

Technik digitaler Editionen mit XML/XSLT

Susanne Haaf

3. Münchner Digital Humanities Summerschool
11.-15. September 2017

XSLT benutzen

- DTABf-XSLT:

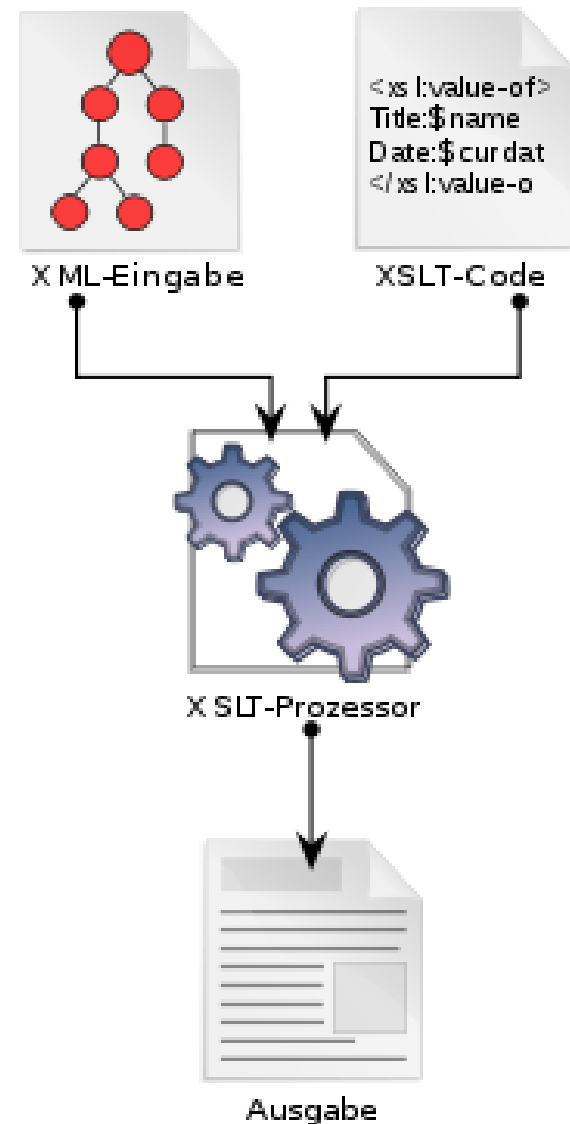
<https://raw.githubusercontent.com/haoess/dta-tools/master/stylesheets/dtabf.xsl>

- Übungs-XSLT:

http://kaskade.dwds.de/~haaf/dhmuc/04_XSL/uebung.xsl

Einführung in XSL(T)

- XSL: Extensible Stylesheet Language
- XSLT: XSL Transformation



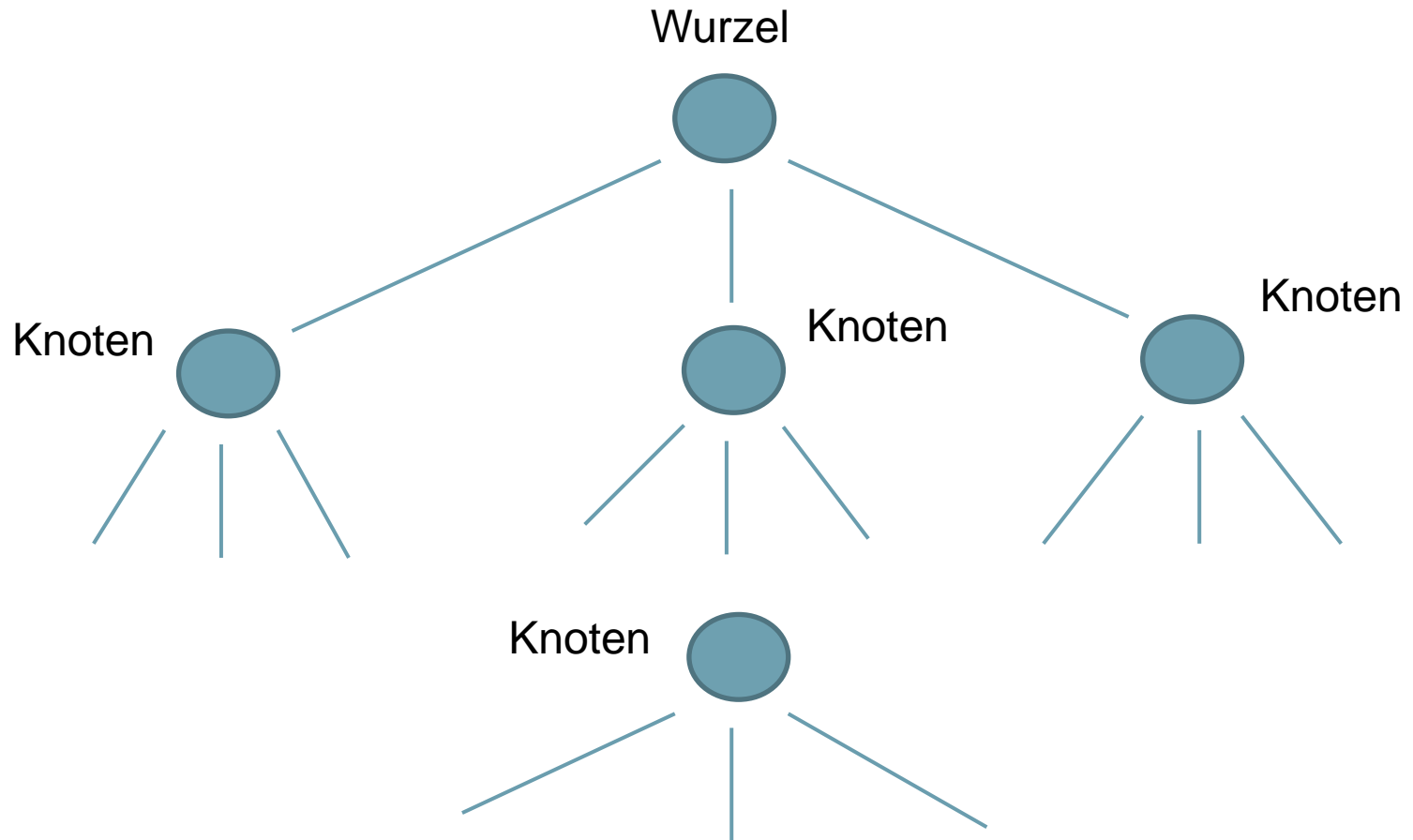
XSL Basics ...

XPath

- “The primary purpose of XPath is to address parts of an XML document.” – d.h. wird in XSLT und XQuery verwendet
- “In support of this primary purpose, it also provides basic facilities for manipulation of strings, numbers and booleans.”
- “XPath operates on the abstract, logical structure of an XML document [...]”
- “XPath gets its name from its use of a path notation as in URLs for navigating through the hierarchical structure of an XML document.”

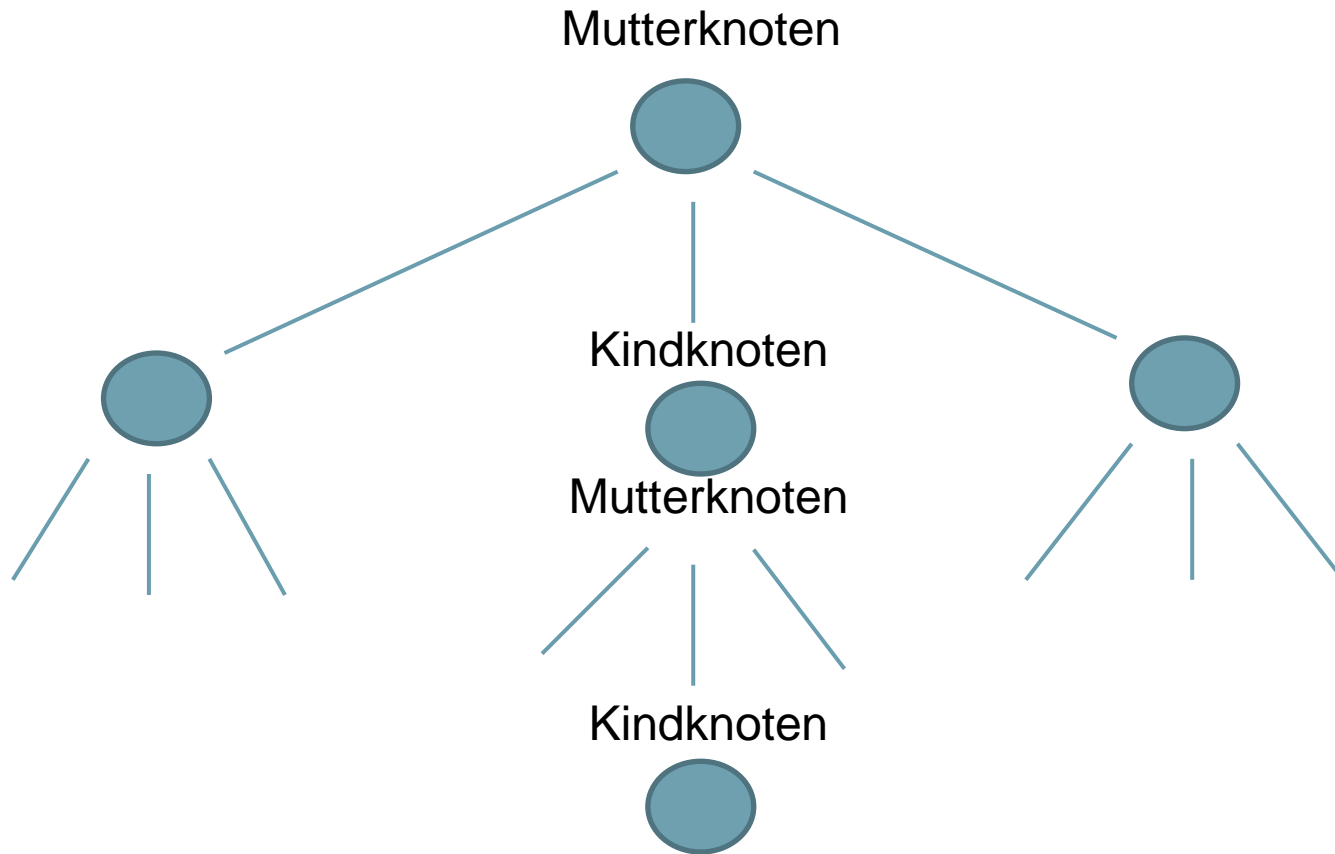
Zitate aus: W3C Recommendation for XPath: <http://www.w3.org/TR/xpath/>

XPath Grundlage: XML-Baum



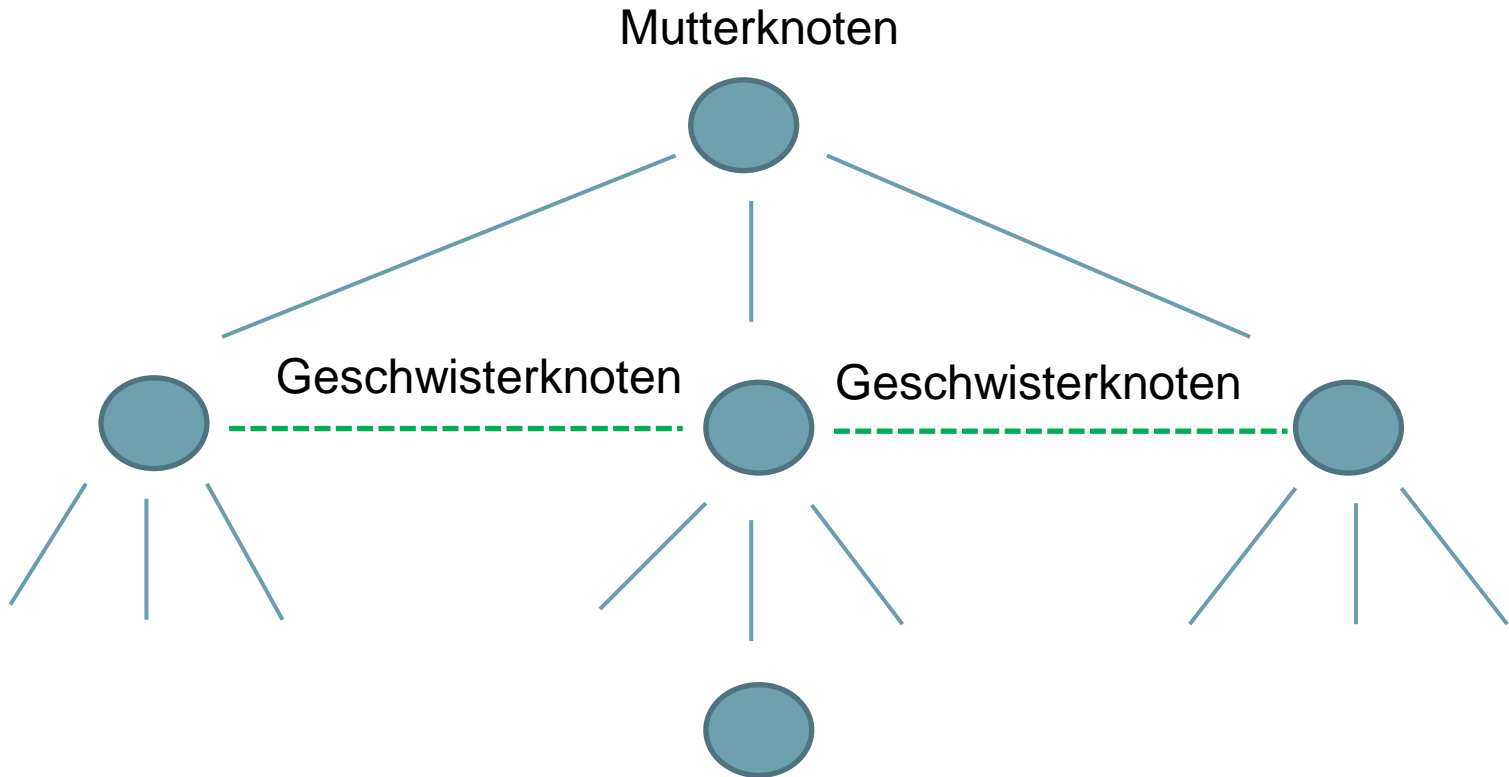
XPath: // (= relativer Pfad)

parent:: und child::



XPath: `//Kindknoten/parent::Mutterknoten`
XPath: `//Mutterknoten/child::Kindknoten`

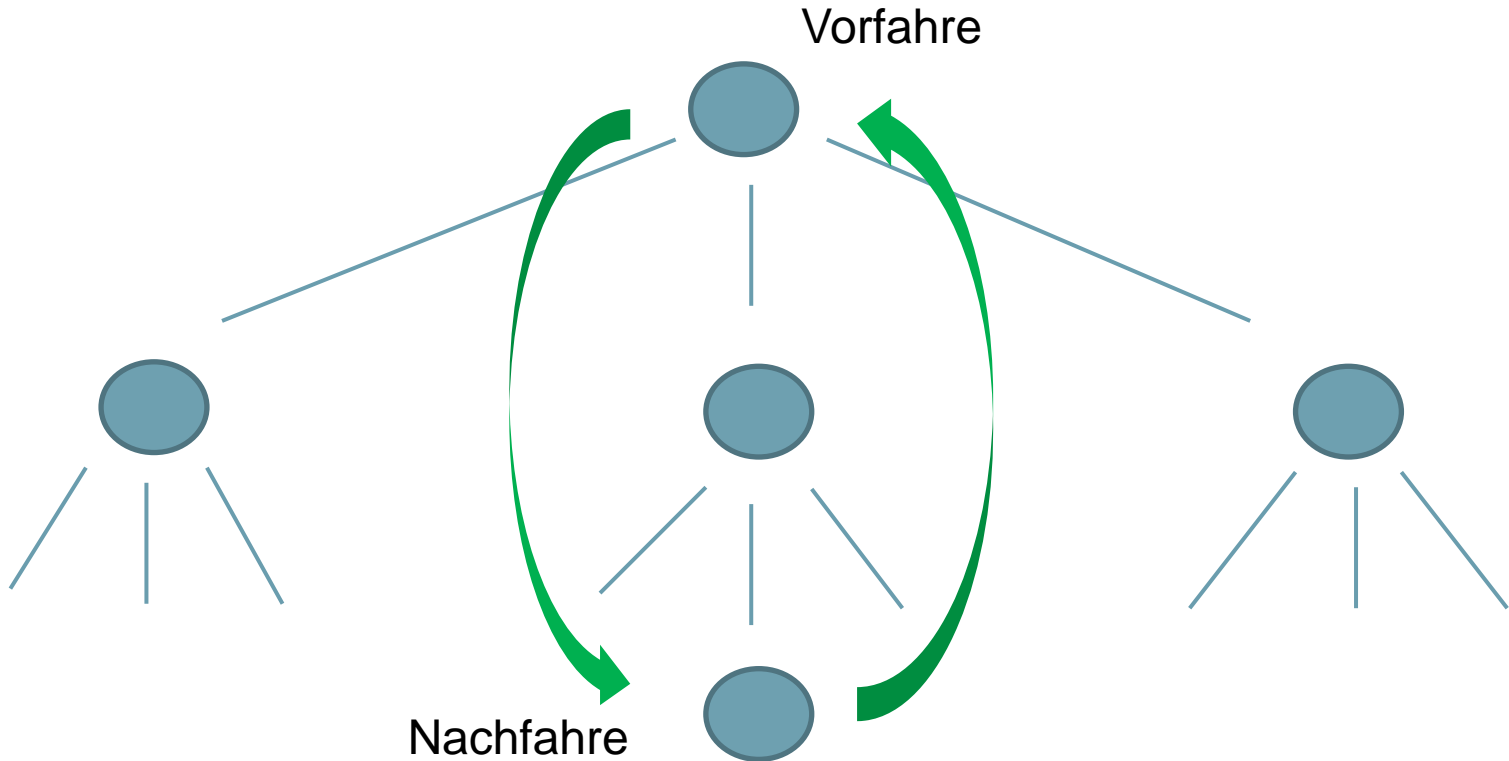
preceding-sibling:: und following-sibling::



XPath: `//Knoten/preceding-sibling::Knoten`

XPath: `//Knoten/following-sibling::Knoten`

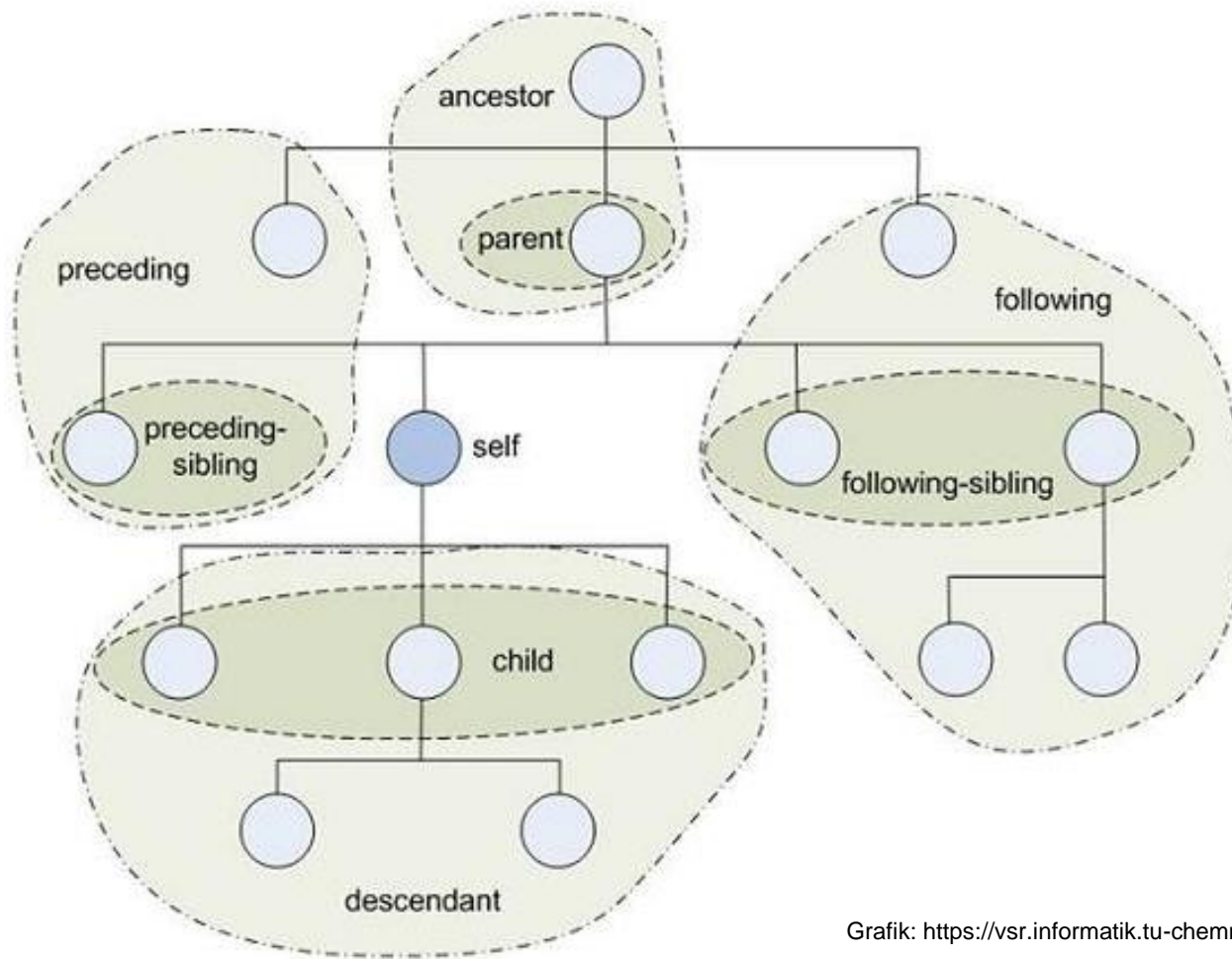
ancestor:: und descendant::



XPath: `//Knoten/descendant::` spätererKnoten

XPath: `//Knoten/ancestor::` frühererKnoten

XPath-Achsen



Grafik: <https://vsr.informatik.tu-chemnitz.de/proseminare/xml04/doku/xpath/>

XPath-Syntax

- Syntax von XPath ähnelt derjenigen, die für die Pfadangaben von Dateiordnern verwendet wird
- Beispiel: `/TEI/teiHeader/fileDesc/titleStmt/title`
- XPath testen in oXygen XML über
 1. XPath-Eingabe-Feld in der Toolbar
 2. Fenster > Ansicht zeigen > XPath/XQuery Builder

XPath-Syntax

| Syntax | Bedeutung | Beispiel |
|--------|----------------------------------|---|
| / | Pfadtrennzeichen | /TEI/teiHeader/fileDesc/titleStmt/titl e |
| // | Verkürzte Notation für Pfad | /TEI//titleStmt/title //title |
| @ | Attributname | //publicationStmt/date/@when |
| * | Wildcard – liefert alle Elemente | //publicationStmt/* |
| [] | Bedingung | //date[@when="1806-10-01"] //p[1] |
| text() | nur Text, keine Elemente | //titleStmt/text() |
| | | |

XPath-Achsen

| Achse | Abkürzung | Beschreibung |
|--------------------|-----------|---|
| self | . | Liefert den Kontextknoten |
| child | | Liefert die Kinder des Kontextknoten |
| parent | .. | Liefert die Eltern des Kontextknoten, falls vorhanden |
| descendant | // | Liefert die Nachkommen des Kontextknoten |
| descendant-or-self | | Liefert den Kontextknoten und die descendant-Achse |
| ancestor | | Liefert die Vorfahren des Kontextknoten bis hin zum Wurzelknoten |
| ancestor-or-self | | Liefert den Kontextknoten und die ancestor-Achse |
| following | | Liefert alle Elementknoten, die in der Dokumentreihenfolge dem Kontextknoten <i>folgen</i> , mit Ausnahme der Nachkommen des Kontextknotens |
| following-sibling | | Liefert alle <i>folgenden</i> Geschwister des Kontextknoten |
| preceding | | Liefert alle Elementknoten, die in der Dokumentreihenfolge dem Kontextknoten <i>vorangehen</i> , mit Ausnahme der Nachkommen des Kontextknotens |
| preceding-sibling | | Liefert alle <i>vorhergehenden</i> Geschwister des Kontextknoten |

XPath – Wichtige Operatoren

| Operator | Bedeutung |
|----------|---|
| = | gleich (Vergleich zweier Werte) |
| != | ungleich (Vergleich zweier Werte) |
| < | kleiner als (Vergleich zweier Werte). Maskierung innerhalb von XSLT erforderlich. |
| > | größer als (Vergleich zweier Werte). Maskierung innerhalb von XSLT zu empfehlen. |
| <= | kleiner als oder gleich (Vergleich zweier Werte). Maskierung innerhalb von XSLT erforderlich. |
| >= | größer als oder gleich (Vergleich zweier Werte). Maskierung innerhalb von XSLT zu empfehlen. |
| and | logische Und-Verknüpfung (beide Ausdrücke) |
| or | logische Oder-Verknüpfung (einer von beiden Ausdrücken) |

Beispiel: `//date[@when >= '1900-01-01']`

XPath-Funktionen

| Funktion | Beschreibung | Beispiel |
|--------------------|---|--|
| starts-with() | Ergibt den Wert <i>wahr</i> , wenn die zweite Zeichenkette in der ersten enthalten ist. | <code>//placeName[starts-with(., 'W')]</code> |
| contains() | Ergibt den Wert <i>wahr</i> , wenn die zweite Zeichenkette in der ersten enthalten ist | <code>//placeName[contains(., 'Hall')]</code> |
| substring-after() | Liefert aus der ersten Zeichenkette das, was <i>hinter</i> dem Teil steht, der mit der zweiten Zeichenkette übereinstimmt | <code>//licence/substring-after(., 'CC-BY-SA ')</code> |
| substring-before() | Liefert aus der ersten Zeichenkette das, was <i>vor</i> dem Teil steht, der mit der zweiten Zeichenkette übereinstimmt | <code>//p/substring-before(., ' ')</code> |