

XML によるシラバスデータベース構築と Web サービスの検討

井田 正明[†] 野澤 孝之[†] 宮崎 和光[†] 芳鐘 冬樹[†] 渋井 進[†] 喜多 一[‡]
 大学評価・学位授与機構[†] 京都大学 学術情報メディアセンター[‡]

1. 研究背景

情報ネットワークの整備とともに高等教育（大学等）においても情報技術の活用が進展しつつある。専門化・多様化が進み質の保証などが求められる高等教育では、カリキュラムの設計・評価のための計算機支援が有効であると考えられる。これまでに我々は高等教育における教育課程の分析のため、電子化された教育課程に関する情報の収集を進めてきた。また収集情報の構造を解析することにより教育課程に関する情報の標準的なデータ構造を検討し、XML に基づくシラバスデータベースを試作してきた。また、データベースの利用による教育課程の分析に関しても検討を行ってきた[1], [2], [3], [4]。

一般に教育課程に関するデータは定型的部分と非定型的な部分（例えば、シラバスにおける各回授業内容、教員情報、参考書等）が存在する。またその利用においても学生への情報提供のみならず大学評価など多様なデータの活用方法が考えられるため、データの有効利用のための XML によるデータベースの開発が必要と考えられる。これまでもデータ形式の検討やデータ管理、検索速度などの検討がなされ[5], [6]、インターネット上での学習に関連する URL 等の情報収集とデータベース化も進展している[7]。

2. シラバスデータベースシステムの開発

これまでに開発を行ったデータベースシステムの概要を Fig. 1 に示す[1], [2]。

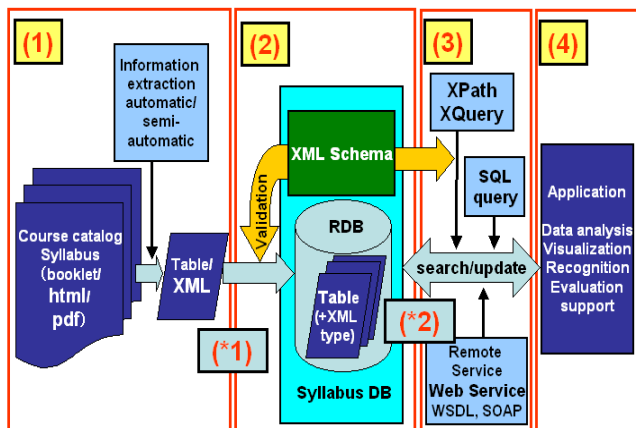


Fig. 1 Structure of a database system

(1), (2) (Fig. 1 の(*1)) : 高等教育機関（大学等）より収集したシラバス等から必要な情報を抽出する[3]。これら教育に関するデータをリレーショナルデータベースに格納する（XML スキーマによりデータ構造を表現）。一般に教育課程の編成方針やシラバスの内容の標準的な記述についてある程度の記載はあるが（学校教育法、各種答申など）、シラバス等は各機関ごとにさまざまな形式をとっているのが実情である。ここでは教育に関する情報の階層構造をつぎのように表現する[1], [2]。

機関（大学等） - 学部 - 学科 - 年度 - 教育課程 (program) - 授業科目 (course)

データベースの元となる資料は、Web, CD などより収集した履修科目表やシラバス等である。階層構造をもつデータをリレーショナルデータベースに格納するため、データ構造を分割した複数のテーブルを設計している：program テーブル（教育課程に関する項目）、linkage テーブル（教育課程と授業科目の関係）、course テーブル（各授業科目）。教育課程と授業科目は一般に多対多の関係にあるため、linkage テーブルによりそれらに関係付けている。これら以外に program や faculty に関するコード表のテーブルを設けている（例えば、course については Fig. 2）。

項目名	説明	minOccurs	maxOccurs
course	〈授業〉	1	1
courseID	授業コード	1	1
title	科目名	1	1
eTitle	英文科目名	1	1
courseObjectives	達成される目標	1	1
abstract	授業概要	1	1
keywords	キーワード	1	1
language	使用言語	1	1
term	開講学期	1	1
day	曜日	1	1
time	時限	1	1
room	教室	1	1
classType	授業形式（講義、実験など）	1	1
lecturers	〈教員リスト〉	1	1
lecturer	〈教員〉	1	N
lecturerID	教員コード	1	1
name	教員名	1	1
tel	教員連絡先電話	1	1
e-mail	教員電子メール	1	1
webSite	教員WebSite	1	1
office	教員研究室	1	1
officeHour	オフィスアワー	1	1

(後続は省略)

Fig. 2 Course table

実際に収集した各大学のシラバス（集合）は 1 科目 1 ファイルの構成が多く、各ファイルは HTML や PDF 形式である。これら文書データは大学（学部）ごとにある程度形式が統一された半構造化データであるため、各ファイルからの情報抽出を半自動的に行うことは可能ではあるが、Fig. 2 のような詳細な項目にデータを分解するためには個別に対応を工夫する必要がある。正確な情報抽出のための時間や労力およびその後のデータの活用方法を考慮にすると、実用上はすべてのファイルについて情報抽出を考えず、(1) HTML および PDF の文書構造をそのまま XML 形式に変換してデータベースに格納し (Fig. 3 参照。検索機能等は SQL に加えて XQuery が可能、Fig. 4 は program テーブルの XML データについての XQuery の例)、(2) (後に) 情報抽出すべきものを項目ごとにリレーショナル形式へと変換する。本年度は Web や CD から収集した教育に関する情報を XML 形式でデータベースに格納している。

列名	データ型	Null を許容
session_preparation14	nvarchar(255)	<input checked="" type="checkbox"/>
session_assignment14	nvarchar(255)	<input checked="" type="checkbox"/>
session_topics15	nvarchar(255)	<input checked="" type="checkbox"/>
session_preparation15	nvarchar(255)	<input checked="" type="checkbox"/>
session_assignment15	nvarchar(255)	<input checked="" type="checkbox"/>
keywords1	nvarchar(50)	<input checked="" type="checkbox"/>
keywords2	nvarchar(50)	<input checked="" type="checkbox"/>
keywords3	nvarchar(50)	<input checked="" type="checkbox"/>
keywords4	nvarchar(50)	<input checked="" type="checkbox"/>
keywords5	nvarchar(50)	<input checked="" type="checkbox"/>
language	nvarchar(50)	<input checked="" type="checkbox"/>
syllabus_xml	xml	<input type="checkbox"/>

Fig. 3 RDB and XML data

```

ALTER PROCEDURE [dbo].[ProgramListGetByInstitutionIDXML2]
AS SELECT xml_c.query(
'<doc>{
  for $c in /program
  return
    <program>
      <programName> "(data(%c/@programName))" </programName>
      <faculty> "(data(%c/@faculty))" </faculty>
      <institution> "(data(%c/@institution))" </institution>
    </program>
}' from program_xml
RETURN

```

Fig. 4 Xquery

(3), (4) (Fig. 1 の(*2)) : このデータベースの後段にキーワード抽出や科目の特徴の可視化など教育課程について解析を行うシステムを開発し研究を進めている [1], [2], [4]. すなわち、キーワード・専門用語の抽出、特徴空間の構成および解析、多変量解析・クラスタリング等の適用、結果検討および解釈等である。

また、検索機能（データ更新）などを Web サービス (SOAP) としてネットワークを通じて提供する方法を検討してきた [1], [2]. SQL 命令をそのまま受け付けるサービスや XQuery (XPath) による Web サービス、また簡便に Web サービスを利用するためのパラメータクエリを SOAP 経由で提供するように開発を進めている。データベースの内部はリレーショナル形式と XML 形式の混在があっても外部からはサービスとして同様の扱いができるよう検討している。

以上、高等教育機関における教育に関する半構造化データをデータベース化する際の実用上の工夫を検討した。またデータベースシステムからさまざまな Web サービスを提供することにより分散アプリケーションの開発を向上させるための方法を検討した。

参考文献

[1] M. Ida, T. Nozawa, F. Yoshikane, K. Miyazaki, H. Kita: Syllabus database and web service on higher education, Proc. of ICACT2005 (2005)

[2] 井田, 野澤, 芳鐘, 宮崎, 喜多: シラバスデータベースシステムの構築と専門教育課程の比較分析への応用, 大学評価・学位研究, 2, PP. 85-97, http://www.niad.ac.jp/n_shuppan/ (2005)

[3] 渡辺, 絹川, 井田, 芳鐘, 野澤, 喜多: Web上のシラバス情報の収集と XML 変換, FIT2004 (2004)

[4] 野澤, 井田, 芳鐘, 宮崎, 喜多: シラバスの文書クラスタリングに基づくカリキュラム分析システムの構築, 情報処理学会論文誌, Vol. 46, No. 1 (2005)

[5] 多田, 遠藤: オープンソースを用いた XML によるシラバスデータベースの試作, 情報処理学会研究報告, 2004-IS-89, pp. 47-53 (2004)

[6] 江本, 荒木, 大河内, 大淵, 那須, 松原, 横田, 国島: XML を基にしたシラバス管理システムの実現, IPSJ データベースシステム, 128-17, pp. 123-130 (2002)

[7] <http://nime-glad.nime.ac.jp/>