

分散型講義サーバシステムの設計

新 麗, 砂原 秀樹

奈良先端科学技術大学院大学 附属図書館研究開発室

〒 630-0101 奈良県生駒市高山町 8916-5

Tel: 0743-72-5103, Fax: 0743-72-5620,

E-mail: {ray,suna}@dl.aist-nara.ac.jp}

概要

奈良先端科学技術大学院大学電子図書館は、次世代の電子図書館の計画と設計を開始した。本年度は電子図書館の内容の充実、他の電子図書館システムとの連携という2点の課題に注目して研究/実装を行っている。その実装の1つとして本学で行なわれる講義、講演等を電子的に蓄積し配信する分散サーバとその分散化を実現するシステム、分散型の講義サーバの設計を行った。

キーワード

電子図書館, 分散型サーバ, メタデータ, 情報作成

The Design of Distributed Lecture Server

Rei S. Atarashi, Hideki Sunahara

Nara Institute of Science and Technology

Research Division, Digital Library

8916-5 Takayama, Ikoma, Nara, 630-0101 Japan

Tel: +81-743-72-5103, Fax: +81-743-72-5620,

E-mail: {ray,suna}@dl.aist-nara.ac.jp}

Abstract

At Nara Institute of Science and Technology, Research Division, Digital Library, we started making the plan and design for Digital Library Next Generation. This year, two points are target topic. One is increasing digital contents, the other is cooperation with other digital library system. As digital contents, we digitalize and archive the university lectures. And we designed distributed server to cooperate with other university lecture archive.

Keyword

Digital Library, Distributed Server, Metadata, Information creation

1. はじめに

情報の電子化技術とインターネットの発達と普及により、情報サーバに関する研究開発、運用はますます重要となってきた。そのなかでも図書館の電子化は特に重要な分野として、様々な試みが行われている。奈良先端科学技術大学院大学 (NAIST) は 1996 年 4 月から電子図書館の運用実験を開始した [1]。現在は書誌情報 60 万ページ、利用アクセス 20 万件/月のシステム [2] となり、様々な成果とまた問題点、課題が見えてきている。3 年の運用実験の間に電子化技術、電子化情報の取り扱い、著作権等、情報の生成、流通、管理のしくみは大きく変化しており、社会の意識や法律までもが改革されようとしている。

我々はこのような変化をふまえ、次世代の電子図書館の計画、設計を開始した。そのなかで、大学の電子図書館の役割として、内容の充実、そして他の電子図書館システムとの連携という観点から、本年度は大学で行われる講義、講演等を電子的に蓄積し配信する講義サーバの分散化に関する研究に重点を置いている。本論文では、本学の次世代電子図書館計画の一部を紹介し、現在設計を行っている分散型講義サーバについて述べる。

2. NAIST の次世代電子図書館計画

我々は NAIST の次世代電子図書館システムの課題は次の 3 点と考えている。

- 内部コンテンツの充実
- 他の情報サーバ (電子図書館) との連携
- より利便性の高いユーザインターフェイスとサービスの提供

本章では、この中で「分散型講義サーバ」に関連する最初の 2 点の課題について解説する。

2.1 内部コンテンツの充実

コンテンツは図書館の特徴を決めるものであり、その充実が図書館の質を決めることになる。冊子体の図書館においては所蔵する蔵書の数と種類であったが、電子図書館では冊子体の電子化情報だけでなく他にも様々なコンテンツが価値を持つ。たとえば、図書館や美術館が所蔵する所蔵品の写真、ビデオ、あるいはある項目への関連情報を集めてあり調査、研究、あるいは検索の指針となるサブジェクトゲートウェイの機能、または特徴ある検索エンジンや表示機能なども電子図書館を特徴づけるコンテンツとなる。

NAIST はこれまで、電子化するコンテンツとして以下のものを対象とし、許諾を得られたものを電子化し蓄積してきている。

- 出版された冊子体情報 (書籍、論文誌等)
- 放映されたビデオ情報
- 学内で執筆された論文 (修士論文、博士論文等)

ここ数年、著作権への意識、課金、認証からなる電子商取引技術の進歩、インターネットの普及により、冊子体情報の電子化は急速に進んでいる。特に最近発表された論文はそのまま電子的に保存できる形式で作成されていることが多く、出版元が電子化で提供すれば印刷物を電子化するよりも効率がよい。出版され

た冊子体情報の電子化は本学電子図書館の大きな特徴の一つであるが、今後は作業に多少の変化が起こると考えられる。

一方で、本学でしか提供できないコンテンツが、学内の研究教育の成果である。本学は大学院のみの大学であり研究教育のレベルも高い。一般教育がなく専門教育のみを行うため、講義内容は最新技術であり充実している。講義は大学の特徴を決める大きな要素であり大学の財産でもある。

これまでは講義の録画、電子化はコストがかかるために実現が困難であった。画像ファイルはデータ量が多いため保存や配信には特別な環境が必要であり、インターネットを通じて公開することは難しいと考えられていた。ここ数年のカメラ技術、画像エンコード技術、データ蓄積技術の発達、またインターネットの高速化により、低コストでの電子化とインターネットでの配信が可能となった。また計算機の表示技術の向上により、ただ動画を再生するだけではなく講義資料と同期させ、実際の授業の再現に近い表示ができるようになった。

本年度は、NAISTでは学内の講義の一部をコンテンツとして取り入れ学内外に提供することにより電子図書館の拡張を図る。目的は次の通りである。

- 教育効果の向上: 学生はいつでも講義を再現でき、全部または一部を見直すことによって理解を深めることができる。また受講していない講義から、必要な部分を取り出し学ぶこともできる。
- 教育研究の公開: 本学の特色ある学内の教育研究活動を公開する。学内外からの関心等を知ることができ、今後の活動の指針のひとつになる。
- 情報流通に関する研究: 本システムを構築、運用するにあたって、録画、電子化、保存、管理、配信、また受講生からの質問の処理まで情報システムに関する研究課題が数多く存在する。講義の録画に留まらず、電子図書館のコンテンツに関する総合的な技術課題に解決を与える。

2.2 外部のサーバとの連携

これまでの電子図書館は、学内の蔵書の電子化し検索、表示する機能が実現されており、システムとして独立して閉じていた。次世代の電子図書館は独立はしてはいるながらも他のサーバとも連携し、利便性の向上を図らなければならない。従来の冊子体の図書館だけでなく、博物館、美術館また他大学なども必要に応じて連携することであらゆる知識、情報を提供する環境が出現する。

現在も各情報サーバがお互いにリンクを作成し合うことで、様々な電子図書館、博物館等を行き来する環境は構築できる。または他のサーバのある情報へのリンクを作成すれば、情報がどのサーバにあるかを意識せずに様々な知識を得ることが可能である。サーバ間の連携はある程度実現されているように見えるが、以下に挙げる問題点がある。

- リンク作成は人手による作業である。連携している相手ごとに保守が必要となる。
- 検索や一覧表示は各サーバで行う必要がある。
- サーバによってインターフェースの差異が大きい。

これらをふまえ、電子図書館/博物館/美術館などの連携モデルを提案する。次世代の情報サーバは独立した個性が重要になる。コンテンツの内容、公開の方針、ユーザーインターフェースなどは各サーバごとに個性があるべきで統一は望ましくない。しかし利用する側からは、検索など一部の機能は統一化され各サーバ上ではなくどこからでも一度に必要な情報を取り出せることが要求される。これを実現するためには標

準化されたメタデータが有効である。蓄積方法、表示方法は異なってもメタデータが統一されていれば、同一の方法で検索を行うことが可能となる。また電子化情報を扱う際に大きな問題となるのは著作権である。著作権はデータの属性となる情報であり、正しく保護するためにはまず統一的に扱う機構が必要となる。これもメタデータにより標準化することで機能が実現できる。よって、次世代の情報サーバはそれぞれ個性があるように見えてメタデータの標準に従うモデルが現実的となる。サーバ間の連携はこの機構の上に成り立つものであり、個性に応じ様々な連携形態が実現される可能性がある。

本論文ではこのモデルに基づき、内部コンテンツの充実と外部サーバとの連携を実現する新しいサーバを設計する。

3. 分散型講義サーバ

我々は、NAISTの次世代電子図書館計画を実現するための新しいサーバとして分散型講義サーバを構築する。講義は大学の電子図書館を特徴づけるには重要なコンテンツである。また電子化してネットワークで接続し他のサーバと連携すれば、利便性が上がるだけでなく研究教育に対してもさらに大きな効果が期待できる。

一方、NAISTは1997年秋からWIDEプロジェクト[3]によるUniversity on the Internet[4][5]プロジェクトに参加している。本プロジェクトは次世代インターネット大学環境を構築するためのシステムを構築し様々な研究、開発、実験を行っている。NAISTでも許諾を得た一部の講義を本システムの講義として登録し、インターネット上の受講生が聴講するための実験を行ってきた。その実験に対する反響からNAISTの講義に対する学外からの関心の高さ、あるいは情報としての貴重性が確認されており、まずコンテンツとしての有用性が評価されている。この実験では1台の講義サーバにすべてのコンテンツを登録し配信していたが、これには、2つの大きな問題がある。1つは計算機、ネットワークの負荷が講義数、受講生数の増加と共に上がることである。もう1つは1台のサーバで一括管理を行う参加組織に対し制約が大きく特徴が出しにくくなり、結果として参加するための準備にかかるコストが上がり新しい組織が参加しにくいということである。この2点の問題は、参加組織ごとにサーバ持つように分散して標準に従って構築し、情報を交換するしくみを構築することによって解決できる。

以上の環境の上で他サーバと連携する分散型の講義サーバの設計を行う。

4. 設計

分散型講義サーバシステムは、講義を電子化しサーバに保存する情報作成システム、受講生の要求に応じ講義を配信する情報配信システム、講義や受講生、または分散されたサーバの管理を行う管理システムの3つで構成される。本章ではそれぞれのシステムについて説明する。

4.1 情報作成システム

講義を録画して電子化し、講義資料と同期したコンテンツとし、メタデータを付与して保存するシステムである。録画は民生用のカメラで行い、オペレータが操作する。電子化のフォーマットはDV, MPEG, Realvideoの3種類を作成し、要求に応じて切り替えられるようにする。これは受講生が持つ様々なネットワークへの接続環境を考慮したものであり、高速な環境では高画質の画面を、低速な環境では画質は落ちるが講義としては成立するような画面を送ることにより、その時に最適な環境で受講できるようにする仕組みである。

次に、講義のメタデータとして必須となる項目は以下である。

- 講義タイトル
- 講演者
- 講義の日付、期間
- 関連情報
- 著作権

本システムでは Dublin Core[6] を基礎としたメタデータを定義する。

4.2 情報配信システム

受講生の要求に応じて講義の録画データ、講義資料を配信するシステムである。受講には Web ブラウザを使用する。サーバは Web サーバと Real サーバ等のビデオサーバから構成され、受講生のネットワーク環境に合わせたデータを配信する。

分散するのはこのシステムであり、実現には 2 つのコンポーネントが必要である。

- データの分散: サーバの負荷を軽減するために複数のサーバが同じデータを持つ。サーバと受講生の物理的、ネットワーク的な距離とサーバの負荷状況に応じて最適なサーバを選択する。このコンポーネントは他のサーバのデータとの同一性を保つ機構が必要である。
- サーバの分散: 各組織ごとに特徴に合わせたサーバを構築する。検索、一覧表示などは同じインターフェースで行えるような機構が必要である。

4.3 管理システム

講義に関するデータと、受講生に関する情報を管理する。講義データは著作権情報を持ち、公開する範囲を規定してサーバの動作を制御する。現在のところはコピー防止の機構は持たない。受講生に関する情報はプライバシーに関わるものであり、アクセスは本人のみに限られる必要がある。本システムでは本人の認証システムを取り入れることによってアクセス制限を実現する。

5. 今後の計画

本設計に基づき、システムを実装し評価する。特に分散化は初めての試みであり、従来のリンク情報よりも密であるがクラスタサーバのように一括管理ではない、ゆるやかな結合状態にあるシステムのモデルはまだ確立されていない。本実装を通じて、メタデータの共通化によるサーバの連携の有効性を評価すると共に、本システムをモデル化し提案する。また本学の電子図書館に組み込むことで、次世代の大学電子図書館の 1 つのモデルケースを実現する。

6. まとめ

NAISTの電子図書館は3年間の運用を通じて、電子図書館の1つのありかたを提案してきた。日本国内、海外ともに電子図書館化の試みは盛んであり、どれも特徴あるシステムを提供している。これまでは蔵書あるいは所蔵品、またはそのメタデータを電子化する、図書館の中に目を向けた実装に重点が置かれてきたが、次世代の電子図書館は外に目を向け他の電子図書館あるいは情報サーバと連携することが重要となる。本論文では他サーバとの連携を実現する1つのモデルとして分散型講義サーバを設計した。様々な方針で設計されているサーバを連携させるには、システムのモデル化とある程度の標準化が必要となる。今後は本システムの実装を行い、運用を通じて分散型サーバのモデル化と評価を行う。

謝辞

本研究を行うにあたって様々な助言をいただいた WIDE プロジェクトの大川恵子氏をはじめ School of Internet グループの皆様と、奈良先端科学技術大学院大学附属図書館研究開発室の羽田久一氏に感謝いたします。

参考文献

- [1] Rei Suzuki, Hideki Sunahara, Masakazu Imai, Kunihiro Chihara: Building Digital Library System - NAIST Challenge, International Symposium on Research, Development and Practice in Digital Libraries 1997, ISDL'97, 1997.
- [2] 今井 正和, 新 麗, 羽田 久一, 西村 亨, 砂原 秀樹, 千原 國宏: 「ある電子図書館の運用と統計」, 第13回デジタル図書館ワークショップ, pp23-32, 1998.
- [3] <http://www.wide.ad.jp/>
- [4] Keiko Okawa, Jun Murai: "School of the Internet: A University on the Internet", In Proceedings Internet Global Summit, INET98, 1988.
- [5] <http://www.wide.ad.jp/soi/>
- [6] "Description of Dublin Core Elements", <http://purl.org/metadata/>