

HTML 文書を活用した XML 技術によるセンターWeb サイトの構築

Development of Center Web Site

with XML Technology Utilizing HTML Documents

松田 勝敬[†], 南木 徹[‡], 西村 亨[†], 藤原 克彦[†], 永井 明[†]

Masahiro MATSUDA, Tooru NANMOKU, Tooru NISHIMURA, Katsuhiko FUJIWARA and Akira NAGAI

[†] 宇都宮大学総合情報処理センター

Advanced Media Network Center, Utsunomiya University

〒321-8585 栃木県宇都宮市陽東 7-1-2

7-1-2 Yoto Utsunomiya Tochigi 3218585 Japan

E-Mail: {mmatsuda, nisimura, fjwr, anagai}@cc.utsunomiya-u.ac.jp

Tel: +81-28-689-6340

[‡] マジケ合資会社

Magiche

〒320-0848 栃木県宇都宮市幸町 9-17

9-17 Saiwai Utsunomiya Tochigi 3200848 Japan

E-mail: nanmoku@magiche.com

Tel: +81-28-684-5856

概要

インターネットを利用して情報を公開する方法として、WWW は最も重要かつ有効な手段である。Web ページを記述する言語としては、現在では HTML が多用されている。一方、より柔軟性の高い記述方式を提供する言語として XML が開発され、それに基づいて拡張された XHTML や XML 自体も Web サイトの記述に用いられるようになった。XML を用いて Web サイトを構築することにより、拡張性と柔軟性が向上し種々の情報を効率よく処理し、公開することが可能となる。筆者らは、まず総合情報処理センターの既存 HTML データを活用し、XML によるサイトの開発、構築をおこなった。ここでは HTML で記述されていた Web サイトのデータを活用した XML 技術による Web サイトの構成に関して、運用とセキュリティの面からの考察及び構築について述べる。

キーワード: WWW, HTML, XML, Web サイト

Abstract

World Wide Web is a most important and effective means for presenting and retreating information on the Internet today. Pages on most Web sites are still described in Hyper Text Markup Language. Some newer sites seem to begin utilizing XHTML, developed as an extension to the HTML using XML technology, or native XML, extensible markup language for site description. By describing a web site using XML, representation will become more extensible and flexible and the data processing will be performed more effectively. We developed an XML based web site using existing previous site data written in HTML. This paper discusses how we performed the transition from HTML based to XML based site, from administration and security.

Keywords: WWW, HTML, XML, Web Site

1.はじめに

現在、大学の情報系センターでは基幹ネットワークシステムを管理、運用する機能が重要な位置を占めている。コンピュータネットワークはセンター内や大学内の情報交換だけでなく、インターネットをはじめとした大学外との情報交換手段として一般的に用いられるようになってきている。特にインターネットの急速な普及により、Web サイトなどによる大学内の情報発信手段は必要不可欠なものとなっている。一般世帯へのインターネット環境の普及率もセンターや大学などの組織の情報を得るときに、まずその組織に関して Web 検索サイトから検索しホームページを見ることが一般的になっている[1]。また、ネットワーク帯域の広帯域化によりネットワークで流通する情報も、様々な分野の多種多様な形態をとるようになってきている。

ネットワーク上の情報を記述するマークアップ言語として、1998年 W3C(World Wide Web Consortium)から XML(Extensible Markup Language)が勧告された[2]。XML は様々な構造化言語をつくることができ、これまで別個の形態で作成、保存されていた情報を XML 文書として同じ規定の上で作成、保存することができる。XML は様々な分野で利用されており、XML 文書として同じ規格に則って情報が蓄積されている。XML 技術を用いることにより、これらの異なる分野の種々の情報を同じ処理システムで扱うことが可能である。様々な情報を XML 文書という規定で蓄積することにより、これまで難しかった種々多様な分野、形式の情報を横断的に扱うことが可能となる。最近ではアプリケーションの XML への対応が進み、XML 文書で情報を直接出力することができるものも現れているが、これまで蓄積されている情報のほとんどは XML ではない形式である。そこで情報の XML への対応技術は、XML の普及に従い需要が増えると考えられる。

我々は、総合情報処理センターでは大学のネットワークやサーバ等の基幹システムの運用や管理だけでなく、情報通信に関する教育や研究を積極的におこなうことが必要であると考え、通信に関連する先端技術の教材や実用的なシステムの開発を行っている。特に我々は次世代の Web 技術として XML に注目している。情報リテラシー教育教材として XML とその関連技術に関する教材を開発し[3]、実際に情報リテラシー授業で大学一年生を対象に授業で利用している。また、XML データベースシステムについて検討[4]をおこなうなど、XML の特性を生かした利用技術の研究

をおこなっている。センターでは情報公開手段として、ネットワークや計算機のシステムの稼動状況や利用方法、各種申請書のフォーマットなどを Web サイトで公開しており、色々な情報が集約されている[5]。

宇都宮大学総合情報処理センターは 2001 年 4 月に「情報処理センター」から「総合情報処理センター」に改組拡充された。当時は情報処理センター時代からの主に HTML を用いた Web サイトを運用していた。総合化に伴い XML 化を念頭に置いた Web サイトの改変を行い、2001 年 8 月に Web サイトのデザインと構成の一新を図るとともに Web ページの記述を HTML から XHTML に変更した。公開内容も携帯電話を主な対象としたモバイル端末向けのサービス、ストリーミングによる動画配信なども新たに追加し、公開コンテンツの種類も増加するとともに Web サイトの構成も複雑化している。さらに積極的に Web サイトに新技術の導入をすべく、既存のセンター Web サイトを生かしつつ XML 技術を用いた公開方法を実装した。そこで、センター Web サイトに関して XML 技術の適用を行い、特に既存の HTML(Hyper Text Markup Language)で記述されている情報の XML 技術による公開への移行について検討、考察を行い実装を行った。

2.センターWebサイトの概要

2.1 設計方針

WWW(World Wide Web)による Web サイト閲覧は、インターネットの急速な普及とともに情報収集手段として広く用いられている。大学の中には、大学のホームページを始め様々な Web サイトが公開されている。一般的に Web サイトを公開する一番の目的は、情報を広く組織内外に公開することである。Web サイトを閲覧する側も、ある組織などに興味を持った人がまず初めに情報を得るために、その組織の Web サイトを確認することも多いであろう。

Web サイトは WWW サーバとネットワークを稼動していれば、管理者の勤務時間や勤務日などに係らず 24 時間 365 日常に閲覧することができる。このような安定した Web サイトの運用のためには、サーバの機能を必要最低限のものとした簡潔なサーバ構成による、サーバ停止の防止と不具合発生時に対する対応の容易さなどが求められる。また簡潔なサーバ構成はセキュリティ性を向上させる上でも有利である。

大学のホームページは受験生から共同研究を行いたい企業などの研究者、公開講座や講演会などに参加を希望する方々など、様々な人が閲覧する Web サイト

である。国内に限らず、留学を希望する学生や海外の研究者など国外からのアクセスも考えられる。そのために、大学のホームページは様々な閲覧者の環境に対応した Web サイトである必要性が重要である。様々な種類やバージョンの WWW ブラウザからのアクセスや、狭帯域ネットワークからのアクセスなど、大学の内外の様々な環境からもページを表示でき、確実に情報を伝達することが求められる。さらに、閲覧する側からも更新などのメンテナンスをおこなう側からも、Web ページの構成が複雑ではなく単純で理解性が高いことが要求される。

一方我々は総合情報処理センターの Web サイトに関しては、大学のホームページに要求されることに加え、先端的な技術や広く一般的には用いられていない有用な技術などを Web サイトに実装し、実際に運用しながら検証をおこなえる実験的な Web サイトとしての一面も持つと考えている。そこで、次の設計方針をもとに Web サイトを構築している。

- ・安定した稼働。
- ・先進的な技術や有用な技術の積極的な実装。
- ・簡潔なサーバ構成。

実験的な運用は安定稼働と相反することになるので、安定した運用ができ様々な環境からのアクセスに対応できる部分と、対象を限定してもかまわないが様々な機能を実装する部分に分けた構成とした。さらに、同じ Web サイトの内容を持つミラーサーバを用意し、サーバ自体の不具合にも備えている。また、データベースなどを用いるサービス、広帯域ストリーミングサービスなどは、別の専用サーバを使い Web サーバの

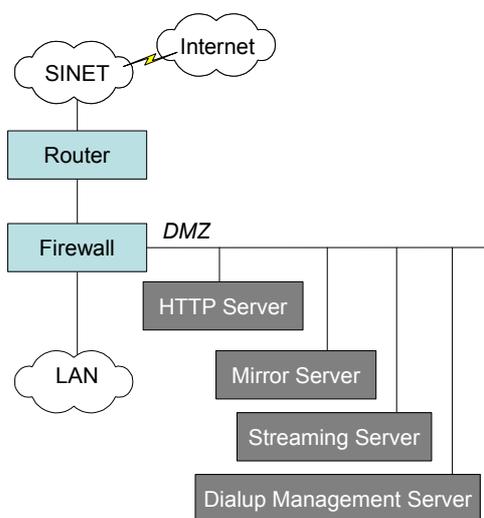


図-1 ネットワーク構成概略

構成の複雑化を避け安定性を向上させるだけでなく、セキュリティ性、整備の容易さも考慮している。

2.2 Web サーバの構成

Web サーバは PC サーバであり、HTTP サーバは Apache[6]を用いている。管理用の Web 運用統計や、2.3 で述べる各ページの生成などで利用するため、Perl[7]を実装している。Web サーバで稼働させるサービスは最低限のものだけ実装し、システムを単純化して安定稼働と整備の容易さ、セキュリティ性の向上を図っている。

ネットワーク構成の概略を図-1 に示す。宇都宮大学は学術情報ネットワーク(SINET)を介してインターネットに接続している。SINET とは筑波大学学術情報処理センターノードに 100Mbps の帯域で接続している。総合情報処理センターの Web サイトの World Wide Web(WWW)サーバは、大学ネットワークの DMZ 上に設置されている。

主なコンテンツは HTTP サーバ で公開しているが、DMZ 上には他にも動画配信用ストリーミングサーバ、RFC やソフトウェアの配布及び教育用コンテンツのミラーサーバ、ダイヤルアップ接続管理サーバなどがあり、センターWeb サイトのコンテンツの一部を成している (図-2)。

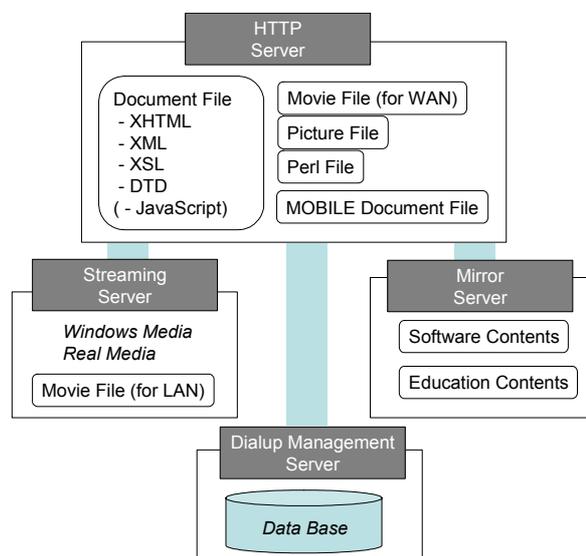


図-2 サーバ構成

2.3 Web サイトの構成

Web サイトの構成を図-3 に示す。Web サイトは大き

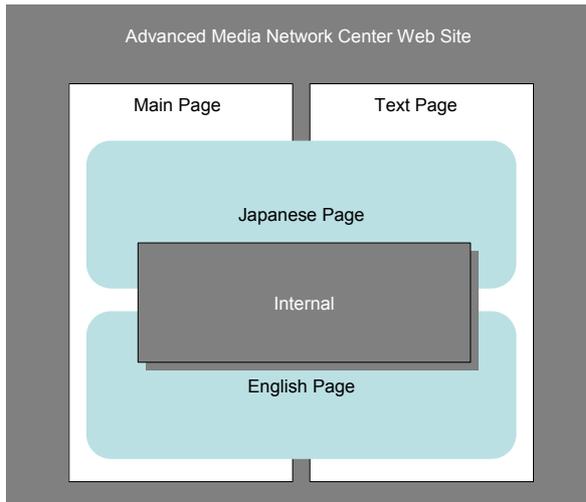


図-3 Web サイト構成

くグラフィカルな「基本ページ」(図-4)と、テキストベースの「テキストページ」(図-5)に分かれている。「テキストページ」はテキスト WWW ブラウザや狭帯域、古い WWW ブラウザなどの閲覧者の環境でも情報提供が可能にするために公開している。「テキストページ」は HTML で記述されている。「テキストページ」を基に「基本ページ」を作成しており、グラフィカルなアイコンやボタンで構成される。「基本ページ」は XML 導入の一環として XHTML での記述をしているが、「テキストページ」を基にした配置変更やアイコンやボタンの追加などの構成の変更を施したのみであり、基本的に HTML による Web サイトの技



図-4 基本ページ

宇都宮大学総合情報処理センター(説明)

宇都宮大学ホーム | English | リンク版 (XHTML 1.0) | XML 試験版 (MSXML 専用)

お知らせ

センターについて

[センター概要 \(PDF詳細版\)](#) | [センター沿革](#) | [関係者紹介](#) | [センターカレンダー](#) | [センター授業時間表 \(学内\)](#)

サービス

[サーバリンク](#) | [マルチメディア編集機器](#) | [ダイヤルアップ接続 \(学内\)](#) | [ストリーミング配信実験 | 互換対応コンテンツ](#)

アーカイブ

[センターニュース](#) | [各種申請書 \(学内\)](#)

学内専用メニュー

[学内専用メニュー入口](#)

サイト内検索

[詳細検索](#) | [検索方法](#)

図-5 テキストページ

術で構築されている。「基本ページ」,「テキストページ」とともにほとんどの公開情報は移行前と同様のものとなっている。

「基本ページ」,「テキストページ」とともに日本語ページと英語ページがある。さらに大学の内外から閲覧することができるページと大学内のネットワークからのみ閲覧を許可しているページに分けてある。トップページは学内外から閲覧可能であるが、メニューによるリンク先で学内外と学内専用ページを分けている。学内専用ページは IP アドレスによってアクセス制限をしている。

学内専用ページでは、総合情報処理センターへの申請書など学内でしか需要がないコンテンツや詳細な利用方法など学外への公開に向かないコンテンツの公開、配布を行っている。学内の主なネットワークは、100Base 及び 1000Base による構成であり、その広帯域性を有効に活用するためのコンテンツも制作し、公開している。同じコンテンツを学内外に配信する際には、複数のビットレートのコンテンツを作成し、環境に応じた利用を可能としている。主に学内専用コンテンツでは 1Mbps, 学外用では 256Kbps, モバイル用は 64Kbps と 32Kbps のビットレートで配信をしている。

そのほかセンターからのお知らせを簡単に入力と自動更新ができる、NewsML[8]を利用した WebUI のシステムや、主に携帯電話向けのモバイル対応ページなども実装している[5]。

3.XML 文書を利用した Web ページ公開

XML はシンプルな文法を持ち、そして柔軟な記述を提供するため広範囲な応用が進められ、そのため多様な関連技術が開発されている。XML で記述された文書情報をソースとして Web サイトで公開する

様々な技術が開発されており、またそれらの技術を組み合わせることにより様々な応用が可能になる。XML は情報を記述する言語であり、XHTML などを除き直接 XML 文書の WWW ブラウザなどによる表示には向いていない。一般的に WWW ブラウザで XML 文書の内容を閲覧するには、元の XML 文書を XSLT など XHTML など WWW ブラウザでの閲覧に適した構造へ変換する。公開する情報の実体が記述されている XML 文書と、表示するための構造変換を記述した XSL 文書から WWW ブラウザで表示するための変換をおこなう技術を実装する必要がある。この変換をサーバでおこなうかクライアントでおこなうか、またどのような技術を用いるかによって XML 文書の公開方法が異なってくる。実際にセンター Web サイトとして XML 文書を利用した Web ページを運用するには、既存の Web サイトも並行して運用することになり、それぞれの Web サイトと関連した運用の容易さも考慮する必要がある。

ここではセンターの Web サイトとして実装し運用するにあたり、

- HTML で記述されている既存の情報の利用
- 簡潔なサーバ構成
- 拡張性
- クライアントの制約
- 既存の Web サイトとの並行運用

などの点から次の 5 つの方法について検討をおこなった。

3.1 XHTML

HTML を XML として再編成したものが XHTML[9] である。そのために、ほとんどの HTML で記述されている Web ページは、容易に XHTML で記述することができる。サーバ、クライアントのどちらも HTML による Web ページと同様な構成で、XHTML に対応することができる (図-6)。

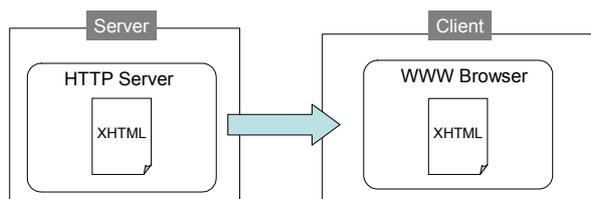


図-6 XHTML による実装

HTML で記述されていた文書を、XHTML に変更することで XML 文書となるため、XML の技術を適用

することが可能となる。3.2 から 3.5 における XML 文書は、HTML 文書を XHTML で記述し直した XML 文書を扱っている。HTML 文書の XML 化は容易におこなうことができ、Web サイトのサーバ構成も複雑ではない。しかし、新たに XML 文書を加える時にも XHTML で記述する必要がある、HTML で記述した文書を加えるのと大差なく、Web サイトとしては、XML 化による利点は少ない。

XHTML は HTML から容易に移行できるため、既存の HTML 文書を XHTML 文書に対応させたり、新たに加えるページを XHTML に準拠して記述することは、XML 技術を用いた Web サイトへの移行をおこなう上では重要である。

3.2 クライアント側 XSLT 処理

XML 文書は XSL[10], XSLT[10]を用いることにより、文書構造を変換することができる。そのままでは WWW ブラウザでの表示には向いていない XML 文書を HTML や XHTML などの WWW ブラウザでの表示に向いている文書に変換をする。元となる情報を記述している XML 文書で XSL ファイルを指定することにより、クライアントの WWW ブラウザで文書の変換をする (図-7)。HTML ファイルから変更した XHTML ファイルに、適用する XSL ファイルを記述するタグを追加するだけでよく、既存ファイルを比較的簡単に利用できる。XSL ファイルを用意すれば、サーバ構成も HTTP サーバだけでよく追加するサービスもない。

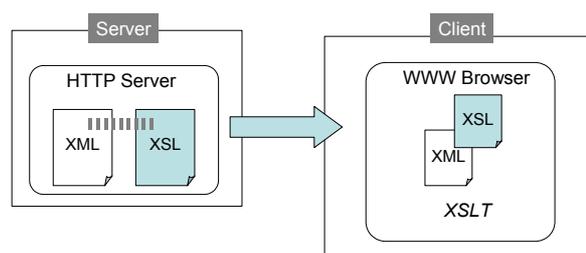


図-7 クライアント側による処理

この方法では、XHTML 文書で指定した XSL ファイルしか適用することができないので、XML の特長である柔軟性はあまり生かせない。同じ XML 文書の内容を異なる XSL ファイルで変換するときには、XSL ファイルの指定だけを変更した同様の XML ファイルが必要である。また、クライアントの WWW ブラウザが、XML 文書を扱うことができ XSLT に対応していないと期待するようにページを表示することが

できない。

3.3 JavaScript によるクライアント側処理

XHTML 文書単体でも XSLT 処理はおこなえるが、適用する XSL ファイルを変更することは難しい。そこで、JavaScript[11] で変換に用いる XML 文書と XSL ファイルを指定することで、柔軟に XML 文書を扱うことが可能である。

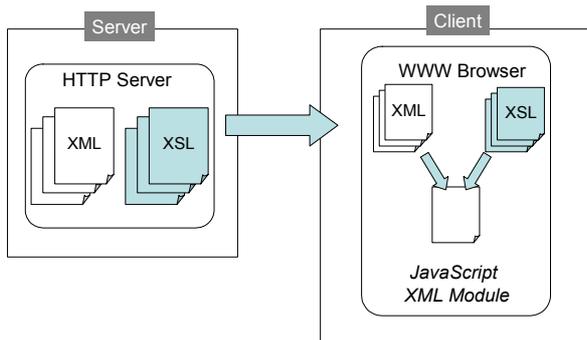


図-8 JavaScript による実装

XSLT の処理をクライアント側の WWW ブラウザでおこなうので、サーバ側には HTTP サーバが実装してあればよい(図-8)。XHTML 文書に読み込ませる XML 文書と XSL ファイルを指定しておき、それらを JavaScript で XML データアイランドに読み込ませ変換処理をおこなう。XML 文書と XSL ファイルの指定は、任意に組み合わせを指定できる。また、JavaScript を用いるので、ボタンなどに対応して XML 文書を変換し出力させるなど、非常に柔軟性と拡張性がある Web サイトを作ることができる。既存の Web サイトの XHTML 文書を指定することにより、並行運用も可能である。新しい XML 文書も簡単に追加することが可能であり、追加する文書は XHTML ではない XML 文書でも構わない。

この方法ではクライアントの WWW ブラウザで変換処理をおこなうので、クライアントの環境が JavaScript が使えるだけでなく、XML 文書が扱え XSLT に対応している WWW ブラウザに限定されてしまう。また、読み込む文書の量によっては、クライアント側のネットワーク帯域や処理能力が貧弱な環境では Web ページの表示が円滑に処理されない場合が起こり得る。

3.4 サーバ側処理

コードがサーバで実行されるスクリプト言語などを用いて、サーバ側で XML 文書と XSL から XHTML 文書に変換する方法である(図-9)。クライアントが読み込んだスクリプトに応じて、サーバにある XML 文書ファイルと XSL ファイルから XSLT プロセッサが HTML(XHTML)に変換する。

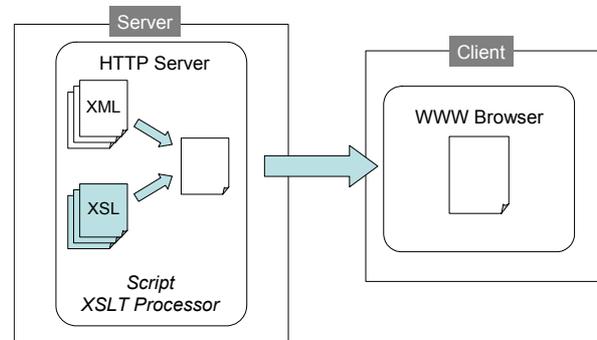


図-9 サーバ側処理による実装

スクリプトとして比較的扱いやすく実装も簡単な、PHP(PHP: Hypertext Preprocessor)[12]をサーバスクリプトとして用いた。XSLT プロセッサは Ginger Alliance の Sablotron[13] を実装した。HTTP サーバは Apache を用いた。

サーバ側で変換処理をおこなうためクライアント側の制約は少なく、様々なクライアントの環境に対応することが容易である。また、XML 文書は既存の Web ページの XHTML ファイルを用いることができ、XSL ファイルを用意しておけば XML 文書の長所を生かした Web サイト構成にすることが可能である。サーバ構成は、HTTP サーバの他にスクリプト言語と XSLT プロセッサを実装すればよく、比較的簡単なサーバ構成になる。この方法では、PHP スクリプトを記述したファイルにしたがって XML 文書の処理をおこなうため、既存のファイルの他にスクリプトを記述したファイルを作成する必要がある。

3.5 ネイティブ XML データベースの利用

ネイティブ XML データベースとは、XML 文書を XML ツリー構造を保持したまま格納可能なデータベースである。現在多く用いられている RDB(Relational Database)でも XML 文書に対応したものはあるが、XML 文書を分解しマッピングする必要があり、格納する XML 文書の構造をなどのスキーマの定義が必要となる。これは XML 文書の特徴であるツリー構造と、表の関係であるリレーショナルデータモデルでは情報

の構造が異なるからである。ネイティブ XML データベースでは、XML 文書をつリー構造のまま格納できるため、スキーマがない整形形式の XML 文書でも格納することができる。

XML のデータベースに関連する仕様の策定が進むにつれ、ネイティブ XML データベースも様々なものが発表されている。我々はこれまでに Apache Xindice[14] を用いて XML データベースシステムの研究・開発を行っている[4]。ここでは、サーバの構造が比較的複雑にならない図-10 の構成を採用した。HTTP サーバは Apache を用い、サーブレットは同じく The Apache Software Foundation の Apache Tomcat[15] を用いている。ネイティブ XML データベースは Tomcat と同様に Java で実装されている Apache Xindice を用いた。

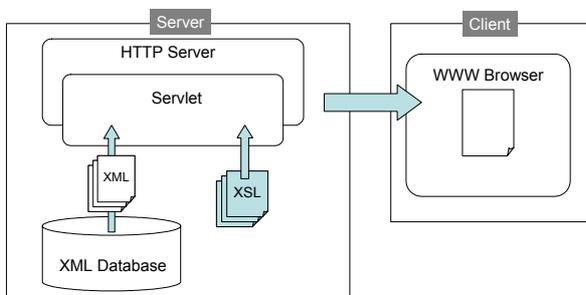


図-10 XML データベースによる実装

このシステムでは、ネイティブ XML データベースに Web サイトで公開するための情報を持つ XML 文書を格納しておき、クライアントからの要求にしたがって対応するファイルを出力する。データベースからの XML 文書と XSL ファイルから XHTML 文書を生成し、HTTP サーバがクライアントの WWW ブラウザに送信する。データベースへの要求や出力された XML 文書と XSL ファイルからの変換などは、サーブレットがおこなう。既存の Web サイトとの更新内容などの一致を図るには、定期的にデータベースの文書を既存の Web サイトの対応する XHTML 文書と比較、更新をおこなうことで実現できる。

本システムの最大の特徴としては、データベースに登録する情報を拡張することにより、単に Web サイトとしての機能だけでなくより高度な WebUI を持つサービスを実装することができる点である。例えば、センターの利用者などの情報と組み合わせ利用者個別に異なる情報を提供する Web サイトや、SOAP[16]の実装により Web サービスと組み合わせた Web サイトへの発展も考えられる。サーバ側で変換処理をおこなうため、柔軟に種々のクライアントの環境に対応する

ことができる。

このシステムでは、HTTP Server のほかに Servlet や Database、さらにそれらを稼働させるために Java をサーバに実装する必要がある。また、Web サーバ自体にデータベースが実装されており、セキュリティ向上を考えデータベースを他のサーバで稼働させるとさらに複雑なサーバ構成になってしまう。

3.6 考察

3.1 のような既存の HTML 文書の XHTML 文書への変換や XHTML による Web ページの記述は、既存の Web サイトからの移行の際には、非常に重要である。今回の既存のセンター Web サイトの「基本ページ」は、XHTML で記述されていたため、既存 Web サイトと XML 技術を用いた Web サイトの並行運用を可能にしている。

3.1 や 3.2 の方法は、XML 文書を Web サイトで扱うには拡張性が乏しく、これらの技術の実装だけでは XML の特長を活かした実用的な Web サイトの構成は難しい。そこで、スクリプトなどによる何らかの動的な文書操作などを実装する必要がある。

3.3 から 3.5 の方法では、スクリプトなどの処理系を実装している。3.3 の方法では柔軟で拡張性のある実装であるが、サーバ側には機能を追加して実装する必要が無いことが特徴である。3.4 や 3.5 では、サーバ側に様々な機能を実装することにより、より高度な Web サーバとして拡張が可能である。しかし、サーバ構成が複雑になってしまう。

サーバ技術に関する検討を経て実装へ移行する以前に、センター Web サイトの利用者の環境を考察する必要がある。現在、XML ベースの Web ネットワークシステムを利用する方式は、大別して、XSLT の処理をクライアント側でおこなう方式(3.2, 3.3)とサーバ側でおこなう方式(3.4, 3.5)に分類することができる。クライアント側で処理をおこなうと、クライアント側で使用する WWW ブラウザやネットワーク帯域、コンピュータの処理能力などに影響を受けてしまう。しかし、XML への対応が比較的進んでいる Internet Explorer 6.x が WWW ブラウザのシェアの 7 割近くを占めている[17]こと、他大学など主な Web 公開対象は SINET に接続が多い、一般家庭でインターネット接続の約半数が 64kbps である[1]こと、最近のコンピュータの高速化とインターネット利用パソコンは半数以上が過去 2 年以内の購入である[17]ことから、クライアント側での処理をおこなっても大きな問題にはならないと考えられる。また、クライアント側で処理をおこなうため、サーバ側に機能を追加する必要がないので

移行も容易である。セキュリティの面からも従来の対策で対応できる。

これらのことから、3.3の方式における Web サイトをセンターWeb サイトに実装をおこなった。

4. センターWeb サイトへの実装

Web サーバ、Web サイトの構成にほぼ変更を加える必要がなく XML サイトを実装できる方法として、JavaScript を用いた XML サイトの実装を採用し実装した (図-11)。

クライアントの様々な環境への対応と安定稼働など



図-11 センターWeb サイトへの JavaScript による実装

を考慮して、トップページからリンクをつくりこれまでの Web サイトと並行して運用できるように実装した (図-12)。WWW ブラウザの JavaScript や XML 処理の対応から利用できる WWW ブラウザ (Microsoft Internet Explorer version 6 以降) が制限されるため、対応していない WWW ブラウザを利用していると警告を表示するようにした。JavaScript により XML 文書と XSL 文書からページを生成するが、XML 文書は「基本ページ」の XHTML 文書である (図-13)。「XML ページ」では「基本ページ」の対応する XHTML 文書を利用しており、どちらのページでも同様な内容が表示される。Web ページに変更、修正を加える時、基本構成を大幅に変えるような変更でなければ、「基本ページ」の XHTML 文書を書き換えるだけで対応する「XML ページ」も変更される。

この方法では WWW ブラウザで閲覧したときに、



図-13 XML 版 (上) と XHTML 版 (下) の センターWeb サイトの比較

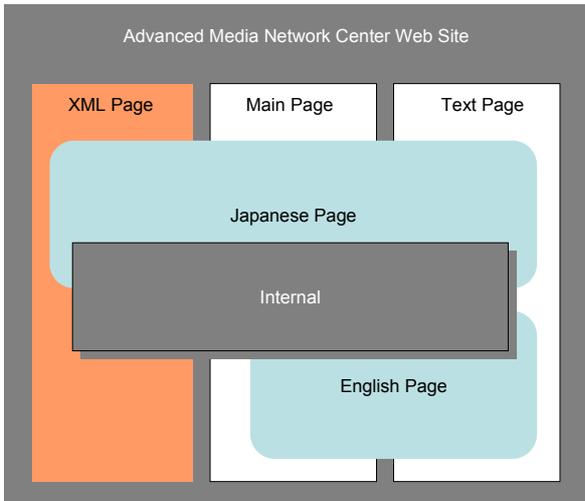


図-12 センターWeb サイトへの XML Web サイト実装後の構成



まずクライアント側に XML ファイルと XSL ファイルの全体を読み込ませるために、トップページの表示に時間がかかってしまう。しかし、学内からの利用であれば通信インフラが整備されていることもあり、全く利用に支障がない程度である。また、ファイルをすべて最初にクライアント側に送信するために、閲覧者がどのページやコンテンツを何回閲覧したかなどのログの記録が残らない。関連するファイルをすべてあらかじめ送信してしまうこの方式では、XML 文書の中の必要な情報だけを取り出すことができるという特長を生かしておらず、無駄な通信が発生してしまう。

5.まとめ

HTML によって記述されていた Web サイトから、柔軟で様々な情報を扱うことができる XML による Web サイトへの移行方法について検討をおこなった。

その結果、既存の HTML 文書を XHTML 文書に変更するだけでも XML 文書として扱うことが可能であるが、Web サイトとして活用するためにはさらに XSL ファイルを用いることが必要であり、XSL, XSLT を用いた変換処理をクライアント側での変換処理による方法とサーバ側の変換処理による方法に大別することができ、そのそれぞれの方式について 2 通りずつ、計 4 種類の方式を検討した。

その結果、クライアント側の処理の方法では、サーバ構成などに変更を加える必要がないが、閲覧可能なクライアント環境が制限されてしまう。サーバ側の処理をする方法では、種々のクライアント側の環境に対応することが可能であるが、サーバの構成が複雑になってしまう。機能の柔軟性と拡張性、セキュリティ性や運用性などを考慮すると、JavaScript を用いてクライアント側で変換処理をおこなう方式が有効であることがわかった。実際にその方式で Web サイトを構築し、既存の Web サイトと並行して XML による Web サイトも公開したところ、現在の環境では十分実用的であることを実証した。

ここでは、従来のコンテンツの利用を前提としたが、新規 Web サイトを構築する場合、今後の公開する情報やコンテンツの増加や様々な機能の実装などをおこなうことを考慮すると、これからはサーバ側の処理に移行すべきであろう。サーバ側の処理ではクライアント側の環境に合わせた Web サイトを公開できるため、既存の Web サイトを並行運用する必要もない。また、閲覧記録を残す機能の実装も比較的簡単におこなえるなど、より柔軟で拡張性が高い Web サイトを構成することができる。

今回実装したセンターWeb サイトは宇都宮大学総

合情報処理センターの Web サイトで公開している (<http://www.cc.utsunomiya-u.ac.jp/>)。

参考文献

- [1] 総務省:情報通信白書、総務省(2003).
- [2] W3C: Extensible Markup Language (XML) 1.0 (1998). <http://www.w3.org/TR/1998/REC-xml-19980210/>
- [3] 西村亨, 松田勝敬, 永井明:次世代 WEB 教材の開発、学術情報処理研究、No.6, pp.77-79(2002).
- [4] 松田勝敬, 西村亨, 永井明:XML データベースシステムの開発、学術情報処理研究、No.6, pp.73-76(2002).
- [5] 松田勝敬, 南木徹, 西村亨, 永井明:センターWeb サイトの開発、学術情報処理研究、No.7, pp.79-82(2003).
- [6] The Apache Software Foundation: The Apache HTTP Server Project (2004). <http://httpd.apache.org/>
- [7] The Perl Foundation :The Perl Directory (2004). <http://www.perl.org/>
- [8] IPTC: NewsML (2004). <http://www.newsml.org/pages/>
- [9] W3C: XHTML 1.0 The Extensible HyperText Markup Language (Second Edition) (2002). <http://www.w3.org/TR/xhtml1/>
- [10] W3C: XSL Transformations (XSLT) (1999). <http://www.w3.org/TR/xslt>
- [11] Netscape: JavaScript Central (2004). <http://devedge.netscape.com/central/javascript/>
- [12] The PHP Group: PHP: Hypertext Preprocessor (2004). <http://www.php.net/>
- [13] Ginger Alliance: Sablotron (2004). http://www.gingerall.com/charlie/ga/xml/p_sab.xml
- [14] The Apache Software Foundation: Apache Xindice (2004). <http://xml.apache.org/xindice/>
- [15] The Apache Software Foundation: Apache Jakarta Tomcat (2004). <http://jakarta.apache.org/tomcat/>
- [16] W3C: Simple Object Access Protocol (SOAP) 1.1 (2000). <http://www.w3.org/TR/2000/NOTE-SOAP-20000508/>
- [17] インターネット協会:インターネット白書 2003、インプレスコミュニケーションズ、東京 (2003).