



Semistrukturierte Daten

Sommersemester 2012

Teil 2: Namespaces (Namensräume)

Motivation

- Problem von Namenskonflikten:
 - Bei Kombination von unterschiedlichen Anwendungen in einem einzigen Dokument
 - Konflikt mit "Schlüsselwörtern" eines verwandten Standards von XML, z.B.: XML-Schema, XSLT, XLink, etc.
- Wie kann ein Programm "seine" Anweisungen erkennen?
- Beispiele:
 - XML: lang, space, ...
 - XML Schema: element, sequence, ...
 - XSLT: element, value-of, ..
 - XLink: title, role, ...
- Lösung: **Namespaces** (kurz "NS")

Beispiel: Namenskonflikt zwischen Anwendungen

Studierende führen Bewertungsschema für Lehrveranstaltungen ein:

```
<veranstaltung>  
  <titel>Semistrukturierte Daten</titel>  
  <bewertung>naja...</bewertung>  
</veranstaltung>
```

*HerausgeberIn eines Vorlesungsverzeichnisses fügt selbes Tag ein
(mit anderer Bedeutung!):*

```
<veranstaltung>  
  <titel>Semistrukturierte Daten</titel>  
  <bewertung>Pflichtfach</bewertung>  
</veranstaltung>
```

Problem: Kombination? Umbenennung?

Lösung: Namespaces

Namespaces

- ursprünglich Namespaces in eigener Recommendation, nachträglich zu XML hinzugefügt
- Eindeutige Identifier sind notwendig
 - URLs (allg. URIs) eignen sich dafür; hier gibt es schon Eindeutigkeit
 - URL dient nur als Identifier, keine Referenz!
 - für NS sind `www.dbai.tuwien.ac.at` und `www.DBAI.tuwien.ac.at` nicht identisch
- Beispiel-Namespaces:
 - HTML 4.0: <http://www.w3.org/TR/REC-html40>
 - XML Schema: <http://www.w3.org/2001/XMLSchema>
 - XSLT: <http://www.w3.org/1999/XSL/Transform>

Beispiel: Namespaces

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

```
<lehre xmlns:stud="http://www.oeh.ac.at/bew/"  
      xmlns:vorl="http://tuwis.tuwien.ac.at/lehre/2.0"  
      xmlns      ="http://tuwis.tuwien.ac.at/lehre/2.0">
```

...

```
<veranstaltung vorl:jahr="2012">  
  <titel>Semistrukturierte Daten</titel>  
  <stud:bewertung>naja...</stud:bewertung>  
  <vorl:bewertung>Wahlpflichtfach</vorl:bewertung>  
</veranstaltung>
```

...

```
</lehre>
```

Namespaces

■ "Umbenennung"

- Anstatt eines URI geben wir Abkürzung für den Identifier an
- dient nur als Abkürzung (versch. Abkürzungen, die auf selben URI zeigen, definieren selben NS)
- Name (des Elements bzw. des Attributs):
präfix:lokaler-name (getrennt durch Doppelpunkt)

■ zusätzlich Deklaration der Präfixe

- im Element-Starttag als Attribut angegeben: gilt für das Element selbst und alle Subelemente
- assoziiert URI mit Präfix
- URI sinnvoll, da URIs immer eindeutig
- Präfix sinnvoll, da URI meistens Sonderzeichen (wie "/") enthält.
- xmlns Attribut ohne NS Angabe erklärt *Default Namespace*

-
- | Government | Percentage |
|---------------------|------------|
| Current government | 85% |
| Previous government | 15% |

Beispiel: Namespaces

```
<?xml version="1.0"?>
<lehre xmlns:stud="http://www.oeh.ac.at/bew/"
      xmlns:vorl="http://tuwis.tuwien.ac.at/lehre/2.0"
      xmlns="http://tuwis.tuwien.ac.at/lehre/2.0">
...
  <veranstaltung xmlns="http://www.wu-wien.ac.at/WINF">
    <titel>Semistrukturierte Daten 1</titel>
    <stud:bewertung>naja...</stud:bewertung>
    <vorl:bewertung>Wahlpflichtfach</vorl:bewertung>
  </veranstaltung>
...
</lehre>
```

Diagram illustrating namespace resolution:

- Arrows point from the `<veranstaltung>` and `<titel>` elements to the text: **liegen im neu definierten Default-NS**.

URIs

■ URIs (Uniform Resource Identifier)

- **URLs** (Uniform Resource Locator): Internetadressen nach „traditionellem“ Adressierungsschema, z.B.: `http://`, `ftp://`, etc.
- **URNs** (Uniform Resource Name): generische Namen, müssen eindeutig sein, z.B. **`urn:ISBN:3-8274-0130-5`**

■ Terminologie

- Präfix **zeigt** auf NS(-Name), z.B. **stud** zeigt auf **`http://www.oeh.ac.at/bew`**
- Element **liegt in** Namespace
- Lokaler Name: **bewertung**
- Qualifizierter Name: [Präfix:]lokaler Name
- Expandierter Name: [NS-Name,]lokaler Name

