**Publizieren Vorlesung 5**

**Datenmodelle** für XML-Dokumente

* XML Information Set
* XDM

**Lokalisierung** von XML-Fragmenten

* XPath

**XML Information Set**

* Definiert ein Datenmodell für XML-Dokumente
* in Form von Komponenten (Information Items)
* mit Attributen für nähere Beschreibung der Komponenten selbst und der Beziehungen

**Datenmodell**

* legt fest, welche Bestandteile eines XML-Dokuments aus Sicht des Datenmodells wesentlich sind (filterndes Abbild der Realität)
* abstrahiert von XML-Syntax
* dient der terminologischen Klarheit und Präzision
* Definiert durch Information Items und ihre Attribute
  + Information Set entspricht Baumstruktur
  + Information Item entspricht Komponente im Baum

**Datenmodell XDM**

Basiert auf XML Infoset als konzeptuelle Sicht auf XML-Dokumente als Strukturen von Knoten / Information Items

* Elementknoten
* Attributknoten
* Textknoten (bzw. Zeichenknoten)
* Namespaceknoten

Datum in XDM ist **Sequenz** von

* **einfachen Daten** vom Typ number, string, boolean, ...
* **Referenzen** auf Komponenten eines XML-Dokuments:

Knoten / Information Items im Sinne von XML Information Set

**Aktivität**

Aspekte der Syntax von XML-Dokumenten, von denen ein XML Information Set abstrahiert

* XML-Deklaration ()
* Begrenzer von Attributwerten (" oder ')
* Kodierung von leeren Elemente
* Leerzeichen in Tags
* Reihenfolge der Attribute

**XPath**

* Teilsprache von **XQuery**
* **Hauptfunktion**:
  + Adressierung von Teilen von XML-Dokumenten, durch **Navigation** durch die hierarchische Struktur (Mini-Query-Sprache)
    - eingängig, auch durch Analogie bei Directories
    - eher prozedural als deklarativ
  + weitere Funktion: Berechnungen allgemeiner Art
* **verwendet** **in**:
  + XML Schema Identity Constraints, Schematron, XSLT, XLink, XForms
* Basiert (wie auch XSLT, XQuery) auf Datenmodell XDM

**Ausdruckstyp Location Path**

* Teilsprache von XPath
* beschreibt schrittweises Navigieren durch Dokumentenbaum
* Aufbau von Step: drei Teile
  + Achse (Richtung für den Schritt), z.B. Kindachse, Elternachse, Attributachse
  + Knotentest (für Typ und Namen)
  + Prädikate mit beliebigen Filter-Ausdrücken
* Auswertung eines Location Path Ausdrucks
  + geht von Kontextknoten aus
  + nimmt in jedem Schritt lokale Filterung vor, mit Knotensequenz als Zwischenergebnis
  + führt zu Knotensequenz als Ergebnis

Syntax für Step:

* **Achse::Knotentest [ Prädikat ] [ Prädikat ] [ Prädikat ] ...**

Syntax für Pfadausdruck:

* Step / Step / … oder / Step / Step ...

**Abkürzungen**

<nichts> child::

@ attribute::

// /descendant-or-self::node()/

. self::node()

.. parent::node()

Welche Datenmodelle für XML gibt es als W3C-Standard?

* XQuery, XPath, XML Information Set
* XDM definiert ein Datenmodell für die Daten, auf denen XPath, XQuery, XSLT operieren

