

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。2009 年度に導入されたシステムより、第四世代目の図書館システム構築が完了した。ここでは、現行システムの設計方針及びその概要を示す。

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 電子図書館システムの現状

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)、Real Networks 社の Real Player または Apple 社の Quick Time Player、Adobe 社の Flash Player が一般的に用いられており、いずれも再生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。

現在、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、標準品質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 2009

年度までは H.264 形式による高品質映像での保存を MOV 形式で行ってきたが、2010 年度からは Web ブラウザとの融合性と検索性を考慮した Flash Video 形式での保存に変更した。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、学内の各部屋にデジタル放送として配信しており、このデジタル放送を利用しすべての授業を自動的に収録し標準品質映像としてアーカイブしている。尚、高品質映像についてはハンディカムカメラにより収録をおこなった後にアーカイブしている。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い、2009 年度にはバイオサイエンス研究科・物質科学研究科も加わり、全学においてすべての授業の収録を行っている。

現在のところ、2.5TB 以上 (6,000 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 920GB (約 420 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等のデータ蓄積およびこれらのバックアップのための大容量ファイルサーバである。2007 年度のシステム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。2009 年度のシステム導入によりシステム全体の総記憶容量は 95TB である。

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、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。なお、高画質映像の保存・配信のために、大容量ディスクを有するハイブリッド型の配信サーバを導入した。また、一次情報等のデータバックアップのために専用ファイルサーバを用意している。これらのファイルサーバの容量が不足した場合、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も利用可能である。

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增強できるようにしている。現在、検索サーバは論理的に9台で構成されるワークステーションクラス構成となっている。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及び文献データベースの検索用に利用されている。2009年度に導入にした高機能サーバ2台を仮想化技術により可用性を保持しつつ4台の検索サーバとして提供し、消費電力と発熱量の削減にも貢献している。

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とシステムの可用性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・複写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、利用者毎の書斎のようなスペース (MyLibrary™) であるポータル画面 (図 1) 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、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。

検索におけ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、複数の文献データベース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たり、文献複写依頼などの文献収集画面へのナビゲート画面が表示される。

現在 NAIST では、授業及び講演の様子をビデオ情報として提供している。図 4 に示すように、ビデオだけでなく資料を同時に表示することによって、授業及び講演内容の充実を図っている。資料は、ビデオの進行状況に合わせて表示されるようになっており、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。資料表示については 2009 年度までは講演者が使用したファイルを利用しているが、動画を表示できない。2010 年度より講演者がプロジェクタに表示している映像をそのまま利用し、高品質映像配信においては、すべての資料表示を可能にした。

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

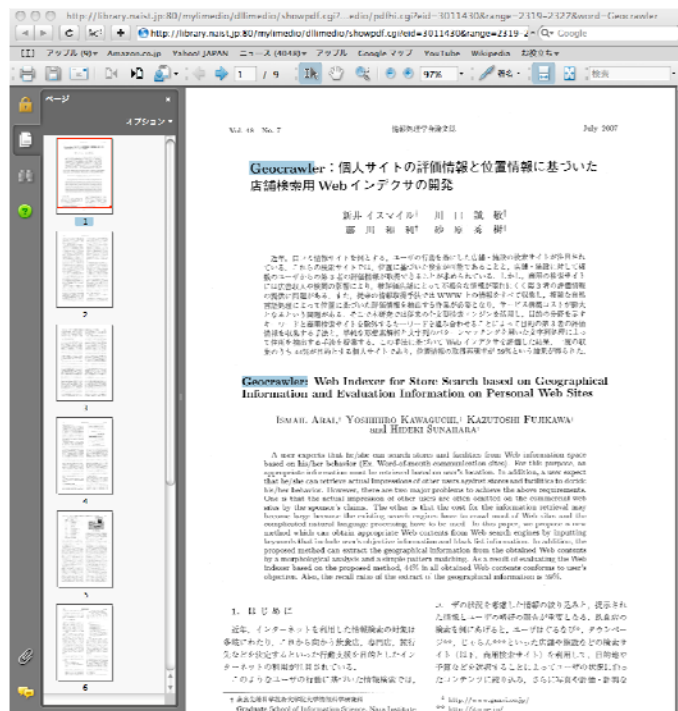


図 5: 冊子体情報の表示

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 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利用者が真に求める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、すべての授業を高品質映像として自動的にアーカイブし配信すること、および本学既設の4Kディスプレイへ高品質映像の講義を効果的に表示することを予定している。

(c) 「NAIST 博物館」構想

本学の電子図書館システムを研究者だけでなく、一般の方を対象にひろくアピールしようという試みである「NAIST 博物館」構想が進められている。電子図書館システムの役割として、単に図書館の延長という枠組みにとらわれず、本学の情報発信局として「NAIST 博物館」のような本学の特徴を十分に伝えるべく、新たなコンテンツ提供、学内の情報基盤整備のほか、スタッフの拡充等の全学的な支援による運用体制ならびに研究開発課題に取り組みたいと考えている。

(d) 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検索の対象となるデータを要約したデータ(メタデータ)を共有するために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される。

(記：辻井 高浩)