

2. NAIST 電子図書館の状況

2.1 電子図書館システムの概要

1995 年度より整備されている本学の電子図書館システムでは、4 年単位でシステム構築設計を行っており、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いる。2007 年度(2008 年 3 月)に導入されるシステムより、第四世代の NAIST 電子図書館システムが稼働する。ここでは、2007 年度導入の電子図書館システムの構成と利用例を紹介する。

2.1.1 システムの設計方針

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、講義映像および講義資料)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

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、かつ、学外機関が管理する資料も学内資料と同様に情報提供サービスを受けることができる。また、学内で生成される種々の情報の収集・管理を行う。
- **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。

2.1.2 2007 年度導入の電子図書館システム

上記の目標を実現するため、原則としてシステム内に一次情報を蓄積するとともに、電子ジャーナルを提供する国内外の出版機関と連携し、本学の研究教育活動に有益な情報を利用者が検索して利用する機能を中核的な機能として実現している。これに加えて、従来の図書館機能としての情報収集・管理を支援する機能が実現されている。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（基本的に、記事/論文単位の PDF）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 3 台のモノクロ・カラー兼用スキャナ装置とこれらを支援するための PC

群で構成されている。特に OCR 処理には高い処理能力が必要とされるため、5 台の PC で構成されるクラスターで処理される。これらの処理は自動的に 5 台の PC に負荷分散される。さらに、入力作業用のデータ領域として独立した領域を用意している。

また、ここで生成される情報は、通常の図書館での書誌に相当する。これらの情報が消失することは、書籍の紛失を意味し、ファイルのバックアップの体制は十分に確立していることが求められる。そのため、第一次のバックアップとして最終的な形式への変換が完了すると書誌単位で DVD-R へのデータのバックアップを作成することとなっている。

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、学外への発信をも検討している。

映像情報のデータ形式において、主に現在使われているのは MPEG-2 形式と MPEG-4 形式であろう。DVD や地上デジタル放送で採用されている方式が MPEG-2 であるが、高画質高圧縮が可能であることから、MPEG-4 形式の中の H. 264 形式が多く採用されるようになってきている。また、映像情報の表示には Microsoft 社の Windows Media Player (WMP) または Real Networks 社の Real Player が一般的に用いられており、いずれも再生用プログラムは無償で提供されているため容易に再生環境を準備することが可能となる。

このため、本システムでは利用者の多様な閲覧環境 (Windows, Mac OS X, Linux など) を考慮し、映像情報の再生には、Real Player を用いることとし、そのため、映像情報を Real 形式で保存している。同時に、H. 264 形式による高品質映像での保存も行っており、次世代インターネット基盤でのサービス展開も視野に入れている。さらに、本学において映像情報を作成するためのバーチャルスタジオ、編集装置なども本システムに含まれる。これにより、本学からビデオ情報を発信する環境が構成されている。

また、授業アーカイブシステムとして各教室に常設のビデオカメラを設置し自動的にすべての授業を収録しアーカイブするとともに、学内の各部屋にデジタル放送として配信することが可能となっている。授業アーカイブについては、2005 年度より情報科学研究科での試験運用を行っているが、最終的には全

学においてすべての授業の収録を行う予定である。

現在のところ、700GB 以上(2,800 時間以上)の閲覧可能な映像情報が保存されている。このうち、H.264 形式の映像情報は、約 130GB(約 20 時間)保存されている。

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。2007 年度導入により磁気テープジュークボックス装置を中心としたシステムを廃止し、ディスクアレイ装置のみで本システムを構成している。システム全体の総記憶容量は、2005 年度導入されたシステムとあわせて 64TB になる。これにより、本学が所有する一次情報はすべてディスクアレイ装置上に保存されており、電子化された雑誌や授業映像等のビデオ情報などすべての一次情報への高速なアクセスが可能となっている。また、データのバックアップには、本学全体の情報環境基盤である曼陀羅システムのファイルサーバを利用している。

4. 検索システム

本学電子図書館システムでは、従来の図書館が提供する二次情報を用いた検索機能に加えて、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している。一般的に、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。その結果として生じる検索性能の低下を防ぐため、ワークステーションクラス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。現在、検索サーバは 9 台で構成されるワークステーションクラスとなっている。このうち、4 台は検索用インデックス情報生成などのハウスキーピング処理をバックグラウンドで行うようにしている。1 台は学外の電子ジャーナル等の検索用に利用されている。

現行システムでは、検索システムとして全文検索サービスを提供している。ここでは単に文字列を含む文書を提示するだけでなく検索した文字列が該当文書のどこにあるのかを表示する機能を提供している。これは、新しい PDF の機能によって OCR によって認識した文字列を透明なテキストとしてイメージデータの前に貼り付け機能を用いて実現している。これにより、PDF の検索機能を用いて画面上で検索文字列の存在場所を示している。

5. ネットワーク接続装置

電子図書館システムは、各システムが独立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、本学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。本装置は、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、Gigabit Ethernet を基本とするスイッチベースのネットワークで構成されている。

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。

現在、図書の管理において IC タグ(RF-ID, ISO/IEC15693 準拠)を用いたシステムが導入されている。また、学生証・職員証も IC カード(ISO/IEC18092 準拠)化されており、図書館カウンターに設置されている図書用 RF-ID リーダおよび学生証・職員証リーダを用いることで、図書の貸出・返却を利用者自身が行うことも可能となっている。

2.1.3 電子図書館サービス

利用者(NAIST に在籍する学生及び教職員)は、電子図書館の Web ページにアクセスし利用者 ID とパスワード を入力すると、図 1 に示すようなポータル画面に誘導される。この画面は利用者単位でカスタマイズ可能である。具体的には、利用者画面は「ウィジェット」とよばれるモジュール群から構成されているため、モジュールのレイアウトの変更やモジュールの追加・削除が可能となっている。モジュールを用いた画面構成は、2008 年 3 月に導入された新電子図書館システムの新しい機能であり、電子図書館サービスの個人化を実現するためのものである。現在、本学電子図書館システムでは、文献検索、図書館からのお知らせ、RSS リーダ、貸出手続きなどの図書館業務に関連するものなど 20 個の

モジュールが提供されている。

基本的な利用方法は、Google や Yahoo などの検索エンジンと同様であり、検索用モジュールに検索キーワードを入力することによって行う(図 2)。

ここでは、学内に蓄積された情報のみを検索対象にするか、オンラインジャーナルを検索するか、すべてを検索対象にするかを選択することができるようになっている。当然、検索対象を広くするほど検索速度が遅くなるため、必要に応じて選択することが望ましい。結果は図 3 のように表示され、冊子体情報、ビデオ情報、オンラインジャーナルを区別無く表示するようになっている。それぞれの資料に対応するアイコンをクリックすると、それぞれの資料の書誌情報または資料本体が表示されるようになっている。

ビデオ情報は、図 4 のように表示される。現在 NAIST では、ビデオ情報の共有を特に授業アーカイブとして活用している。そのため図に示したようにビデオだけでなく、講義資料を同時に表示するようにしている。講義資料は、ビデオの進行状況に合わせて更新されるようになっているとともに、資料上のキーワードも検索に用いられるようになっている。

冊子体情報は、図 5 に示すように、PDF ファイル表示プログラムによって表示される。表示では、「Geocrawler」という単語がハイライトされているが、これは検索時に用いたキーワードがハイライトされているものである。多少画像上でのずれがあるのは埋め込まれている透明テキストの場所が微妙に異なっているためであるが、実用上支障となるレベルではないと考える。



図 1: NAIST 電子図書館 ポータル画面



図 2: 検索キーワードの入力



図 3: 検索結果

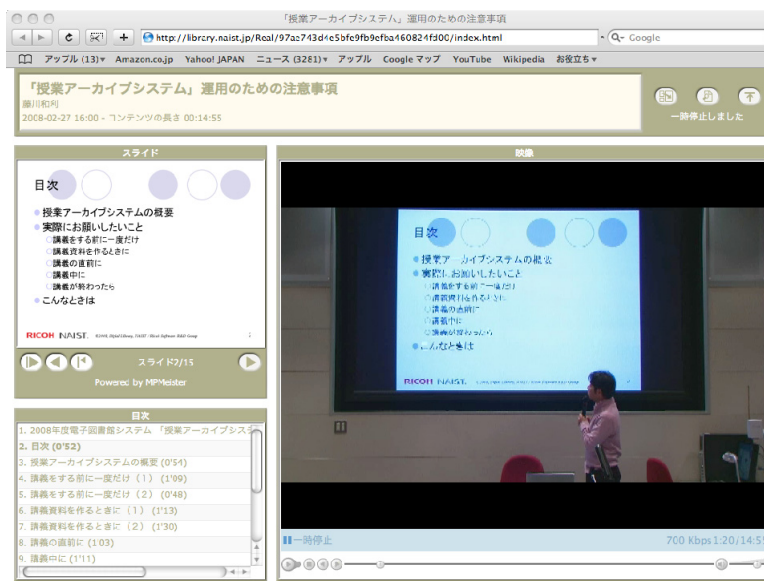


図 4: ビデオ情報の表示

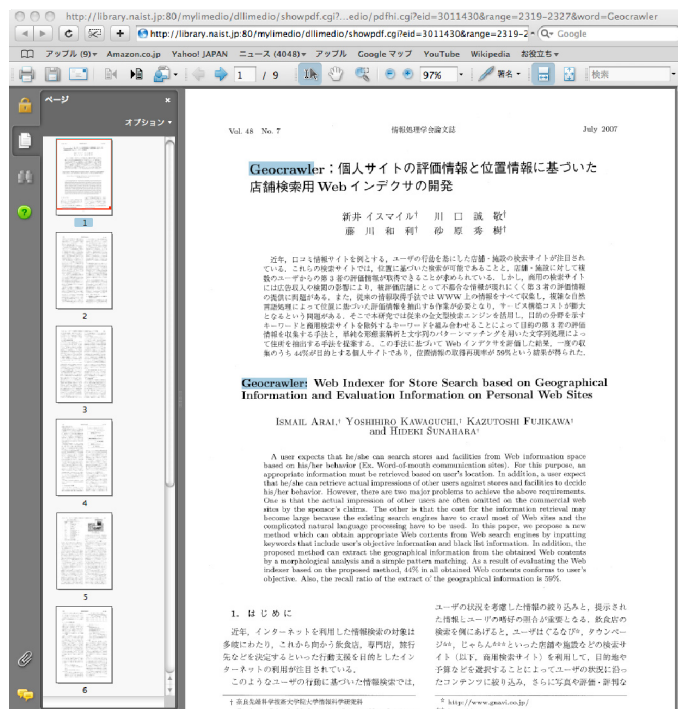


図 5: 冊子体情報の表示

2. 1. 4 次世代電子図書館サービスへ向けて

現在のシステムは、「情報館」構想に基づき新しい機能の導入を行っている。ここでは、次世代の電子図書館システムへ向けての展望について述べる。

(a) 電子書齋/司書サービス

利用者が電子図書館に求めるものは、単純に情報を提供することではなく、どのような情報を見れば良いのかを示唆することである。現行のシステムでも、利用者の興味のある単語を登録してもらい、新着図書からその単語を含むものを抽出し、メールで通知する機能を提供している。しかし、利用者の検索履歴、閲覧履歴、ブックマークなどを用いることで、利用者の興味を抽出しより適切な示唆を与えることが可能であると考えている。電子図書館システム内にユーザ毎の書齋のようなスペース (MyLibrary™) を提供し、利用者の情報を蓄積し適当な情報の示唆を行う電子司書の機能の提供を検討している。

こうした機能の基盤として、ユーザ毎の情報保存スペースを用意し、定期購読や検索・閲覧履歴、ブックマークなどを管理する機能を提供する。これらの技術は近年注目されている Web2.0 と呼ばれる仕組みによって実現することが可能であり、RSS による新着情報の管理や類似利用者の抽出や推薦書誌の提示等の機能を実現するための開発を開始している。

(b) 情報発信機能

電子図書館に求められる機能は、単に情報の閲覧だけではない。論文、テクニカルレポート、テクニカルメモなど、利用者が作成した情報を発信する機能が不可欠であり、これを自由に行えるようにすることでインターネットを経由した情報の広いサーキュレーションを実現することができるようになる。ユーザ毎の書齋スペース(MyLibrary™)では、情報発信のための機能も提供し電子図書館に組み込んでいく。また、従来の書誌形式の情報だけでなく、ソフトウェアや遺伝子配列情報、材料成分等の本学が有する「知」を発信する仕組みとしての活用を前提に柔軟性に富んだ情報管理システムの開発を進めている。

さらに前述の通り大学が有する重要な情報として「講義」を捉え、講義を収録し公開していく機能を用意しているが、これらをポッドキャスト等の技術で配信する実験を開始している。

(c) デジタルコンテンツの扱い

デジタルコンテンツの扱いは、それらを収集し検索エンジンを構成することから、大量のデータからさまざまな関係、利用者の動向を抽出し、より有益な情報を構成することへと移行しつつある。これは、すべてのデジタル情報が同じ共通の基盤の上に格納されているというインターネットの特徴を活用したものであり、今後より大きな役割を果たすと考えられる。爆発的に増殖する大量のデータはもはや一箇所に集約することは不可能であり、電子図書館システムやオンラインジャーナルシステム等が連携し機能することが求められている。そういった意味においても、電子図書館システムの今後の注目がされる。