Kapitel 10

XML

Die *eXtensible Markup Language (XML)* ist ein vom World-Wide-Web-Konsortium vorgeschlagener Dokumentenverarbeitungsstandard, veröffentlicht unter `http://www.w3c.org/XML` . XML ermöglicht dem Benutzer die Strukturierung seiner Daten mit Hilfe von selbstgewählten *Tags*:

```
<Titel>Selber Atmen</Titel>
<Autor>Willi Wacker</Autor>
```

Mit Hilfe der *eXtensible Stylesheet Language (XSL)* (`http://www.w3c.org/style/xsl`) kann für die einzelnen Tags eine individuelle Darstellungsweise festgelegt werden, welche auf die zugehörigen Daten angewendet wird.

Auf diese Weise wird eine Trennung zwischen Struktur, Inhalt und Layout erreicht. Typischerweise verteilen sich daher die Angaben zu den Benutzerdaten, z.B. einen Zeitschriftenartikel, auf drei Dateien:

- `artikel.dtd`: *Document Type Definition* mit der Strukturbeschreibung
- `artikel.xml`: *XML-Datei* mit den durch *Tags* markierten Daten
- `artikel.xsl`: *Stylesheet* mit Angaben zum Rendern des Layout

Ein XML-Parser kann zu einer vorliegenden XML-Datei ohne Angabe der zugehörigen DTD überprüfen, ob die XML-Datei *wohlgeformt* ist, d.h. ob die grundsätzlichen Syntaxregeln eingehalten werden. Bei Vorlage der DTD kann der XML-Parser zusätzlich überprüfen, ob die Datei *gültig* ist, d.h. ob ihr Inhalt der Strukturbeschreibung gehorcht. Ein XSLT-Prozessor (*eXtensible Stylesheet Language Transformation)* rendert das Layout für die Daten der XML-Datei unter Anwendung des Stylesheets. Anstelle einer DTD kann die Struktur auch durch ein *XML Schema* beschrieben werden, welches selbst wiederum eine wohlgeformte XML-Datei darstellt (`http://www.w3.org/XML/Schema`).

Die Beispiele der folgenden Seiten lassen sich im Microsoft Internet Explorer ausführen.

`http://xml.apache.org/xerces-j/index.html` bietet den XML-Parser Apache Xerces zum Download.

`http://www.xmlme.com/Validator.aspx` bietet Online-Validierung von XML-Dokumenten.

`http://www.xmlme.com/XsltTransformer.aspx?mid` bietet Online-XSLT-Transformation.

10.1 Strukturierte Texte

Listing 10.1 zeigt die Document Type Definition für die typische Struktur eines Zeitschriftenartikels, Listing 10.4 zeigt als Alternative ein XML-Schema. Listing 10.2 zeigt den Inhalt des Dokuments, welches diese Strukturvorgaben erfüllt. Abbildung 10.1 zeigt die von Listing 10.3 definierte Ausgabe.

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<!ELEMENT artikel      (Kopf,Rumpf)>
<!ELEMENT Kopf         (Titel,Autor)>
<!ELEMENT Rumpf        (Kapitel*)>
<!ELEMENT Kapitel      (Titel, Absatz*)>
<!ELEMENT Titel        (#PCDATA)>
<!ELEMENT Autor        (#PCDATA)>
<!ELEMENT Absatz       (#PCDATA|betont)*>
<!ELEMENT betont       (#PCDATA)>
```

Listing 10.1: artikel.dtd

```
<?xml version='1.0' encoding="ISO-8859-1" ?>
<!DOCTYPE artikel SYSTEM "artikel.dtd" >
<artikel>
  <Kopf>
    <Titel>Selber Atmen</Titel>
    <Autor>Willi Wacker</Autor>
  </Kopf>
  <Rumpf>
    <Kapitel>
      <Titel>Einleitung</Titel>
      <Absatz>
        In einer Reihe von aufsehenerregenden Experimenten
        wurden krzlich am <betont>Max-Planck-Institut</betont>
        die Vorteile des selbstndigen Denkens herausgearbeitet.
      </Absatz>
      <Absatz>
        Unsere Forschungen lassen erwarten, da analoge
        Aussagen auch fr den Atmungsvorgang gelten knnten.
      </Absatz>
    </Kapitel>
    <Kapitel>
      <Titel>Ausblick</Titel>
      <Absatz>
        Es gibt viel zu tun; warten wir es ab !
      </Absatz>
    </Kapitel>
  </Rumpf>
</artikel>
```

Listing 10.2: artikel.xml

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">

  <xsl:template match="artikel">
    <html>
      <head><title><xsl:value-of select="Titel"/></title></head>
      <body><xsl:apply-templates/></body>
    </html>
  </xsl:template>

  <xsl:template match="artikel/Kopf/Titel">
    <center><h1><xsl:apply-templates/></h1></center>
  </xsl:template>

  <xsl:template match="artikel/Kopf/Autor">
    <center><h4><xsl:apply-templates/></h4></center>
  </xsl:template>

  <xsl:template match="artikel/Rumpf/Kapitel/Titel">
    <h3><xsl:apply-templates/></h3>
  </xsl:template>

  <xsl:template match="Absatz">
    <p><xsl:apply-templates/></p>
  </xsl:template>

  <xsl:template match="betont">
    <I><xsl:apply-templates/></I>
  </xsl:template>

</xsl:stylesheet>
```

Listing 10.3: artikel.xsl

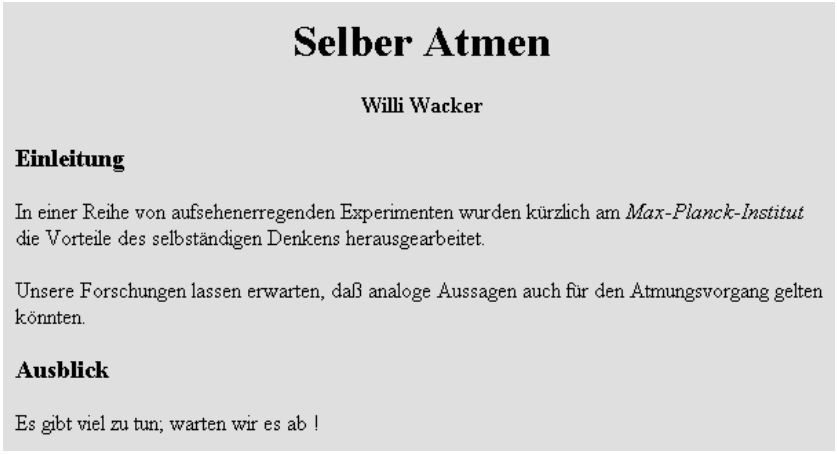

Abbildung 10.1: Ausgabe von artikel.xml mit Hilfe von artikel.xsl

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
    <xs:element name="artikel">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="Kopf"/>
                <xs:element ref="Rumpf"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>
    <xs:element name="Kopf">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="Titel"/>
                <xs:element ref="Autor"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>
    <xs:element name="Titel" type="xs:string"/>
    <xs:element name="Autor" type="xs:string"/>
    <xs:element name="Rumpf">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="Kapitel" minOccurs="0" maxOccurs="unbounded"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>
    <xs:element name="Kapitel">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="Titel"/>
                <xs:element ref="Absatz" minOccurs="0" maxOccurs="unbounded"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>
    <xs:element name="Absatz">
        <xs:complexType mixed="true">
            <xs:choice minOccurs="0" maxOccurs="unbounded">
                <xs:element ref="betont"/>
            </xs:choice>
        </xs:complexType>
    </xs:element>
    <xs:element name="betont" type="xs:string"/>
</xs:schema>
```

Listing 10.4: Strukturbeschreibung für artikel.xml durch ein XML-Schema

10.2 Strukturierte Daten

Listing 10.5 zeigt die Document Type Definition für einen Entity-Typ *personen*, bestehend aus den Feldern *Vorname, Nachname* und *Photo* und den Attributen *PersNr, Chef* und *Geschlecht*. Das Feld *Photo* ist leer und besteht nur aus einem Attribut *src* zur Angabe einer Bilddatei. Das Attribut *PersNr* vom Typ `ID` muß eindeutig sein, das Attribut *Chef* bezieht sich auf ein ID-Attribut. Listing 10.6 zeigt einige zur DTD passende Daten.

```
<?xml version='1.0' encoding="ISO-8859-1" ?>
<!ELEMENT personen             (person*)>
<!ELEMENT person               (vorname,nachname,photo)>
<!ATTLIST person   persnr      ID #REQUIRED>
<!ATTLIST person   chef        IDREF #IMPLIED>
<!ATTLIST person   geschlecht  (maennlich|weiblich) #REQUIRED>
<!ELEMENT vorname              (#PCDATA)>
<!ELEMENT nachname             (#PCDATA)>
<!ELEMENT photo                 EMPTY>
<!ATTLIST photo    src          CDATA #REQUIRED>
```

Listing 10.5: personen.dtd

```
<?xml version="1.0"?>

<personen>

  <person persnr="P4711" chef="P4712" geschlecht="weiblich">
    <vorname>Susi</vorname>
    <nachname>Sorglos</nachname>
    <photo src="erika.gif"/>
  </person>

  <person persnr="P4712" geschlecht="maennlich">
    <vorname>Willi</vorname>
    <nachname>Wacker</nachname>
    <photo src="willi.gif"/>
  </person>

</personen>
```

Listing 10.6: personen.xml

Listing 10.7 zeigt die zur DTD von Listing 10.5 gleichwertige Schema-Definition; Listing 10.8 zeigt eine XML-Datei, welche auf dieses Schema Bezug nimmt.

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="personen" type="PersonenTyp"/>
  <xsd:complexType name="PersonenTyp">
    <xsd:sequence>
      <xsd:element name="person" type="PersonTyp"
                                 minOccurs="0" maxOccurs="unbounded"/>
    </xsd:sequence>
  </xsd:complexType>
  <xsd:complexType name="PersonTyp">
    <xsd:sequence>
      <xsd:element name="vorname"    type="xsd:string"/>
      <xsd:element name="nachname"   type="xsd:string"/>
      <xsd:element name="photo"      type="PhotoTyp"/>
    </xsd:sequence>
    <xsd:attribute  name="persnr"     type="xsd:ID"/>
    <xsd:attribute  name="chef"       type="xsd:IDREF">
    <xsd:attribute  name="geschlecht" use="required">
                    <xsd:simpleType>
                      <xsd:restriction base="xsd:NMTOKEN">
                          <xsd:enumeration value="maennlich"/>
                          <xsd:enumeration value="weiblich"/>
                      </xsd:restriction>
                    </xsd:simpleType>
    </xsd:attribute>
  </xsd:complexType>
  <xsd:complexType name="PhotoTyp">
    <xsd:attribute name="src" type="xsd:string"/>
  </xsd:complexType>
</xsd:schema>
```

Listing 10.7: personen.xsd

```
<?xml version="1.0"?>
<personen xmlns="http://www-lehre.inf.uos.de/~dbs/2005/XML/personen.xsd">
  <person persnr="P4711" chef="P4712" geschlecht="weiblich">
    <vorname>Susi</vorname>
    <nachname>Sorglos</nachname>
    <photo src="erika.gif"/>
  </person>
  <person persnr="P4712" geschlecht="maennlich">
    <vorname>Willi</vorname>
    <nachname>Wacker</nachname>
    <photo src="willi.gif"/>
  </person>
</personen>
```

Listing 10.8: personen-mit-xsd-verweis.xml

Listing 10.9 zeigt eine Document Type Definition für den Entity-Typ *Dozenten*, der den Entity-Typ *Professoren* erweitert um eine variable Anzahl von Feldern mit dem Namen *Amt*. Als Alternative zeigt Listing 10.11 das passende XML-Schema. Abbildung 10.2 zeigt Teil einer Ausprägung des XML-Baums. Listing 10.10 zeigt dazu passende Daten.

```
<?xml version='1.0' encoding="ISO-8859-1" ?>
<!ELEMENT dozenten    (dozent*)>
<!ELEMENT dozent      (PersNr,Name,Rang,Raum,Amt*)>
<!ELEMENT PersNr      (#PCDATA)>
<!ELEMENT Name        (#PCDATA)>
<!ELEMENT Rang        (#PCDATA)>
<!ELEMENT Raum        (#PCDATA)>
<!ELEMENT Amt         (Bezeichnung,Termin?)>
<!ELEMENT Bezeichnung (#PCDATA)>
<!ELEMENT Termin      (#PCDATA)>
```

Listing 10.9: dozenten.dtd

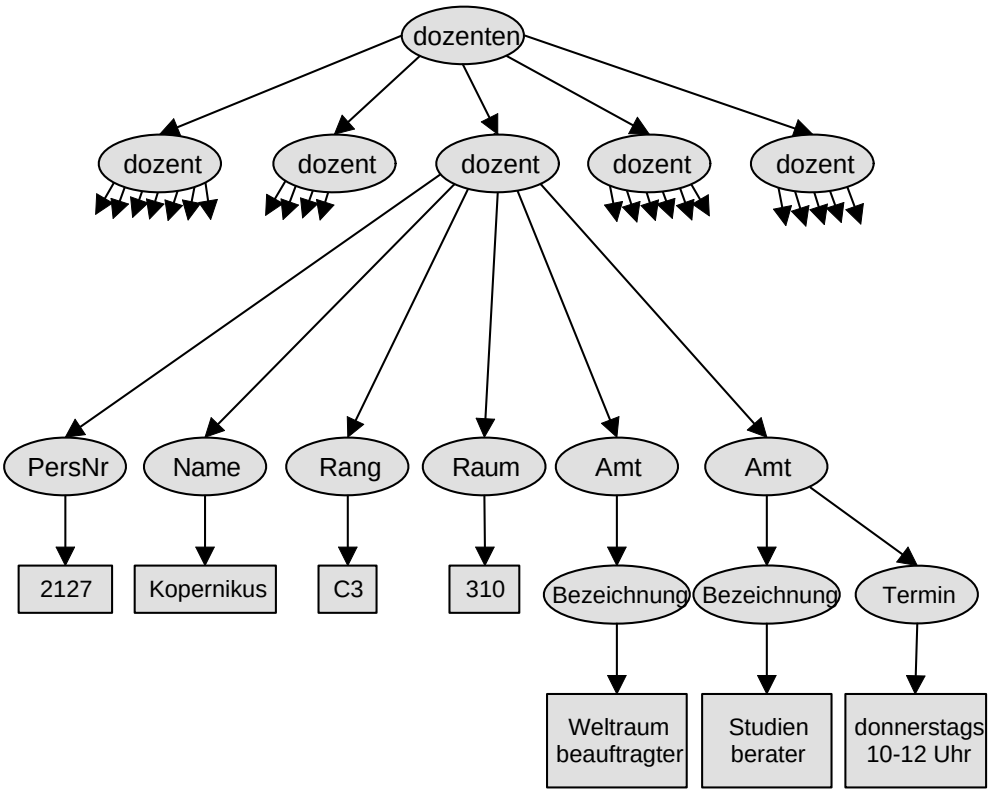

Abbildung 10.2: XML-Baum zur Datei dozenten.xml

```
<?xml version='1.0' encoding="ISO-8859-1" ?>
<!DOCTYPE dozenten SYSTEM "dozenten.dtd" >
<dozenten>
  <dozent>
    <PersNr>2125</PersNr>
    <Name>Sokrates</Name>
    <Rang>C4</Rang>
    <Raum>226</Raum>
    <Amt>
      <Bezeichnung>Dekan</Bezeichnung>
      <Termin>vormittags</Termin>
    </Amt>
  </dozent>
  <dozent>
    <PersNr>2126</PersNr>
    <Name>Russel</Name>
    <Rang>C4</Rang>
    <Raum>232</Raum>
  </dozent>
  <dozent>
    <PersNr>2127</PersNr>
    <Name>Kopernikus</Name>
    <Rang>C3</Rang>
    <Raum>310</Raum>
    <Amt>
      <Bezeichnung>Weltraumbeauftragter</Bezeichnung>
    </Amt>
    <Amt>
      <Bezeichnung>Studienberater</Bezeichnung>
      <Termin>donnerstags 10-12 Uhr</Termin>
    </Amt>
  </dozent>
  <dozent>
    <PersNr>2133</PersNr>
    <Name>Popper</Name>
    <Rang>C3</Rang>
    <Raum>52</Raum>
  </dozent>
  <dozent>
    <PersNr>2134</PersNr>
    <Name>Augustinus</Name>
    <Rang>C3</Rang>
    <Raum>309</Raum>
  </dozent>
</dozenten>
```

Listing 10.10: dozenten.xml

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

    <xs:element name="dozenten">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="dozent" minOccurs="0" maxOccurs="unbounded"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>

    <xs:element name="dozent">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="PersNr"/>
                <xs:element ref="Name"/>
                <xs:element ref="Rang"/>
                <xs:element ref="Raum"/>
                <xs:element ref="Amt" minOccurs="0" maxOccurs="unbounded"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>

    <xs:element name="PersNr" type="xs:string"/>

    <xs:element name="Name" type="xs:string"/>

    <xs:element name="Rang" type="xs:string"/>

    <xs:element name="Raum" type="xs:string"/>

    <xs:element name="Amt">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="Bezeichnung"/>
                <xs:element ref="Termin" minOccurs="0"/>
            </xs:sequence>
        </xs:complexType>
    </xs:element>

    <xs:element name="Bezeichnung" type="xs:string"/>

    <xs:element name="Termin" type="xs:string"/>

</xs:schema>
```

Listing 10.11: Strukturbeschreibung für dozenten.xml durch ein XML-Schema

10.3 XML-Ausgabe auf Web-Seiten

Alle Beispiele dieser Sektion (mit Ausnahme des letzten) beziehen sich auf die Document Type Definition `dozenten.dtd`, gezeigt in Listing 10.9, sowie die XML-Datei `dozenten.xml`, gezeigt in Listing 10.10.

Listing 10.12 zeigt den Quelltext einer HTML-Datei `dozenten-tabelle.html`, in der die Elemente der XML-Datei `dozenten.xml` zu einer Tabelle aufbereitet werden. Abbildung 10.3 zeigt die Ausgabe.

```
<html>
  <head>
    <title>dozenten-tabelle.html</title>
  </head>
  <body bgcolor="DDDDDD">
    <xml id="dozenten" src="dozenten.xml"></xml>
    <table border="1" bgcolor="ABCDEF" datasrc="#dozenten">
      <tr>
        <td><span datafld="Name"></span></td>
        <td><span datafld="Rang"></span></td>
      </tr>
    </table>
  </body>
</html>
```

Listing 10.12: HTML-Datei mit XML-Datei

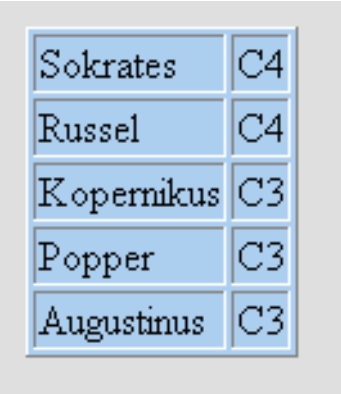

Abbildung 10.3: Ausgabe einer XML-Datei als HTML-Tabelle

Listing 10.13 zeigt den Quelltext der HTML-Datei `dozenten-traverse.html`, in der mittels VB-Script der Microsoft XML-Parser verwendet wird, um die Knoten des XML-Baumes der XML-Datei `dozenten.xml` zu traversieren. Abbildung 10.4 zeigt die Ausgabe.

```
<html>
  <head>
    <title>dozenten-traverse.html</title>
  </head>
  <body bgcolor="dddddd">
    <script type="text/vbscript">
      set xmlDoc=CreateObject("Microsoft.XMLDOM")
      xmlDoc.async="false"
      xmlDoc.load("dozenten.xml")
      document.write("<h1>Traversieren der XML Knoten</h1>")
      for each x in xmlDoc.documentElement.childNodes
        document.write("<b>" & x.nodename & "</b>")
        document.write(": ")
        document.write(x.text)
        document.write("<br><br>")
      next
    </script>
  </body>
</html>
```

Listing 10.13: Traversieren der XML-Knoten

Abbildung 10.4: Traversieren der XML-Knoten

Listing 10.14 zeigt den Quelltext der HTML-Datei `dozenten-navigation.html`, in der mit Hilfe von Javascript durch die Elemente der XML-Datei `dozenten.xml` navigiert wird. Abbildung 10.5 zeigt die Ausgabe.

```
<html>
  <head><title>dozenten-navigation.html</title></head>
    <script type="text/javascript">
      function naechste(){
        x=vor.recordset
        if (x.absoluteposition < x.recordcount){ x.movenext()}
      }
      function vorige(){
        x=vor.recordset
        if (x.absoluteposition > 1){ x.moveprevious()}
      }
    </script>
  </head>
  <body bgcolor="dddddd">
    <xml src="dozenten.xml" id="vor" async="false"></xml>
    <B>Name:</B> <span datasrc="#vor" datafld="Name"></span><br>
    <B>Rang:</B> <span datasrc="#vor" datafld="Rang"></span><br>
    <B>Raum:</B> <span datasrc="#vor" datafld="Raum"></span><br>
    <input type="button" value="Voriger Dozent"  onclick="vorige()">
    <input type="button" value="Nchster Dozent" onclick="naechste()">
  </body>
</html>
```

Listing 10.14: Navigation durch eine XML-Datei mit Hilfe von Javascript

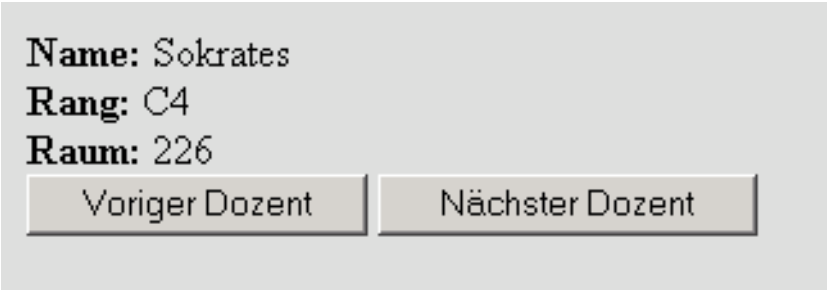

Abbildung 10.5: Navigation durch die XML-Datei mit Hilfe von Javascript

Listing 10.15 zeigt den Quelltext der Cascading-Stylesheet-Datei `dozenten.css`, auf die von der XML-Datei `dozenten.xml` Bezug genommen werden kann, nachdem dort die Zeile

```
<?xml-stylesheet type="text/css" href="dozenten.css"?>
```

eingefügt worden ist. Abbildung 10.6 zeigt die Ausgabe.

```
dozenten{background-color: #dddddd;width: 100%;}
dozent {display: block;   margin-bottom: 10pt;margin-left: 15pt;}
PersNr {color:   #0000FF; font-size: 20pt;}
Name   {color:   #FF0000; font-size: 28pt;}
Rang   {color:   #0000FF; font-size: 20pt;}
Raum   {color:   #0000FF; font-size: 20pt;}
Amt    {Display: block;   color: #000000;margin-left: 20pt;}
```

Listing 10.15: Cascading Stylesheet für dozenten.xml

Abbildung 10.6: Ausgabe von dozenten.xml mit Hilfe von dozenten.css

Listing 10.16 zeigt den Quelltext der XSL-Datei `dozenten.xsl`, auf die von der XML-Datei `dozenten.xml` Bezug genommen werden kann über die Zeile

```
<?xml-stylesheet type="text/xsl" href="dozenten.xsl"?>
```

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">

<xsl:template match="/">
  <html>
    <body>
      <table border="1" bgcolor="ABCDEF">
        <tr>
          <th>Pers-Nr</th><th>Name</th><th>Rang</th><th>Raum</th><th>mter</th>
        </tr>
        <xsl:for-each select="/dozenten/dozent">
          <tr>
            <td><xsl:value-of select="PersNr" /></td>
            <td><xsl:value-of select="Name"   /></td>
            <td><xsl:value-of select="Rang"   /></td>
            <td><xsl:value-of select="Raum"   /></td>
            <td>
               <xsl:for-each select="Amt">
                 <xsl:value-of select="Bezeichnung" />
                 <xsl:if test="position() != last()">, </xsl:if>
               </xsl:for-each>
            </td>
          </tr>
        </xsl:for-each>
      </table>
    </body>
  </html>
</xsl:template>
</xsl:stylesheet>
```

Listing 10.16: dozenten.xsl

Pers-Nr	Name	Rang	Raum	Ämter
2125	Sokrates	C4	226	Dekan
2126	Russel	C4	232	
2127	Kopernikus	C3	310	Weltraumbeauftrager, Studienberater
2133	Popper	C3	52	
2134	Augustinus	C3	309	

Abbildung 10.7: Ausgabe von dozenten.xml mit Hilfe von dozenten.xsl

Eine Alternative zur Anwendung einer XSL-Datei auf eine XML-Datei besteht in der Formulierung einer HTML-Seite, in der mit Hilfe des Microsoft XML-Parsers zunächst die XML-Datei und die XSL-Datei geladen werden und danach die Transformation angestoßen wird. Listing 10.17 zeigt den Quelltext der HTML-Datei `dozenten-mit-xsl.html`.

In beiden Fällen kommt dieselbe XSL-Datei zur Anwendung , das Ergebnis wird in Abbildung 10.7 gezeigt.

```
<html>
  <head><title>dozenten-mit-xsl.html</title></head>
  <body bgcolor="dddddd">
    <script type="text/javascript">

      var xml = new ActiveXObject("Microsoft.XMLDOM") // XML-File laden
      xml.async = false
      xml.load("dozenten.xml")

      var xsl = new ActiveXObject("Microsoft.XMLDOM") // XSL-File laden
      xsl.async = false
      xsl.load("dozenten.xsl")

      document.write(xml.transformNode(xsl))          // transformieren

    </script>
  </body>
</html>
```

Listing 10.17: HTML-Datei zur Ausgabe von dozenten.xml mit Hilfe von dozenten.xsl

Das folgende Beispiel bezieht sich auf `personen.dtd`, eine im Listing 10.5 gezeigte Document Type Definition. Neben dem Attribut *persnr* vom Typ `ID` zur eindeutigen Kennung einer Person und dem Attribut *chef* vom Tpy `IDREF` zur Referenz auf eine persnr gibt es ein leeres Feld *Photo* mit dem Attribut *src* zur Angabe einer Bilddatei.

Listing 10.18 zeigt den Quelltext der XSL-Datei `personen.xsl`, auf die von der XML-Datei `personen.xml` Bezug genommen werden kann, nachdem die Zeile

```
<?xml-stylesheet type="text/xsl" href="personen.xsl"?>
```

eingefügt wurde. Abbildung 10.8 zeigt die Ausgabe.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">

<xsl:template match="/">
  <html>
    <body bgcolor="dddddd">
      <table border="1" cellpadding="5" bgcolor="abcdef">
        <tr>
          <th>Pers-Nr</th><th>Anrede</th><th>Vorname</th><th>Nachname</th>
          <th>Chef</th><th>Photo</th>
        </tr>
        <xsl:for-each select="/personen/person">
          <tr>
            <td><xsl:value-of select="@persnr"/></td>
            <td>
             <xsl:if test="@geschlecht='weiblich'">Frau</xsl:if>
             <xsl:if test="@geschlecht='maennlich'">Herr</xsl:if>
            </td>
            <td><xsl:value-of select="vorname"    /></td>
            <td><xsl:value-of select="nachname"   /></td>
            <td><xsl:value-of select="@chef"      /></td>
            <td><IMG SRC="{photo/@src}"/></td>
          </tr>
        </xsl:for-each>
      </table>
    </body>
  </html>
</xsl:template>
</xsl:stylesheet>
```

Listing 10.18: personen.xsl

Pers-Nr	Anrede	Vorname	Nachname	Chef	Photo
P4711	Frau	Erika	Mustermann	P4712	
P4712	Herr	Willi	Wacker		

Abbildung 10.8: Ausgabe von personen.xml mit Hilfe von personen.xsl

10.4 XML vom Microsoft SQL Server 2000

Der Microsoft SQL-Server 2000 unterstützt XML sowohl bei der Überführung von XML-Dokumenten in die Datenbank als auch bei der Wiedergabe von Tabellendaten in XML-Notation.

Durch Aufruf der stored procedure `sp_xml_prepare_document` kann der wohlgeformte Inhalt eines XML-Strings in eine interne Darstellung überführt werden. Durch Aufruf von `OPENXML` kann der Inhalt des XML-Baums als Rowset geliefert werden. Listing 10.19 zeigt den gesamten Aufruf, Abbildung 10.9 zeigt die erzeugte Tabelle.

```
DECLARE @hdoc int                       -- Handle fuer interne Darstellung
DECLARE @doc varchar(1000)              -- Variable fuer XML-String
SET @doc =                              -- initialisiere XML-String
'<dozenten>
    <dozent>
      <persnr>4711</persnr>
      <name>Willi</name>
      <rang>C4</rang>
      <raum>801</raum>
    </dozent>
 </dozenten>'

EXEC sp_xml_preparedocument             -- lege interne Darstellung @hdoc
     @hdoc OUTPUT, @doc                 -- fuer den XML-String @doc an

SELECT * FROM OPENXML                   -- werte XML-Baum aus
       (@hdoc,'/dozenten/dozent')       -- ueber Handle @hdoc

EXEC sp_xml_removedocument @hDoc        -- entferne interne Darstellung
```

Listing 10.19: Aufruf von OPENXML

	id	parentid	nodetype	localname	prefix	namespaceuri	datatype	prev	text
1	2	0	1	dozent	NULL	NULL	NULL	NULL	NULL
2	3	2	1	persnr	NULL	NULL	NULL	NULL	NULL
3	7	3	3	#text	NULL	NULL	NULL	NULL	4711
4	4	2	1	name	NULL	NULL	NULL	3	NULL
5	8	4	3	#text	NULL	NULL	NULL	NULL	Willi
6	5	2	1	rang	NULL	NULL	NULL	4	NULL
7	9	5	3	#text	NULL	NULL	NULL	NULL	C4
8	6	2	1	raum	NULL	NULL	NULL	5	NULL
9	10	6	3	#text	NULL	NULL	NULL	NULL	801

Abbildung 10.9: XML-Baum als Tabelle, erstellt von OPENXML

Der Inhalt der von OPENXML gelieferten Tabelle kann mit Hilfe eines `INSERT`-Statements in eine bestehende Datenbanktabelle eingefügt werden. Hierbei müssen über das Schlüsselwort `WITH` die zu verwendenden Spaltentypen angegeben werden. Dem OPENXML-Kommando wird durch die Flagge `1` eine attributbezogene, durch die Flagge `2` eine elementbezogene XML-Struktur signalisiert. Listing 10.20 zeigt den kompletten Ablauf.

```
DECLARE @hdoc int                       -- Handle fuer interne Darstellung
DECLARE @doc varchar(1000)              -- Variable fuer XML-String
SET @doc =                              -- initialisiere XML-String
'<dozenten>
    <dozent>
      <persnr>4711</persnr>
      <name>Willi</name>
      <rang>C4</rang>
      <raum>801</raum>
    </dozent>
    <dozent>
      <persnr>4712</persnr>
      <name>Erika</name>
      <rang>C3</rang>
      <raum>802</raum>
    </dozent>
 </dozenten>'

EXEC sp_xml_preparedocument             -- lege interne Dartstellung @hdoc
     @hdoc OUTPUT, @doc                 -- fuer den XML-String @doc an

insert into professoren                 -- fuege in Tabelle professoren ein
            (persnr,name,rang,raum)     -- Persnr, Name, Rang, Raum
SELECT * FROM OPENXML                   -- werte XML-Baum aus
       (@hdoc,'/dozenten/dozent',2)     -- Parameter 2: Elementbezogen
with (persnr int, name varchar(20),     -- Datentypen fuer persnr und Name
      rang char(2), raum int)           -- Datentypen fuer Rang und Raum

EXEC sp_xml_removedocument @hDoc        -- entferne interne Darstellung
```

Listing 10.20: Einfügen mit Hilfe von OPENXML

Eine Ausgabe von Tabelleninhalten in geschachtelter XML-Notation ist beim SQL-Server 2000 möglich im Rahmen einer erweiterten SQL-Syntax:

```
select persnr, name from professoren
where rang='C3'
for xml auto, elements
```

Hierbei wird für jede Zeile aus dem Result-Set ein Tupel erzeugt, welches geklammert wird durch Tags mit dem Namen der Tabelle. Jede Spalte des Result-Set wird geklammert durch Tags mit dem Namen der Spalte:

```
<professoren><persnr>2127</persnr><name>Kopernikus</name></professoren>
<professoren><persnr>2134</persnr><name>Augustinus</name></professoren>
```

10.5 XQuery

Als Abfragesprache für XML-Dokumente scheint sich *XQuery* durchzusetzen. Im April 2005 wurde unter der Adresse `http://www.w3.org/TR/xquery` der letzte W3C Working Draft veröffentlicht.

In *XQuery* werden Abfragen durch eine Kombination von *FLWR*-Ausdrücken, Pfadausdrücken und Element-Konstruktoren formuliert. Ein *FLWR*-Ausdruck (gesprochen wie das englische Wort *Flower*) besteht aus einer Sequenz der Schlüsselworte `FOR, LET, WHERE, RETURN`. Ein Pfadausdruck nach der **XPath**-Syntax (`http://www.w3.org/TR/xpath.html`) fasst das Dokument als Baum von Element-Knoten auf und liefert durch gezielte Navigation den gewünschten Teilbaum ab. Dabei trennt ein einzelner Schrägstrich ('/') Elemente in direkter Vater/Sohn-Beziehung, ein doppelter Schrägstrich ('//') bedeutet eine Nachkommen-Beziehung über eine oder auch mehrere Baum-Ebenen. Element-Konstruktor klammert das Ergebnis der Query mit öffnenden und schließenden Tags.

Da XML-Fragmente bereits gültige XQuery-Ausdrücke darstellen, müssen sie durch geschweifte Klammern ({,}) gekennzeichnet werden, andernfalls würden sie wie normaler Text behandelt und ausgegeben.

- Liste alle Dozenten, die Vorlesungen mit mehr als 2 SWS halten:

```
<aktive-Dozenten>
  {FOR $d IN distinct(document("dozenten.xml")//dozent)
  LET $v := document("vorlesungen.xml")//vorlesung[gelesenVon=$d/PersNr]
  WHERE $v/SWS > 2
  RETURN <Dozent>$d/Name</Dozent>}
</aktive-Dozenten>
```

  Das Ergebnis würde lauten:

```
<aktive-Dozenten>
  <Dozent>
    <name>
      Sokrates
    </name>
  </Dozent>
  <Dozent>
    <name>
      Kant
    </name>
  </Dozent>
  <Dozent>
    <name>
      Russel
    </name>
  </Dozent>
</aktive-Dozenten>
```

- Liste alle Dozenten zusammen mit ihren Lehrveranstaltungen:

```
<Vorlesungsverzeichnis>
  {FOR    $d IN document("dozenten.xml")//dozent,
         $v IN document("vorlesungen.xml")//vorlesung
  WHERE  $d/PersNr = $v/gelesenVon
  RETURN <Veranstaltung>$d/Name, $v/Titel</Veranstaltung>}
</Vorlesungsverzeichnis>
```

- Liste zu jedem Dozenten alle seine Lehrveranstaltungen:

```
<Vorlesungsverzeichnis>
  {FOR      $d IN document("dozenten.xml")//dozent
  RETURN
    <Dozent>$d/Name</Dozent>
    {FOR    $v IN document("vorlesungen.xml")//vorlesung
     WHERE  $d/PersNr = $v/gelesenVon
     RETURN <Vorlesung>$v/Titel</Vorlesung>}}
</Vorlesungsverzeichnis>
```

- Liste zu jedem Dozenten die Summe der Semesterwochenstunden seiner Lehrveranstaltungen:

```
<Dozenten>
  {FOR    $d IN document("dozenten.xml")//dozent
   LET    $s := sum(document("vorlesungen.xml")
                  //vorlesung[gelesenVon=$d/PersNr]/SWS)
   WHERE  $s > 0
   RETURN <Dozent>
            <Name>$d/Name</Name>
            <Lehrbelastung>$s</Lehrbelastung>
          </Dozent>}
</Dozenten>
```

- Liste alle Dozenten mit demselben Rang wie Sokrates:

```
<Rang-wie-Sokrates>
  {FOR   $d1 IN document("dozenten.xml")//dozent[Name="Sokrates"],
         $d2 IN document("dozenten.xml")//dozent[Rang=$d1/Rang]
   RETURN <Name>$d2/Name</Name>}
</Rang-wie-Sokrates>
```

10.6 Apache Xindice

Seit November 2002 pflegt die Apache Software Foundation eine Website zu **Xindice**, gesprochen *sin-di-tschee*: `http://xml.apache.org/xindice`. Xindice ist eine XML-Datenbank, die dem Benutzer das Abspeichern und Wiederfinden von XML-Dokumenten und ihren Bestandteilen ermöglicht. Als Abfragesprache wird XPath verwendet (`http://www.w3.org/TR/xpath`).

10.7 Web Services

Bisher wurden Informationen im Internet über von Menschen lesbare Webseiten angeboten; zum Finden der passenden Webseiten helfen Suchmaschinen, wie z.B. Google. Inzwischen existieren auch Angebote in maschinenlesbarer Form, sogenannte *Web Services*, die über genormte Schnittstellen Informationen bereitstellten, die von Programmen abgefragt und ausgewertet werden können. Zum Auffinden dieser maschinenlesbaren Angebote gibt es Verzeichnisdienste, in denen der Anbieter seine Dienste kategorisiert ankündigt.

Mittlerweile existieren folgende auf XML basierende Standards:

1. **UDDI:** Universal Description, Discovery and Integration: Auf dieser Norm sind die Verzeichnisdienste organisiert, mit denen interessierte Clienten sich über angebotetene Dienste informieren und mit deren Hilfe sie die zuständigen WSDL-Dateien finden können.
2. **WSDL:** Web Services Description Language: In dieser Notation kündigt der Diensteanbieter an, welche Art von Anfragen er zulässt, welche Art von Antworten zu erwarten sind und an welchen Server die Anfragen gerichtet werden können. Dieses Dokument kann, aber muss nicht in einem UDDI-Verzeichnisdienst gepflegt werden und muss auch nicht auf demselben Server liegen, der den Dienst bereitstellt.
3. **SOAP:** Simple Object Access Protokoll: Hiermit sendet der anfragende Client gemäß der im WSDL-Dokument genannten Spezifikation einen Request in Form eines XML-Dokuments an den anbietenden Dienst und erhält von dort die Antwort, wiederum als XML-Dokument.

Die Firmen Google und Amazon bieten beispielsweise Web Services an, welche dieselben Anfragen erlauben, die auch über die Menschen-lesbare Webseite möglich sind. Details sind zu finden unter `http://www.google.com/apis` bzw.

```
http://www.amazon.com/gp/aws/sdk/002-5975367-0251255?v=2005%2d03%2d23&
s=AWSEcommerceService
```

Die Amazon-WSDL-Datei lautet
`http://soap.amazon.com/schemas2/AmazonWebServices.wsdl`.

Wenn man sich beim Amazon Web Service registrieren lässt, erhält man eine Subscription-ID, mit der man sich in Anfrage authentifizieren kann. Z.B. kann die Suche nach *Harry Potter* wie folgt an den Amazon Web Service gestellt werden:

```
http://webservices.amazon.de/onca/xml?Service=AWSECommerceService&
SubscriptionId=0BB9F8ZNDA7HG04QG1G2&Operation=ItemSearch&
SearchIndex=Books&Keywords=Harry%20Potter&ResponseGroup=Small
```

Das Testen von Web Services unter Angabe einer WSDL-Datei wird angeboten von
`http://www.mindreef.net`.

Ein einfacher Verzeichnisdienst liegt unter `http://www.xmethods.net`.

Zur Demonstration wurde auf dem Apache Tomcat Server `terra` im Informatik-Netz ein sehr einfacher Web Service installiert, welcher folgende Funktionalität bietet:

> Die Anfrage `GetLehrumfangVonProfessorRequest`, versehen mit einem `ProfName` wird beantwortet als `GetLehrUmfangVonProfessorResponse`, versehen mit dem `LehrUmfang`.

Die zugehörige WSDL-Datei liegt unter
`http://www-lehre.inf.uos.de/~dbs/2005/WS/UniVerwaltung.wsdl`

Listing 10.19 zeigt den Quelltext der Klasse InquireDB.java, welche vom Apache Tomcat Server beim Eintreffen einer Anfrage an den Webservice gestartet wird.

```
import java.sql.*;

public class InquireDB {

  public static int getLehrUmfangVonProfessor(String profName) {

    int sumsws    = 0;
    String url    = "jdbc:microsoft:sqlserver://maestro:1433";
    String user   = "erika";
    String pwd    = "mustermann";

    try {

      Class.forName("com.microsoft.jdbc.sqlserver.SQLServerDriver");
      Connection con = DriverManager.getConnection(url, user, pwd);

      PreparedStatement stmt = con.prepareStatement(
        "SELECT SUM(v.SWS) AS SUMSWS " +
        "FROM Vorlesungen v, Professoren p " +
        "WHERE v.gelesenVon = p.PersNr " +
        "AND p.Name = ?");

      stmt.setString(1, profName);
      ResultSet rset = stmt.executeQuery();
      rset.next();
      sumsws = java.lang.Integer.parseInt(rset.getString("SUMSWS"));
      rset.close();
      stmt.close();
      con.close();
    } catch (Exception e) {e.printStackTrace();}

    return sumsws;

  }
}
```

Listing 10.19: InquireDB.java

Listing 10.20 zeigt den Quelltext des WSDL-Dokuments `UniVerwaltung.wsdl`, in dem die Funktionalität angekündigt wird.

```
<?xml version="1.0" ?>
<definitions name="UniVerwaltung"
   targetNamespace="http://www-lehre/inf.uos.de/~dbs/2005/WS/UniVerwaltung.wsdl"
   xmlns:tns="http://www-lehre/inf.uos.de/~dbs/2005/WS/UniVerwaltung.wsdl"
   xmlns:xsd="http://www.w3.org/2001/XMLSchema"
   xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
   xmlns="http://schemas.xmlsoap.org/wsdl/">

  <message name="GetLehrUmfangVonProfessorRequest">
    <part name="ProfName" type="xsd:string"/>
  </message>
  <message name="GetLehrUmfangVonProfessorResponse">
    <part name="LehrUmfang" type="xsd:int"/>
  </message>

  <portType name="UniVerwaltungPortType">
    <operation name="getLehrUmfangVonProfessor">
      <input message="tns:GetLehrUmfangVonProfessorRequest"/>
      <output message="tns:GetLehrUmfangVonProfessorResponse"/>
    </operation>
  </portType>

  <binding name="UniVerwaltungSOAPBinding" type="tns:UniVerwaltungPortType">
    <soap:binding style="rpc" transport="http://schemas.xmlsoap.org/soap/http"/>

    <operation name="getLehrUmfangVonProfessor">
      <soap:operation soapAction=""/>
      <input>
        <soap:body use="encoded"  namespace="UniVerwaltung"
                   encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"/>
      </input>
      <output>
        <soap:body use="encoded"  namespace="UniVerwaltung"
                   encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"/>
      </output>
    </operation>
  </binding>

  <service name="UniVerwaltungService">
    <port name="UniVerwaltung" binding="tns:UniVerwaltungSOAPBinding">
      <soap:address location="http://terra.inf.uos.de/axis/services/UniVerwaltung"/>
    </port>
  </service>

</definitions>
```

Listing 10.20: UniVerwaltung.wsdl

Listing 10.21 zeigt den SOAP-Envelope, in dem die Anfrage nach dem Lehrumfang von Sokrates an den Web Service geschickt wird. Listing 10.22 zeigt den SOAP-Envelope, der die Antwort vom Web Service enthält.

```
<?xml version="1.0" encoding="UTF-8"?>
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <ns1:getLehrUmfangVonProfessor
    soapenv:encodingStyle= "http://schemas.xmlsoap.org/soap/encoding/"
    xmlns:ns1="UniVerwaltung">
    <ProfName xsi:type="xsd:string">Sokrates</ProfName>
    </ns1:getLehrUmfangVonProfessor>
  </soapenv:Body>
</soapenv:Envelope>
```

Listing 10.21: Request, an den Web Service gerichtet

```
<?xml version="1.0" encoding="UTF-8"?>
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <ns1:getLehrUmfangVonProfessorResponse
    soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
    xmlns:ns1="UniVerwaltung">
    <LehrUmfang href="#id0"/>
    </ns1:getLehrUmfangVonProfessorResponse>
    <multiRef id="id0"
    soapenc:root="0"
    soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
    xsi:type="xsd:int"
    xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/">10</multiRef>
  </soapenv:Body>
</soapenv:Envelope>
```

Listing 10.22: Response, vom Web Service geliefert

Listing 10.23 zeigt den Quelltext eines "handgestrickten" Klienten `ClientUniVerwaltung.java`, welcher das zur Anfrage "*Welchen Lehrumfang hat Sokrates ?*" passende XML-Dokument als SOAP-Envelope erstellt, verschickt, das Ergebnis wieder als SOAP-Envelope entgegennimmt und ausgibt.

```
import java.io.*;
import java.net.*;

public class ClientUniVerwaltung {
   private static final int BUFF_SIZE = 100;

   public static void main(String[] argv) throws Exception {
      String request =
      "<?xml version='1.0' encoding='UTF-8'?>"+
         "<soapenv:Envelope " +
           "xmlns:soapenv='http://schemas.xmlsoap.org/soap/envelope/' " +
           "xmlns:xsd='http://www.w3.org/2001/XMLSchema' " +
           "xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'> " +
          "<soapenv:Body> "+
            "<ns1:getLehrUmfangVonProfessor " +
               "soapenv:encodingStyle= " +
                  "'http://schemas.xmlsoap.org/soap/encoding/' " +
               "xmlns:ns1='UniVerwaltung'> " +
                 "<ProfName xsi:type='xsd:string'>Sokrates</ProfName>" +
                 "</ns1:getLehrUmfangVonProfessor>" +
          "</soapenv:Body>"+
        "</soapenv:Envelope>";

      URL url = new URL(
          "http://terra.inf.uos.de/axis/services/UniVerwaltung");
      HttpURLConnection conn = (HttpURLConnection) url.openConnection();

      conn.setDoOutput(true); conn.setUseCaches(false);
      conn.setRequestProperty("Accept", "text/xml");
      conn.setRequestProperty("Connection", "keep-alive");
      conn.setRequestProperty("Content-Type", "text/xml");
      conn.setRequestProperty("Content-length",Integer.toString(request.length()));
      conn.setRequestProperty("SOAPAction", "\" \"");

      OutputStream out = conn.getOutputStream();
      out.write(request.getBytes()); out.flush();

      StringBuffer response = new StringBuffer(BUFF_SIZE);
      InputStreamReader in =
         new InputStreamReader(conn.getInputStream(), "UTF-8");
      char buff[] = new char[BUFF_SIZE];   int n;
      while ((n = in.read(buff, 0, BUFF_SIZE - 1)) > 0) {
         response.append(buff, 0, n);
      }
      out.close(); in.close();
      System.out.println( response.toString() );
   }
}
```

Listing 10.23: ClientUniVerwaltung.java