

Vergleich von XPath, XSLT und XQuery

Philipp Pobitzer

23. Mai 2015

1 Einleitung

Diese Arbeit gibt einen kurzen Einblick in XML und vergleicht die darauf aufbauenden Sprachen XPath, XSLT und XQuery, die vom W3C definiert wurden. Sie soll des Weiteren die Anwendungen der Sprache und gegebenenfalls deren Zusammenhang aufzeigen. Sie gibt keine Einführung oder Anleitung zum Umgang mit diesen Sprachen. Dafür sind im Schlussteil Literaturempfehlungen hierzu zu finden.

2 XML

Die Sprachen XPath, XSLT und XQuery bauen auf XML auf. Daher möchte ich zuerst in aller Kürze XML ansprechen. Der *Austausch von Daten* zwischen Programmen, auch zwischen verschiedenen Plattformen, ist eine sehr wichtige Funktionalität. XML wurde für diesen Zweck 1998 geschaffen und stellt heute einen wichtigen Standard dar[5]. Ein weiterer Aspekt ist, dass Daten von der Repräsentation getrennt werden. Dies wurde erforderlich, damit eine Kompatibilität zwischen den verschiedensten Plattformen möglich ist[1]. Anhand von 1, ist es ersichtlich wie Informationen anhand von Elementen in einer XML Datei abgespeichert werden.

3 XPath

Im Jahr 1999 wurde die Abfragesprache XPath 1.0 zusammen mit XSLT 1.0 veröffentlicht. Sie wurde als Subsprache für XSLT entwickelt und wird heute

```

<contacts>
  <person>
    <age>22</age>
    <sex>female</sex>
    <name>Julia</name>
  </person>
  <person>
    <age>27</age>
    <sex>male</sex>
    <name>Max</name>
  </person>
</contacts>

```

Listing 1: XML Beispiel

in vielen anderen Sprachen eingebunden. XPath selbst kann hingegen keine andere Sprache einbinden. Ein XPath Ausdruck kann unter anderem dafür benützt werden, um numerische Berechnungen durchzuführen, Strings zu berechnen und/oder boolesche Ausdrücke auszuwerten. Die wichtigste Funktion jedoch ist die *Auswahl von Knotenmengen und deren Adressierung*, woher sich auch der Name ableiten lässt[1]. XPath verwandelt eine XML Datei in eine Baumstruktur, um darauf zu arbeiten. Bei der Bearbeitung der Knoten wird das Prinzip der Tiefensuche angewandt[5]. Es wird auch in XQuery verwendet und die daraus resultierenden Anforderungen trugen maßgeblich zu der Definition von XPath 2.0 aus dem Jahr 2007 bei[4]. Als Beispiel der Ausdruck `/contacts/person/age` liefert die Knotenmenge aller Elemente `age`, was aus 1 zwei Knoten liefern würde.

4 XSLT

Das Akronym steht für Extensible Stylesheet Language: Transformation. Die Syntax entspricht der von XML. Die Aufgabe von XSLT ist *die Struktur und den Inhalt einer XML Datei zu transformieren*. Dabei steht XPath zur Verfügung. Dabei wird die originale Datei nicht modifiziert. Die Anwendungen sind hier sehr zahlreich. So ist es unter anderen möglich als Ausgabe PDF, Sounddateien oder XML zu erzeugen. Die wichtigste und häufigste Transformation ist jedoch nach HTML. Mit XSLT werden damit Daten auch mit einer

Repräsentation verbunden[1]. Bei der Umwandlung zu XML ist es möglich zwischen verschiedenen XML-Vokabularen zu wechseln und die Quelldaten auszuwählen, umzusortieren und/oder zu erweitern. *XSLT ist keine read only Sprache wie XPath*, sondern kann die Daten auch verändern[5]. Zudem ist die Sprache turing-vollständig.¹

5 XQuery

Die Abfragesprache wurde 2007 veröffentlicht. Sie sollte eine dem *SQL ähnlich einfache Abfragen von XML Dateien* erlauben. Dies spiegelt sich in den Ähnlichkeiten der Syntax wieder. Als Subsprache verwendet es ebenfalls XPath[2]. Aus der Zusammenarbeit der Arbeitsgruppen von XPath 2.0, XSLT 2.0 und XQuery 1.0 ergibt sich auch, dass alle drei Sprachen das gleiche Datenmodell XDM (XQuery and XPath Data Model) verwenden. Die Ähnlichkeit zu SQL wird gut anhand der wichtigsten Klauseln in XQuery verdeutlicht. Die sogenannten FLWOR², ein Akronym der entsprechenden Anfangsbuchstaben: for, let, where, order, return[5]. Sie entsprechen in gewisser Weise dem SELECT-FROM-WHERE aus SQL, wobei SELECT von XPath übernommen wird. Es entstand, um aus großen Mengen von XML Daten bestimmte Datensätze zu finden. Es können Datensätze manipuliert werden. Die Originaldaten werden jedoch nicht verändert, sondern eine Ausgabedatei, häufig eine XML Datei, erstellt. Es gibt allerdings Erweiterungen, um diese Funktionalität bereitzustellen. Die Sprache ist ebenfalls turing-vollständig³ und funktional[3].

6 Zusammenfassung

Alle drei Sprache erfüllen grundsätzlich einen durchaus unterschiedlichen Zweck. XPath ist zur Adressierung in beiden anderen enthalten und wurde maßgeblich durch die Anforderungen bei den Weiterentwicklungen der Sprachen beeinflusst. XSLT und XQuery haben, obwohl sie für unterschiedliche Aufgaben entworfen wurden, durchaus eine Schnittmenge, in der beide

¹<http://conferences.idealliance.org/extreme/html/2004/Kepser01/EML2004Kepser01.html>
(23. Mai 2015)

²gesprochen wie englisch »flower«

³<http://conferences.idealliance.org/extreme/html/2004/Kepser01/EML2004Kepser01.html>
(23. Mai 2015)

Sprachen für eine Aufgabe benützt werden. So erfüllt XSLT 1.0 alle Anforderungen der ersten und zweiten Anforderungsspezifikation aus dem Jahr 2000 von XQuery[1]. XQuery wird mehr im Umfeld von Datenbanken eingesetzt. Da XSLT schon länger verfügbar ist, hat es eine größere Nutzerbasis als XQuery. Für kleinere Datenmengen wird häufig XSLT eingesetzt. Jedoch ist die Wahl der Sprache wie Michael Kay, ein Mitglied der XSLT Gruppe beim W3C, in [1] beschreibt, in diesem Bereich, mehr die Frage des Stils, als wirklich der Substanz. Abschließend möchte ich noch auf die Bücher [5], als Einführung in die gesamte Materie rund um XML (inklusive XPath, XSLT und XQuery), und [1], welches eine gute Einführung und Vertiefung in XPath und XSLT gibt, verweisen.

Literatur

- [1] Michael Kay. *XSLT 2.0 and XPath 2.0; programmer's reference*. Wrox programmer to programmer. Wiley, Indianapolis, Ind., 4. ed. edition, 2008.
- [2] Jim Melton and Stephen Buxton. Chapter 10 - introduction to {XQuery} 1.0. In Jim Melton and Stephen Buxton, editors, *Querying {XML}*, The Morgan Kaufmann Series in Data Management Systems, pages 261 – 328. Morgan Kaufmann, Burlington, 2006.
- [3] Jim Melton and Stephen Buxton. Chapter 11 - {XQuery} 1.0 definition. In Jim Melton and Stephen Buxton, editors, *Querying {XML}*, The Morgan Kaufmann Series in Data Management Systems, pages 329 – 406. Morgan Kaufmann, Burlington, 2006.
- [4] Jim Melton and Stephen Buxton. Chapter 9 - {XPath} 1.0 and {XPath} 2.0. In Jim Melton and Stephen Buxton, editors, *Querying {XML}*, The Morgan Kaufmann Series in Data Management Systems, pages 215 – 259. Morgan Kaufmann, Burlington, 2006.
- [5] Helmut Vonhoegen. *Einstieg in XML; [Grundlagen, Praxis, Referenz ; für Anwendungsentwicklung und E-Publishing ; Transformation, Formattierung, Schnittstellen ; inkl. XML Schema, DTD, XSLT, CSS, XSL, XPath, DOM, SAX, SOAP, XQuery, XForms]*. Galileo Computing. Galileo Press, Bonn, 6., aktualisierte und erw. aufl. edition, 2011.