

Datenbanken und Informationssysteme (Sommersemester 2017)

Übung 8

Abgabe bis 27. Juni 10:00 Uhr. Zu spät eingereichte Übungen werden nicht berücksichtigt.

Bitte reichen Sie Ihre Lösung in Dreiergruppen ein. Die Lösung zu diesem Übungsblatt wird in den Übungen am 27. und 28. Juni vorgestellt. Bitte beachten Sie auch die aktuellen Ankündigungen im L²P-Lernraum zur Vorlesung.

Aufgabe 8.1 (XML Schema)

(8 Punkte)

Bitte die XML Schema-Datei und die XML-Datei zu der Aufgabe 8.1 *zusätzlich* ins L2P hochladen. Dies vereinfacht die Korrektur.

Mit dem folgenden Befehl lässt sich eine XML-Datei anhand eines XML Schema validieren:

```
xmllint --schema Bundestagswahl.xsd Bundestagswahl-daten.xml
```

(xmllint gehört zu libxml2 und lässt sich z.B. unter Ubuntu Linux per `apt-get install libxml2-utils` installieren)

- a) Entwerfen Sie zunächst ein XML Schema, welches die folgenden Anforderungen zur Speicherung der Ergebnisse einer Bundestagswahl erfüllt. Das Schema muss ein valides XML Schema sein.
- Alle Informationen werden in dem Element Bundestagswahl gespeichert. Zur Bundestagswahl wird das Datum gespeichert. Für die Bundestagswahl werden die Daten mehrere Wahlkreise gespeichert. Ein Wahlkreis wird über seine Wahlkreisnummer identifiziert, zusätzlich hat ein Wahlkreis eine Bezeichnung.
 - Für jeden Wahlkreis werden als Erststimmen die Anzahl der Stimmen für mehrere Kandidaten und zusätzlich die Anzahl ungültigen Stimmen gespeichert. Ein Kandidat wird über seine Sozialversicherungsnummer identifiziert. Es wird sein Name, Beruf und eine beliebige Anzahl Telefonnummern gespeichert.
 - Außerdem werden für jeden Wahlkreis als Zweitstimmen die Anzahl der Stimmen für mehrere Parteien und die Anzahl ungültigen Stimmen gespeichert. Eine Partei wird über ihr Kürzel identifiziert.
- b) Geben Sie nun ein Beispiel XML-Dokument an, welches eine Ausprägung ihres XML Schemas ist. Ihr XML-Dokument muss mindestens eine Instanz für jedes Element aus ihrem XML Schema enthalten. Die XML-Datei muss valide zum Schema sein. Hinweis: Es ist nicht erforderlich Schlüsselbedingungen zu modellieren.

Bitte die XPath Queries in eine separate Text-Datei kopieren und *zusätzlich* ins L2P hochladen. Dies vereinfacht die Korrektur.

Wir haben die XPath Queries auf dieser Seite getestet:

<http://www.freeformatter.com/xpath-tester.html>

Gegeben sei der folgende Ausschnitt eines XML-Dokuments mit Informationen über eine Firma (s. auch XML-Dokument `firma.xml` als Anlage zum Übungsblatt).

Hinweis: Jeder Angestellte ist im Element Personal nur ein mal enthalten.

Listing 1: firma.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<Firma>
  <Personal>
    <Abteilung bezeichnung="Buchhaltung">
      <Angestellter personalnummer="43865">
        <Name>Angelica Mueller</Name>
        <Telefonnummern>
          <Telefonnummer>+49 555 43285761</Telefonnummer>
        </Telefonnummern>
        <EinstellungsDatum>07.04.2013</EinstellungsDatum>
        <Rolle>BuchhalterIn</Rolle>
        <Beaufsichtigt />
      </Angestellter>
    ...
  </Abteilung>
  <Abteilung bezeichnung="Marketing">
    <Angestellter personalnummer="27109">
      <Name>Thomas Schneider</Name>
      <Telefonnummern>
        <Telefonnummer>+49 555 90128745</Telefonnummer>
      </Telefonnummern>
      <EinstellungsDatum>01.09.2008</EinstellungsDatum>
      <Rolle>Abteilungsleiter</Rolle>
      <Beaufsichtigt>
        <Angestellter personalnummer="31826">
          <Name>Lisa Froehlich</Name>
          <Telefonnummern>
            <Telefonnummer>+49 555 73610283</Telefonnummer>
          </Telefonnummern>
          <EinstellungsDatum>15.11.2015</EinstellungsDatum>
          <Rolle>SekretaerIn</Rolle>
          <Beaufsichtigt />
        </Angestellter>
      </Beaufsichtigt>
    </Angestellter>
  </Abteilung>
  ...
</Personal>

  <Kunden>
    <Kunde kundennummer = "K1837246">
      <Name>Matthias Schroeder</Name>
      <Telefonnummern>
        <Telefonnummer>+49 555 23172639</Telefonnummer>
        <Telefonnummer>+49 555 13182310</Telefonnummer>
      </Telefonnummern>
    </Kunde>
    <Kunde kundennummer = "K3783440">
      <Name> Bert Lange</Name>
      <Telefonnummern>
        <Telefonnummer>+49 555 88327204</Telefonnummer>
      </Telefonnummern>
    </Kunde>
  </Kunden>
</Firma>
```

```

        <Telefonnummer>+49 555 16575461</Telefonnummer>
    </Telefonnummern>
</Kunde>
...
</Kunden>

<Auftraege>
    <Auftrag auftragsnummer = "A1392784">
        <Auftraggeber>
            <Kunde>K1837246</Kunde>
        </Auftraggeber>
        <Beauftragte>
            <Angestellter>33992</Angestellter>
            <Angestellter>43865</Angestellter>
        </Beauftragte>
        <Auftragsdatum>23.02.2017</Auftragsdatum>
        <Deadline>01.04.2017</Deadline>
    </Auftrag>
    <Auftrag auftragsnummer = "A0093747">
        <Auftraggeber>
            <Kunde>K5039127</Kunde>
            <Kunde>K3783440</Kunde>
            <Kunde>K7564029</Kunde>
        </Auftraggeber>
        <Beauftragte>
            <Angestellter>27109</Angestellter>
            <Angestellter>12938</Angestellter>
            <Angestellter>31826</Angestellter>
        </Beauftragte>
        <Auftragsdatum>06.01.2017</Auftragsdatum>
        <Deadline>15.09.2017</Deadline>
    </Auftrag>
    ...
</Auftraege>
</Firma>

```

Geben Sie für die folgenden geforderten Informationen/Knoten die jeweils passende XPath Anweisung an:

- Die Telefonnummern aller Kunden.
- Die Daten der Einstellung aller Sekretäre/Sekretärinnen.
- Die Anzahl aller Mitarbeiter der Abteilung Marketing.
- Die Auftragsnummern aller Aufträge von Kunden mit dem Namen "Matthias Schroeder".
- Die Personalnummern aller Angestellten, die zusammen mit dem Angestellten 12938 an Aufträgen zusammenarbeiten. Der Angestellte 12938 soll selbst nicht zurück gegeben werden.

Tipp: Für die Umsetzung sind Funktionen wie `text()` und `count()` erforderlich. Außerdem benötigen Sie unter Umständen logische Verknüpfungen wie `and` oder `or` bei Prädikaten. Informationen hierzu finden Sie im Internet, z.B. in der W3C Spezifikation von XPath.

Aufgabe 8.3 (RDF Datenmodell)

(5 Punkte)

Gegeben sind die folgenden Informationen bzgl. des fiktiven Universums aus der Fernsehserie und Buchreihe “The Expanse”:

- Es gibt einen Planeten mit dem Namen “Erde”.
- Die “Vereinten Nationen” sind eine politische Organisation.
- Die “Erde” wird von den “Vereinten Nationen” regiert.
- “Chrisjen Avasarala” ist der Name einer Person, die auch ein Politiker ist.
- “Chrisjen Avasarala” arbeitet für die “Vereinten Nationen”.

Verwenden Sie die angegebenen Präfixe:

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix exp: <http://example.org/the-expanse/> .

- a) Zeichnen Sie einen RDF Graphen, welcher die gegebenen Informationen ausdrückt. Beachten Sie dabei die folgenden Anweisungen zur Modellierung des RDF Graphen:
- Verwenden Sie soweit möglich Prädikate und Klassennamen aus dem Standard RDF Vokabular, und dem FOAF Vokabular (wie zum Beispiel `rdf:type`).
 - In allen anderen Fällen, können Sie Prädikate frei wählen. Sie müssen jedoch den Namespace `exp:` verwenden.
 - URIs für Entitäten können frei gewählt werden, müssen aber Teil des `exp:` Namespace sein.
- b) Geben Sie nun eine Serialisierung ihres RDF Graphen im Turtle Format an.