点で可能なサービスを固定的に提供するのではなく、常にサービスの種類・質と図書館の役割を検討しつつシステム構築を進め、時代の変化に柔軟に対応することを目指している。また、附属図書館は利用者サービスの観点からは以下の三つの性格を有することを目標としている。

- メディアセンター – 利用者が資料の種別（図書、雑誌、音声、ビデオ等）を意識することなく情報提供サービスを受けることができる。
- 居ながら図書館 – 図書館に出向くことなく、利用者に最も近いところから検索・閲覧を行なうことができる。
- 24 時間図書館 – 情報提供の時間が限定されず、開館時間の概念がない。

電子図書館システムは、本学の情報環境（曼陀羅ネットワークおよび曼陀羅システム）を基盤として、上述の理念を具体化するための中核設備である。

こうしたことを実現するため、原則としてシステム内に一次情報を蓄積し、それらを利用者が検索して利用する機能を提供することを中核として、これらを支援する機能とともに設計が行われている。具体的には、システムは以下の6つのサブシステムで構築されている。

1. 一次情報入力システム

本システムは書籍情報の一次情報を電子化するために必要不可欠なデータ入力・メディア変換用機器群である。附属図書館が入力の対象とするメディアには、冊子体、CD-ROM 及びネットワーク経由のファイル入力があり、特に冊子体の場合はカラーで表現された写真が含まれている。ここでは、冊子体情報をスキャナで読み込むとともに、OCRによる全文検索用のデータの生成、目次情報の入力、システム内で利用されるデータ形式への変換編集等が行われる。また、CD-ROM やネットワーク経由で入手されたファイルについては必要なデータ形式への変換を経て、システム内に取り込まれている。

現在5台のモノクロスキャナ装置、6台のカラースキャナ装置とこれらを支援するためのワークステーション群で構成されている。

2. デジタルビデオシステム

今日の教育研究活動では、映像情報の重要性が増大している。本システムは、外部から映像情報を収集するとともに、本学において映像情報を製作し、他の図書情報と統合化した形でデジタルメディアに蓄積している。また、それを学内利用するだけでなく、学外への発信をも目指している。システムでは、ビデオ情報を MPEG-2 形式で保存し、ネットワーク経由での閲覧機能を提供している。さらに、本学において映像情報を作成するためのスタジオ、編集装置なども本システムに含まれる。

現在、4Mbps ストリームの MPEG-2 ビデオを約600時間格納する容量を有している。

3. 一次情報蓄積システム

デジタル化された冊子体一次情報を蓄積するための大容量ファイルサーバである。最終的に十数 TB の容量を実現することを前提に設計されたため、容量のすべてをディスクアレイで構成するのではなく、光磁気ディスクジュークボックス、磁気テープジュークボックスを用いた階層型の記憶システムによって構成されている。データの配置は、データの利用頻度等によって自動的に行われるように管理されている。

現在、システム全体で 7TB の記憶容量を有している。

4. 検索システム

従来の図書館が提供する二次情報を用いた検索機能及び、OCR によって生成された一次情報を用いた全文検索機能を提供する。現在2機の Web ベースの検索サービスを提供している。

5. ネットワーク接続装置

電子図書館のシステムは、これらの装置単独で動作するわけではなく、これらが有機的に協調することによって全体のサービスを提供している。また、学内の情報システムである曼陀羅システム及び曼陀羅ネットワークとの協調も不可欠である。本システムは、これらシステム間の協調を支援するための高速ネットワークである。現在、スイッチベースの FDDI で構成されている。

6. 業務支援システム

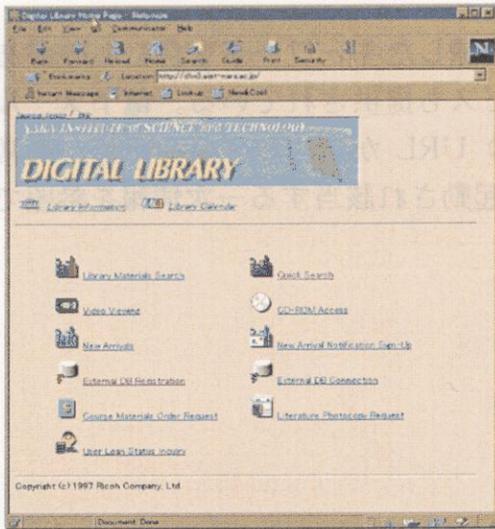
電子図書館といえども、従来型の図書館機能を失っているわけではない。これは、著作権の許諾の関係で電子化できない書籍等に対しての従来型サービスの提供だけでなく、電子化作業の工程管理など電子図書館としても業務を支援する機能を必要としている。これらの機能を提供するため従来型の図書館業務支援システムを拡張する形での実現を行っている。

電子図書館サービス

現行のシステムが提供しているサービスは、基本的に図 4.1.1 に示した通りである。利用者は手元にあるシステムより Web ブラウザを用いて、電子図書館システムにアクセスする。通常の場合、キーワードを入力し全文検索機能で該当する文献を探し出すということを行うことになる。また、MPEG-2 のビデオ閲覧は MPEG-2 のデコード機能を有する端末のみで行うことができるようになっている。

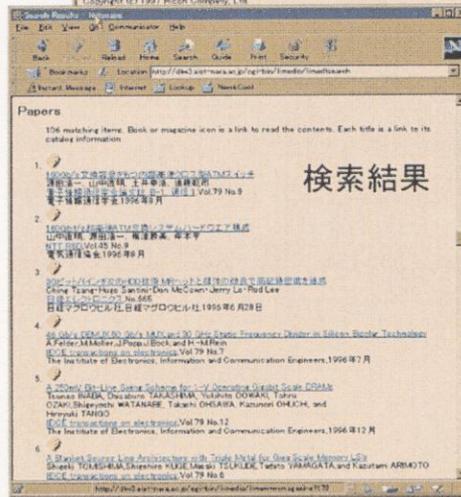
また、雑誌リストから巻号にしたがって、雑誌をブラウズすることも可能である。

これ以外に、新着図書の中から利用者が登録したキーワードを含むものがあつた場合、それを電子メールで知らせるサービスも提供されている。電子メールには文献名とともにそれにアクセスするための URL が含まれているため、通常はこれをクリックするだけで Web ブラウザが起動され該当する一次情報を参照できるようになっている。



電子図書館ホームページ
<http://dlw3.aist-nara.ac.jp>

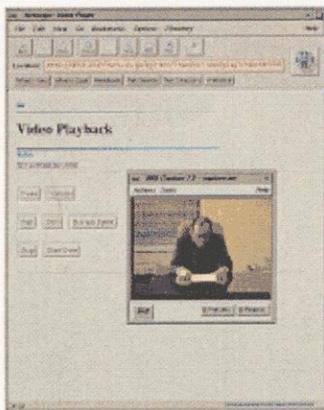
キーワード入力(全文検索)



検索結果



ページブラウジング



MPEG-2ビデオ表示

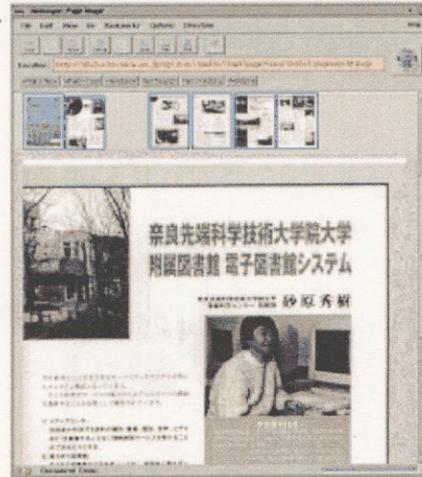


図 4.1.1 Web による電子図書館サービスアクセス例

現行システムの問題点と次期システムでの改善点

4年間の運用経験によって、現行のシステムの問題点が明らかになってきた。これらの問題点の中には、すぐに解決することは困難な問題もあり、研究開発室を中心として情報科学研究科及び情報科学センターにおいて研究課題として問題の解決に努めている。しかし一方において、現行システムの設計見直しによって解決可能な問題もあり、これらは1999年度末に導入される次期システムの設計に反映されている。

次期システムにおいて以下のような改善が行われている。

1. 一次情報蓄積システム構成の見直し

現行システムでは、一次情報蓄積システムは、ディスクアレイ/光磁気ディスクジュクボックス/磁気テープジュクボックスの3階層の構成を採用している。しかし、当初考えていたよりも磁気テープジュクボックスのアクセス速度が高速であることが判明し、3階層の複雑構成を採用するよりも、ディスクアレイ/磁気テープジュクボックスの2階層構成の方が扱いやすく十分な性能が得られることが明らかになってきた。そこで、現行システムの一部からこの構成を採用している。また実際の利用統計によると、利用者がファイルを参照する状況は、個々の利用者によってさまざまであることが明らかになってきている。したがって、頻繁に利用されるファイルを格納するキャッシュであるディスクアレイの容量が当初予想していたものよりも多く必要であることがわかった。当初の設計では、全容量(磁気テープジュクボックスが提供する容量の総計)の10%程度としていたが、現行システムの一部から25%としている。

2. 記事/論文単位でのファイル管理とPDF形式の導入

この問題は階層構成となっている一次情報蓄積システムと密接な関連がある。現行のシステムでは一次情報はページ単位でファイルに格納されている。そのためあるページを参照している状況で、次のページを参照しようとするとは異なるファイルへの参照が発生する。このとき運悪くこのページが磁気テープジュクボックス中に格納されていた場合、そのファイルがディスクアレイにコピーされるまで利用者はそのページの参照を待たされることになる。こうした問題を避けるため、次期システムではPDF形式を採用し記事/論文単位でのファイル管理を行うよう変更されている。これにより、記事/論文が一括して管理されることとなり前述のような問題は解消される。

3. ネットワーククラスタ構成による検索サーバ

電子図書館システムにおいてデータは単調に増加することになる。その結果、同一の能力を持つ検索サーバを用いている限り検索性能は少しずつ低下してゆくことになる。そのため、データの増加に合わせて検索サーバの増強が必要になってくる。しかし、現行のシステムでは、マルチプロセッサ型のサーバ構成を用いており、検索サーバの強化のためには検索サーバ全体の更新が必要となってくるが、このような更新はレンタルによる導入にはそぐわない点がある。現行のシステムにおいては、2機のサーバを用いているが、これらは目的別に利用を分割しており、能力の高いシステムを主に利用者のサービスに用いるようにしている。

こうした状況を解決するために、次期システムにおいてはネットワーククラスタ構成による検索サーバアーキテクチャを採用し、必要に応じてエンジンを追加していくことで全体の検索性能を増強できるようにしている。また、このような構成を採用することでこれまで夜間に行ってきたハウスキーピング処理をクラスタ中の一台のエンジンに割り当てバックグラウンドで行うことができるようになる。これによりハウスキーピング処理による、サービス性能の低下は最小限に留めることができるようになる。

この他、一次情報の入力作業をよりスムーズにするためのシステム構成の見直しや、ネットワーク機能の強化が行われている。この次期システムは、今年度末に導入され試験運用とデータの移行作業を経て来年度より実運用を開始する。