

Datenbanksysteme II

SS 2017 – Übungsblatt 6

1. Aufgabe (XML Schema)

- a) Geben Sie zu der XML-Instanz aus dem Übungsblatt 5 Aufgabe 3 ein konformes XML Schema an für die Elemente *biblioentry*, *authorgroup* und *title* an. Die Elemente *author* und *othercredit* können als *person* verallgemeinert werden. Diesbezüglich soll *person* als komplexer Typ definiert werden und *author* und *othercredit* als Elemente dieses Typs. Die Korrektheit einer XML-Instanz bzgl. eines XML-Schemas lässt sich ebenfalls mit einem Validator überprüfen, z.B. <http://www.xmlforasp.net/SchemaValidator.aspx>.

```
<?xml version="1.0" ?>
<bib>
  <biblioentry id="Rade97">
    <authorgroup>
      <author>
        <surname>Rade</surname>
        <firstname>Lennart</firstname>
      </author>
      <author>
        <surname>Westergren</surname>
        <firstname>Bertil</firstname>
      </author>
      <othercredit>
        <surname>Vachenaer</surname>
        <firstname>Peter</firstname>
        <address>
          <email>pv@springer.de</email>
        </address>
      </othercredit>
    </authorgroup>
    <title>Springers Mathematische Formeln</title>
    <subtitle>Taschenbuch für Ingenieure, Naturwissenschaftler,
              Informatiker, Wirtschaftswissenschaftler</subtitle>
    <publisher>
      <publishername>Springer-Verlag</publishername>
      <address>
        <city>Berlin</city>
      </address>
    </publisher>
    <pubdate>1997</pubdate>
    <pagenums start="33" end="36" />
  </biblioentry>
</bib>
```

2. Aufgabe (XMLTable)

Gegeben sei die Tabelle *GolfClubverzeichnis*(Ort, Golfclub). Dabei sei *Golfclub* vom Typ XML

| Ort | Golfclub |
|---------|--|
| Leipzig | <pre> <tns:GolfCountryClub xmlns:tns="http://www.example.org/GolfCountryClub" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.example.org/GolfCountryClub GolfCountryClub.xsd "> <GolfCourse NumberOfHoles="18"> <Name>Bird's Nest Course</Name> <Yardage>7125</Yardage> <Par>72</Par> <Handicap>74.6</Handicap> <Slope>138</Slope> </GolfCourse> <GolfCourse NumberOfHoles="18"> <Name>Eagle's Nest Course</Name> <Yardage>7198</Yardage> <Par>72</Par> <Handicap>73.5</Handicap> <Slope>136</Slope> </GolfCourse> <Member handicap="20"> <RegularMember> <FirstName>Keith</FirstName> <LastName>Chong</LastName> <MembershipExpiry>2010-10-31</MembershipExpiry> </RegularMember> </Member> <Member handicap="0"> <ExclusiveMember> <FirstName>James</FirstName> <MiddleName>Lion</MiddleName> <LastName>Chan</LastName> <ValetParkingID>124</ValetParkingID> <LockerNumber>10</LockerNumber> </ExclusiveMember> </Member> <Member handicap="25"> <RestrictedMember> <FirstName>Allan</FirstName> <LastName>Chong</LastName> <MembershipExpiry>2015-10-31</MembershipExpiry> <LimitedPlay>Weekend Only</LimitedPlay> </RestrictedMember> </Member> <Member handicap="10"> <RegularMember> <FirstName>Murray</FirstName> <MiddleName>Izzun</MiddleName> <LastName>Beaton</LastName> <MembershipExpiry>2015-09-30</MembershipExpiry> </RegularMember> </Member> </tns:GolfCountryClub> </pre> |

Formulieren Sie eine SQL- Query, die alle Mitglieder durch die Angabe von Vorname, Nachname und Handicap in einer Relation ausgibt, die reguläre Mitglieder sind und ein Handicap kleiner als 15 haben. Dabei sei der Golfclub in „Leipzig“.

Skizzieren Sie die resultierende Relation.

3. Aufgabe (XPath/XQuery)

Zum Ausführen von XPath/XQuery-Anfragen steht im Rahmen von **LOTS** (Leipzig Online-Test-System) ein Online-XQuery-Trainer unter <http://lots.uni-leipzig.de/xqtrain> bereit. Um im XQuery-Trainer auf einer bestimmten XML-Instanz/Datei Anfragen stellen zu können, ist der Anfrage `doc("datei.xml")` voranzustellen, z.B.:

```
doc("bib.xml")/bib/biblioentry[@id="Rade97"]
```

Nachfolgende Aufgaben beziehen sich auf das Schema der XML-Instanz von Aufgabe 1:

3.1 Erstellen Sie die folgenden Anfragen in *XPath* unter Verwendung von *bib.xml*

- a) Liste aller *ids* von Publikationen
- b) Liste sämtlicher Emailadressen
- c) Liste der Titel aller Publikationen von Verlegern aus Berlin
- d) Liste der *ids* aller Publikationen, in deren Titel oder Untertitel der Ausdruck "wissen" vorkommt
- e) Liste aller Publikationen, die mehr als 2 Seiten umfassen
- f) Liste der Titel der Publikationen, an denen (mindestens) ein Autor beteiligt war, dessen Nachname mit 'W' beginnt
- g) Anzahl aller Publikationen
- h) Liste aller Publikationen, die mehr als einen Autor haben
- i) Liste aller Publikationen, die zwischen 2001 und 2003 erschienen sind

3.2 Erstellen Sie die folgenden Anfragen in *XQuery* unter Verwendung von *bib.xml*. Beispiel:

```
for $b in doc("bib.xml")/bib/biblioentry
where $b/@id="Rade97"
return $b
```

- a) Geben Sie eine alphabetisch sortierte Liste aller Vornamen von Autoren aus.
- b) Geben Sie für jede Publikation den Buchtitel sowie die Autoren aus. Fügen Sie dabei den Vornamen und Nachnamen (mit einem Leerzeichen) zusammen und geben diesen im Element *aname* nach folgendem Muster aus:

```
<pub>
  <title>...</title>
  <aname>...</aname>
  <aname>...</aname>
</pub>
```

c) Geben Sie eine Liste aller Städte zusammen mit den dort ansässigen Verlegern nach folgendem Muster aus:

```
<L>
  <stadt>...</stadt>
  <verleger>...</verleger>
  <verleger>...</verleger>
</L>
```

d) Geben Sie eine aufsteigend sortierte Liste aller Jahre zusammen mit den Titeln der in diesem Jahr veröffentlichten Bücher nach folgendem Muster aus:

```
<year value="2000">
  <title>...</title>
  <title>...</title>
</year>
```

4. Aufgabe (Vorteile NoSQL- Systeme)

Welche Vorteile versprechen NoSQL-Systeme gegenüber relationalen DBS? Nennen Sie für die nachfolgenden Aussagen die Eigenschaft, weshalb NoSQL Systeme sich für diesen Aspekt eignen.

1. Facebook hat 300PB Daten in seinem Data Warehouse¹.
2. Dabei kommen 600TB Daten täglich neu hinzu.
3. Viele Datenquellen haben kein festvorgegebenes Schema, somit können Objekte eine beliebige Anzahl von Attributen besitzen und beliebige Beziehungen zu anderen Objekten aufweisen.

¹ <https://code.facebook.com/posts/229861827208629/scaling-the-facebook-data-warehouse-to-300-pb/>