



Interoperabilität

Workflow Systems and Technology
Faculty of Computer Science
University of Vienna
Stefanie Rinderle-Ma, Juergen Mangler

26.03.2015

2. Übung: XML und Datenbanken (12 Punkte)

Datenbasis - (Vorsicht, dauert etwas bis es angezeigt wird). Alle Aufgaben auf diesem Übungsblatt sind auf der Basis dieser online abrufbaren Daten zu lösen.

Team 1: <http://online.univie.ac.at/vlvz?kapitel=501&semester=S2015&format=xml>
Team 2: <http://online.univie.ac.at/vlvz?kapitel=508&semester=S2015&format=xml>
Team 3: <http://online.univie.ac.at/vlvz?kapitel=701&semester=S2015&format=xml>
Team 4: <http://online.univie.ac.at/vlvz?kapitel=1001&semester=S2015&format=xml>
Team 5: <http://online.univie.ac.at/vlvz?kapitel=1004&semester=S2015&format=xml>
Team 6: <http://online.univie.ac.at/vlvz?kapitel=106&semester=S2015&format=xml>
Team 7: <http://online.univie.ac.at/vlvz?kapitel=107&semester=S2015&format=xml>
Team 8: <http://online.univie.ac.at/vlvz?kapitel=108&semester=S2015&format=xml>
Team 9: <http://online.univie.ac.at/vlvz?kapitel=202&semester=S2015&format=xml>
Team 10: <http://online.univie.ac.at/vlvz?kapitel=204&semester=S2015&format=xml>
Team 11: <http://online.univie.ac.at/vlvz?kapitel=405&semester=S2015&format=xml>
Team 12: <http://online.univie.ac.at/vlvz?kapitel=509&semester=S2015&format=xml>
Team 13: <http://online.univie.ac.at/vlvz?kapitel=707&semester=S2015&format=xml>
Team 14: <http://online.univie.ac.at/vlvz?kapitel=708&semester=S2015&format=xml>
Team 15: <http://online.univie.ac.at/vlvz?kapitel=2302&semester=S2015&format=xml>
Team 16: <http://online.univie.ac.at/vlvz?kapitel=2303&semester=S2015&format=xml>
Team 17: <http://online.univie.ac.at/vlvz?kapitel=2306&semester=S2015&format=xml>
Team 18: <http://online.univie.ac.at/vlvz?kapitel=2401&semester=S2015&format=xml>
Team 19: <http://online.univie.ac.at/vlvz?kapitel=2403&semester=S2015&format=xml>
Team 20: <http://online.univie.ac.at/vlvz?kapitel=2501&semester=S2015&format=xml>
Team 21: <http://online.univie.ac.at/vlvz?kapitel=2503&semester=S2015&format=xml>
Team 22: <http://online.univie.ac.at/vlvz?kapitel=2703&semester=S2015&format=xml>

Sie können NICHT davon ausgehen dass die Datenbasis korrektes XML ist - eventuell müssen entities wie ` `; die nicht teil von XML sind, vor dem einlesen umgewandelt werden (z.b. ` `; ind ` `;))

Die Daten sind in eine saubere Hierarchie abzubilden, die mindestens folgende Aspekte umfassen muss:

- Modul
- Name, Titel, ID, Typ, Prüfungsimmanenz, Gruppen, ECTS
- bei den Gruppen: Gebäudensprache, Streaming, Link zur Plattform, Block, einzelne Termine, Leiter (ev. Mehrere), verschiedene Beschreibungen
- anbietende SPL
- und noch andere je nach Fragestellung (siehe unten).

Die Hierarchie soll grob sein: Studium > Modul > LV > Gruppe

Aufgabe 2-1: Edge-Ansatz (3 Punkte)

- a) Erstellen Sie eine saubere hierarchische XML File, ohne unnötige Attribute und Elemente aus der gegebenen Datenbasis. Alle Attribut und Element Namen soll in englischer Sprache sein. Bilden Sie die Daten anhand des im Papers "Storing and Querying XML Data using an RDMBS" vorgestellten Edge-Ansatzes auf Tabellen eines relationalen DBMS ab. Benutzen Sie für das Speichern der Werte jeweils eigene Tabellen (siehe Abschnitt 2.2.1 im Paper). Erstellen Sie zu diesem Zweck eine neue Datenbank in DB2.
- b) Stellen Sie folgende Anfragen an Ihre Datenbank.
 - 1) Die am längsten pro Einheit dauernde(n) Lehrveranstaltungen(n) (Stunden, Gruppe, LV-Leiter)
 - 2) Die am längsten pro Semester dauernde(n) Lehrveranstaltungen(n) (Stunden, Gruppe, LV-Leiter)
 - 3) Alle nach 18 Uhr stattfindenden LVs, die prüfungsimmanent sind, und nur einen LV-Leiter haben.

Abzugeben: Datenbasis als XML, Saubere Datenbasis als XML, ein PDF mit queries und der jeweiligen Lösungsmenge

Aufgabe 2-2: Anfragen mit XQuery mit BaseX (3 Punkte)

Starten Sie BaseX (Windows download unter <http://basex.org/products/download/>; Ubuntu: apt-get install basex) und laden Sie das saubere hierarchische XML File aus Aufgabe 2-1.

Erstellen Sie folgende Anfragen in XQuery:

- a) Geben Sie die Kurstitel mit Modul, Datum und Ort, an dem sie stattfinden, aus.
- b) Geben Sie alle Verantwortlichen LV-Leiter aus.
- c) Geben Sie die Anzahl aller Module im Studium aus.
- d) Geben Sie (nur) KursNr und Titel der Kurse der Kursleiter, der Nachnamen mit M beginnt aus..
- e) Alle Vorlesungen die am Vormittag stattfinden.
- f) Alle Gruppen deren Unterrichtssprache NICHT Englisch ist.

Abzugeben: Datenbasis als XML, Saubere Datenbasis als XML, ein PDF mit queries und der jeweiligen Lösungsmenge

Aufgabe 2-3: Datenintegration mittels XSLT 1.0 (3 Punkte)

In dieser Aufgabe geht es darum, von Hand ein Mapping zwischen zwei Datenformaten, der eigenen sauberen Datenbasis und der sauberen Datenbasis eines Partnerteams, mittels XSLT zu erstellen. Die Idee wie folgt:

Bitte lassen Sie Ihre saubere Datenbasis (xml File) bis spätestens 06.04.2015 dem Ihnen zugeordneten Team per Email zukommen (cc an Herrn Herrnbrod). Dabei bitte wie folgt die Dokumente schicken:

- Team 1 -> Team 2 -> Team 3 -> Team 4 -> Team 5 ... Team 22 -> Team 1

Falls Sie kein Dokument an das andere Team schicken, wird Ihnen die Aufgabe „aberkannt“. Sollten Ihre XML Dokument und das XML Dokument des Partner Teams identisch sein, bitte melden.

Nehmen Sie nun Ihr eigenes Dokument und das Dokument der KollegInnen als Grundlage für die Transformation. Diese können Sie unter <https://cewebs.cs.univie.ac.at/awi-interop/ss15/index.php?t=xslt> entwickeln, testen und abgeben.

Aufgabe 2-4: Schemaintegration – Tool support (3 Punkte)

Machen Sie sich mit der mit der Funktionalität zum Mapping von XML Schemata in Altova Mapforce vertraut (30-Tage-Testversion unter: <http://www.altova.com/de/download-trial.html>).

Erstellen Sie Ihre für eigene saubere XML Datenbasis ein XML Schema im Format RelaxNG (<http://relaxng.org/tutorial-20011203.html>). Konvertieren Sie diesen Modell mittels „trang“ in ein XSD Modell und testen Sie ob Ihre saubere Datenbasis sowohl mit dem RNG als auch mit dem XSD validieren (Multi Schema Validator – MSV).

Bitte lassen Sie Ihr XSD Schema bis spätestens 10.04.2015 dem Ihnen zugeordneten Team per Email zukommen (cc an Herrn Herrnbrödt). Dabei bitte wie folgt die Dokumente schicken:

- Team 22 -> Team 21 -> Team 20 -> Team 19 -> Team 18 ... Team 1 -> Team 22

Definieren Sie dann in Altova Mapforce ein geeignetes Mapping zwischen ihrem eigenen Schema und dem Schema ihres Partnerteams. Geben Sie die Resultate als XQuery, XSLT und Java-Programm aus. Transformieren Sie dann unter Verwendung dieser Resultate book in book2. Dokumentieren Sie den gesamten Prozess in Ihrer Lösung.

Abzugeben: Datenbasis als XML, Saubere Datenbasis als XML, RNG Schema, XSD Schema, Partner XSD Schema, Mapping als XQuery, Mapping als XSLT, Mapping als JAVA Programm.