

XML – Aufgaben zu Merkblatt 1

1 Grundbegriffe

Die folgenden Begriffe und gegebenenfalls das englischsprachige Pendant sollten Sie verstehen, einordnen und erklären können: Auszeichnungen, Attribut, de-serialisiertes XML, Dokumentknoten, Element, Elementknoten, End-Tag, Fluchtsymbol, hierarchische Daten, HTML, Inhalt (eines Elements), inneres Element, semi-strukturierte Daten, serialisiertes XML, SGML, Start-Tag, Tag, Tag-Name, wohlgeformt, Wurzelknoten, Wurzelement, XML, XML-Baum, XML-Bezeichner, XML-Deklaration, XML-Dokument, Zeichenreferenz.

2 Tag-Namen

Welche Tag-Namen sind syntaktisch korrekt? Welche davon sollten nicht verwendet werden? `ä` ist wie in HTML auch in XML ein Fluchtsymbol für ä.

Tag-Name	syntaktisch korrekt?	verwenden?
first-name		
7zehn		
M.A.S.H		
a...z		
Anfänger		
überzahl		
b:rechnung		
XlmX		
geld\$oeder\$leben		
_.:7		

3 Baumdarstellung

Geben Sie für das folgende XML-Dokument die Baumdarstellung an. Ein Attribut eines Elements ist ein Attributknoten unterhalb dessen Elementknotens. Der Wert eines Attributs ist ein Textknoten unterhalb des Attributknotens.

```
<?xml version="1.0" encoding="utf-8"?>
<rechneruebung name="XML">
  <aufgabe>
    1. Aufgabe
    <student matrikelnummer="60000">
      <name>Arnold Adams</name>
```

```

        <bestanden/>
    </student>
    <student matrikelnummer="60001">
        <name>Anna Albers</name>
        <bestanden/>
    </student>
</aufgabe>
<aufgabe>
    2. Aufgabe
    <student matrikelnummer="60000">
        <name>Arnold Adams</name>
        <nicht-bestanden/>
    </student>
</aufgabe>
<!-- mehr Aufgaben -->
</rechneruebung>

```

4 Wohlgeformtes XML

Welche der folgenden XML-Dokumente sind wohlgeformt und welche nicht?

```
<Hello>
```

```
<Hello text="World"></Hello>
```

```
<Message>Hallo</Message>
<Message>Welt</Message>
```

```
<a><b><c></b><d/></a>
```

```
<a>
    Ein
    <b>bisschen</b>Text
    <!-- und -->
</a>
<!-- mehr -->
```

5 Text

Was ist der Text der folgenden Element, nachdem das XML eingelesen und dabei die Fluchtsymbole ersetzt wurden? Kommentare stellen wie Attribute einen eigenen Knotentyp dar und haben ihren Text als Inhalt.

```
<a>H A L L O</a>
```

```
<a><![CDATA[[Hi!]]]></a>
```

```
<a>&auml;<![CDATA[nder]]>n</a>
```

```
<a>
  Hallo
</a>
```

```
<a>&#10;xXx&#10;</a>
```