

**Aufgabe 1a)**

**[4+1 Punkte]**

Erzeugen Sie ein detailliertes<sup>1</sup> XML Schema für folgendes XML Dokument:

```
<?xml version="1.0" encoding="UTF-8"?>
<node>
  <child id="1" name="Wurzelement">
    <node ref="4" role="must-branch"/>
    <node ref="2" role="optional"/>
  </child>
  <child id="2" name="left_child">
    <node ref="3" role="optional"/>
  </child>
  <child id="3"/>
  <child id="4"/>
</node>
```

Passen Sie das o.g. XML Dokument ggf. an, damit dies auf ihr erstelltes Schema referenziert (vgl. Foliensatz XML, Folie 25)!

**Aufgabe 1b)**

**[2+2 Punkte]**

Erweitern Sie ihr XML Schema aus Aufgabe 1a), indem Sie für **id** einen Schlüssel, für **ref** ein Constraint definieren.

**Aufgabe 1c)**

**[2 Punkte]**

Die **role** muss innerhalb des Element-Kontextes eindeutig sein!

**Allgemeiner Bearbeitungshinweis:**

Geben Sie zu ihren Lösungen stets auch eine grobe Lösungsskizze mit an, damit ihre Ergebnisse nachvollziehbar sind!

Geben Sie die Lösung bis zum 20.06.2012, 8:00 Uhr ab, indem Sie sie entweder in Fach 45 einwerfen (links dreimalig als Buch getackert!), oder bevorzugt per E-Mail mit dem Betreff-prefix [is2] an wi2@ismll.de schicken. Sie erhalten eine Eingangsbestätigung bei E-Mailabgabe.

---

<sup>1</sup> Detailliert bedeutet hier, dass das Schema die Definitionen der beteiligten Elemente beinhalten muss. Eine Definition nach Folie 29 des Foliensatzes ist nicht ausreichend detailliert:

```
<xs:schema version="1.0" xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="node"/>
</xs:schema>
```

Bearbeiten Sie die folgenden Aufgaben mit Hilfe der eXist-XML Datenbank (oder einer eigenen XPath/XQuery Software wie z. B. Oxygen). Eine Mini-Anleitung finden Sie im Moodle. Die zu Grunde liegende XML Datei befindet sich ebendort.

Eine sehr gute Dokumentation zu XPath (und später: XQuery)-Funktionen finden Sie unter:

<http://www.w3.org/TR/xpath-functions/>

Hilfreich sind hier die Kapitel 7 (String-Manipulationen) und 15 (insb. Abschnitt 15.4). Hinweis: in dem Dokument werden Sie Funktionen mit Namensraum-Präfix (`fn:...`) finden. Lassen Sie diesen für ihre Anfragen bitte einfach weg!

Funktionen mit dem Präfix `op:` können Sie **nicht** direkt aufrufen (siehe den letzten Absatz des Abschnitts „Namespaces and Prefixes“: <http://www.w3.org/TR/xpath-functions/#namespace-prefixes>). Diese Funktionen werden auch nicht für die Lösung der Aufgaben benötigt!

Schreiben Sie XPath Ausdrücke:

### Beispielaufgabe

[0 Punkte]

Geben Sie eine Liste der Filmtitel aus.

*Lösung:*

```
                /descendant-or-  
self::node()/element(casting)/attribute(movieTitle)/string(.)
```

oder in Kurzform:

```
//casting/@movieTitle/string(.)
```

*Warum?*

`//casting` → „für alle Elemente mit dem Namen casting“ ...

`@movieTitle` → verweise auf das Attribut movieTitle ...

`/string(.)` → und gebe für den aktuellen Kontext (der Punkt, verweist in diesem Fall auf das Attribut movieTitle) als Zeichenkette zurück.

Warum wird das letzte `string(.)` benötigt? Der logischer „Rückgabewert“ eines XPath-Ausdrucks ist immer der Knoten selbst, der in den „Rückgabekontext“ eingebettet wird. Da ein Attribut jeweils nur innerhalb eines Element-Kontextes existieren kann, ist die Rückgabe des Attributs unzulässig. `string(.)` wandelt hier die komplette Rückgabe in eine Zeichenkette um (und dies ist valide).

**Allgemeiner Hinweis für alle Aufgaben:** Bei der Rückgabe von Attributen kann es erforderlich werden, explizit deren Zeichenkettenrepräsentation anzufragen. Dies gelingt am einfachsten, wenn ans Ende des XPath-Ausdrucks ein `/string(.)` gehängt wird. Attribute **können** in eXist **nicht** zurückgegeben werden!

**Aufgabe 2a)** [1 Punkt]

Geben Sie die Namen der Regisseure (directors) aus.

**Aufgabe 2b)** [1 Punkt]

Geben Sie die Namen der Regisseure jeweils nur ein mal aus (sodass bei mehrfachem Auftreten eines Regisseurs dieser nur ein Mal als Ergebnis aufgegeben wird).

**Aufgabe 2c)** [1 Punkt]

Geben Sie eine Liste der Schauspieler zurück, die „Grant“ im Namen haben (z. B. Hugh Grant, Beth Grant). Namen dürfen nicht doppelt auftreten!

**Aufgabe 2d)** [1 Punkt]

Geben Sie für die Schauspielerliste aus Aufgabe 2c die Filmtitel aus, in denen die Schauspieler mitgespielt haben!

**Aufgabe 2e)** [2 Punkte]

Geben Sie eine Liste von Filmen zusammen mit den Jahren zurück, die nach Jahren sortiert ist. Eine Beispielausgabe würde so aussehen:

1922 : Nosferatu, eine Symphonie des Grauens  
1925 : Gold Rush, The  
1927 : General, The  
1929 : Un chien andalou  
1930 : All Quiet on the Western Front

**Aufgabe 2f)** [2 Punkte]

Geben Sie alle Filmtitel aus, in denen mehr als 10 Schauspieler gecastet wurden. Geben Sie auch die Gesamtzahl der Schauspieler mit an!

Aufgabe 2g)

[2 Punkte]

Erstellen Sie eine Statistik über die Anzahl der Filme, die in einem Jahr vorkommen. Das Ergebnis soll genau 2 Informationen enthalten: Das Jahr, und die Anzahl der Filme in diesem Jahr, also z. B. :

1922 : 1  
1925 : 1  
1927 : 1  
1929 : 1  
1930 : 1

...

Aufgabe 2h)

[3 Punkte]

Geben Sie für jeden Schauspieler eine Liste von Jahresangaben zurück, in welchem er in einem Film mitgespielt hat. Ein Ergebnis kann z. B. so aussehen:

```
Found 1 in 0 seconds.
<<
1 <actors>
  <actor name="Adam Arkin">
    <year> 1998 </year>
    <year> 2000 </year>
  </actor>
  <actor name="Adam Baldwin">
    <year> 1987 </year>
    <year> 1990 </year>
    <year> 1995 </year>
    <year> 1996 </year>
    <year> 2000 </year>
  </actor>
  <actor name="Adam Beach">
    <year> 1998 </year>
  </actor>
  <actor name="Adam Bryant">
```

**Hinweis:** Aus Performance-Gründen sollten Sie für Aufgabe 2h die Liste der Schauspieler reduzieren, indem Sie sich beispielsweise auf Schauspieler mit dem Vornamen „Adam“ beschränken. Sollten Sie dies nicht tun, warten Sie im besten Fall „nur“ einige Minuten auf ein Ergebnis. Der gezeigte Screenshot entspricht der Realität, sodass eine korrekte Anfrage tatsächlich sofort (0 Sekunden) bearbeitet werden kann.

**Wichtig:** Da jede Anfrage (egal wie komplex) vom Server bearbeitet wird und **nicht abgebrochen werden** kann, kann es insb. am Dienstag Abend zu Performance-Problemen bis hin zur Unerreichbarkeit kommen! Eine rechtzeitige Bearbeitung dieses Aufgabenblatts wird daher empfohlen...