

## XML を利用した技術用語検索システムの構築

建設情報研究室

学生氏名 中島 健蔵

指導教員 皆川 勝

### 1. 研究背景

eXtensible Markup Language(以下、XML)は 1998 年 2 月に World Wide Web Consortium が勧告(XML1.0 勧告)したマークアップ言語である。ISO(国際標準機構)が 1986 年に電子的に文書を管理するために国際標準の規格として定めた Standard Generalized Markup Language (以下 SGML) という難解な記述言語を、整理、簡素化し、インターネット上で利用ができるようにしたものである<sup>1)</sup>。XML は、SGML からハイパーテキストを作成するために必要な機能に絞りこんで開発された Hyper Text Markup Language (以下 HTML) にはない情報の一部を抽出したり並べ替えるといったデータ処理について考慮された記述言語である。また、XML は、すでに文書管理システムや Web データベース、電子商取引といった分野で導入が進められている。建設業界では、2000 年に国土交通省(旧建設省)が成果品の電子納品の標準化に向けて XML を採用した。

本研究では、教育分野での XML の利用に着目し、Web 上で利用できる XML データベースを用い技術用語検索システムを構築した。

本研究で構築したシステムは、「応用力学」「水理学」「土木計画」「コンクリート工学」「土質力学」の 5 科目の計 1500 語のデータベースとその検索システムである。データ処理機能に優れた XML データを利用したため、データの共有化、再利用が可能となっている。また、必要な科目や用語が増えても検索システムを拡張することができる。この技術用語検索システムは学習支援システムとして活用することができる。

### 2. XML データの構造

XML データの構造は図 1 で示すように三階層とした。XML ではタグを自由に設定できるので、図 1 のようにタグを記述した。XML データは「番号」「履修学年」「科目名」「頭文字」「用語名」「用語の意味」以上 6 個の項目からなる。

### 3. XSLT 及び XPath の利用

XML 文書には体裁情報が存在しない。XML 専用のスタイルシートである eXtensible Stylesheets Language Transformation (以下 XSLT) を作成することにより、XML データに体裁情報を持たせることができ、1つの XML データをクライアントの要求に応じてブラウザ表示することができる<sup>2)</sup>。本研究

では XML 文書をブラウザ表示させるために 2 種類の XSLT を作成した。又、XML 関連規格である XPath を、XSLT に組み込むことにより、XML データの一部のみをブラウザ表示させることができる。本研究では XML データ及びこれらの XML 関連規格を利用して検索システムを構築した。

### 4. XML データベースの構築

プログレスソフトウェア社が開発した eXtensible Information Server (以下 XIS) は、XML のアプリケーション開発環境であり、Tool box, eXcelon Data Server 等で構成されている<sup>3)</sup>。本研究では XIS を利用してデータベースを作成した。XIS に XML データを格納し、格納した XML データを、Web server である Internet Information Services 5.0 (以下 IIS) を経由させて Web クライアントに供給する。

```

<all>
  <word>
    <number>番号</number>
    <field>科目</field>
    <grade>履修学年</grade>
    <yougo jp="頭文字">用語</yougo>
    <explanation>意味</explanation>
  </word>
</all>

```

図 1 作成した XML データの抜粋

図2に構築したデータベースの構造を示す。ルートフォルダ名を sotsuron,ディレクトリ名をそれぞれの科目の名前とした。ブラウザ表示させるために必要な HTML ファイル, 科目ごとの XML 文書, XSLT をすべてのフォルダに格納した。

5. 検索システムの概要

図3で示すように XML データを XIS のルートフォルダに格納する。次に, XIS の Web クライアントから XML データベースへのアクセスを提供するサーバエクステンション機能を使用し, Web サーバを介してインターネットで閲覧, 管理などをできるようにした4)。

検索方法としては任意の科目の中から用語を検索する方法と, 全部の科目から用語を検索する方法を用意した。どちらの方法も, 検索項目をクリックし, 画面に表示される頭文字の中から用語の頭文字をクリックすることにより目的の用語が検索される。

6. 考察および研究成果

完成した検索システムのブラウザ画面を図4に示す。本研究の成果は, 建設業界で必要とされる様々な用語を, XML データベースとして, Web サーバを介してインターネットで閲覧できる検索システムを構築したことである。「XML を利用した技術用語検索システム」の構築により, 5つの科目の用語を1つの Web サイトから検索できるようになった。このことから本システムは効率の良い学習支援システムとして利用できる。

参考文献

- 1) 大久保 博樹：超入門 XML, 株式会社すばる舎, 2002.3.
- 2) 坂田 健二: windows で学ぶ XML, 技術評論社, 2002.5.
- 3) プロGRESSソフトウェア社：eXcelon ユーザーガイド, 2000.6.
- 4) 田村 郷司：知識情報のための XML による電子文書データベース統合管理システムの構築, 武蔵工業大学卒業論文, 2001.3.

謝辞

本研究を進めるにあたり, 佐藤安雄技師並びに修士2年の田村郷司氏には多大なご指導を頂きました。ここに感謝の意を表します。

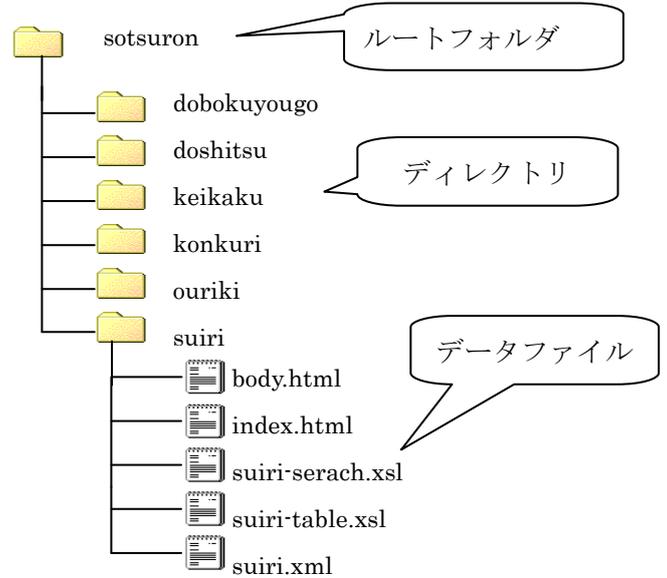


図2 XML データベースの構造

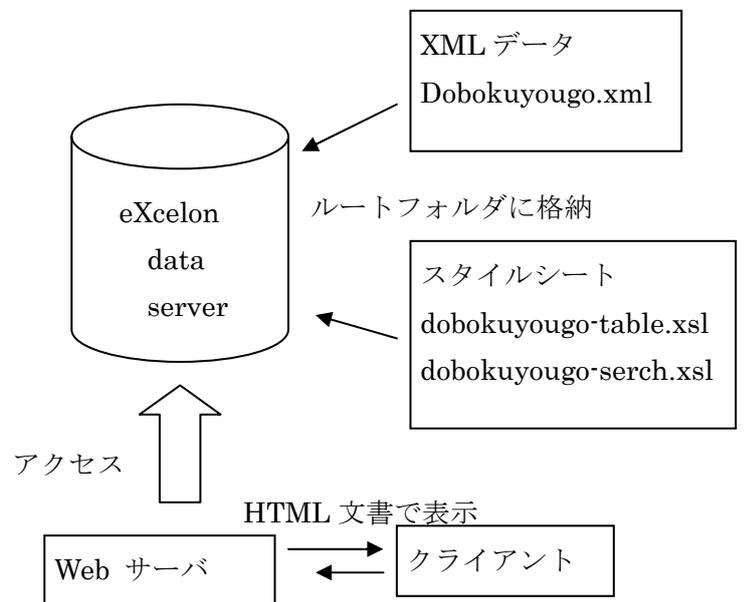


図3 検索システムの概要

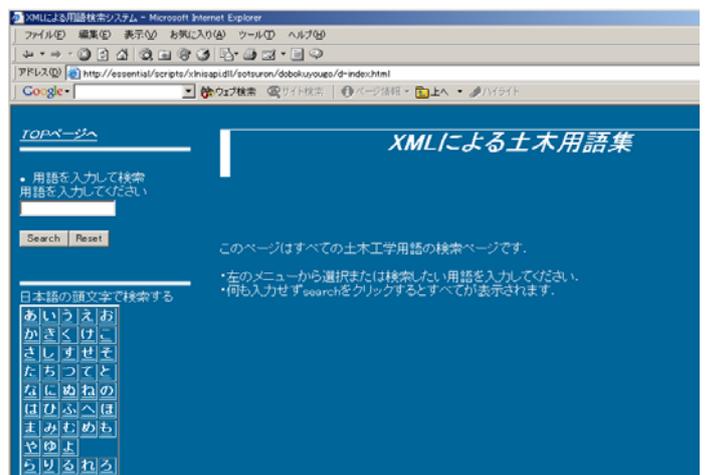


図4 検索システムのブラウザ画面

本研究を進めるにあたり、皆川勝教授、佐藤安雄技師並びに修士2年の田村郷司氏には多大なご指導を頂きました。ここに感謝の意を表します。

HTMLは段落、見出し、引用文といった文章の構成要素をタグと呼ばれる記号で指定します。これらの記号を該当する文に指定していく作業(あるいは結果)をマークアップといいます。