

XML Anfragesprachen

Martin Steffen

Universitetet i Oslo

Zusammenfassung. *XML* (“ Extensible Markup Language”) ist ein etablierter Standard zur Modellierung und zum Austausch (semi-)strukturierter Daten. Gestützt auf eine Vielzahl von Werkzeugen und zusammen mit der Buchstabensuppe seiner Begleitsprachen und -technologien wie z.B. XPath, XML-Schema, XSLT und vielen Anderen mehr, gilt XML heute als universelles Format für den Datenaustausch.

Daten sollen nicht nur repräsentiert, gespeichert (z.B., in Datenbanken) oder ausgetauscht werden, sie wollen auch gefunden und aufbereitet sein. *Anfragesprachen* erlauben es, Fragen und Suchkriterien für XML-repräsentierte, das bedeutet im Kern, *baumartig* strukturierte Daten zu formulieren und so diese zu selektieren, zu transformieren und zu kombinieren. XMLs baumstrukturiertes Datenmodell läßt sich als natürliche Verallgemeinerung des relationalen Datenmodells traditioneller Datenbanken auffassen, eine Verallgemeinerung eben auf Bäume anstelle von Datenbanktabellen. Mithin kann man *XML-Anfragesprachen* als die entsprechende Übertragung der Ideen traditioneller Datenbankanfragesprachen wie SQL auf baumrepräsentierte Daten verstehen.

Aufgrund der Vielfalt unterschiedlicher Daten im Internet und der großen wirtschaftlichen Bedeutung der Standardisierung des Datenaustausches nimmt es nicht Wunder, daß eine Vielzahl von Vorschlägen für XML-(Anfrage)-Sprachen existiert, und zwar sowohl seitens der W3C-Gremien, dem Konsortium für die “Standardisierungen im WWW”, als auch aus dem akademischen Bereich.

Der Vortrag stellt die Prinzipien gängiger Vertreter von XML-Anfragesprachen vor, speziell von *XQuery* und *XSLT*, zweier aktueller W3C-Vorschläge, und ihrer Teilsprachen. Darüber hinaus geht der Vortrag auf moderne, typbasierte Ansätze für Anfragesprachen ein, die auf statischer Typüberprüfung, Typinferenz und “Pattern Matching” basieren.