

Johannes Gutenberg-Universität Mainz  
Fachbereich Informatik  
Seminar Internet-Technologie  
Seminarleitung: Prof. Dr. Herbert Göttler, Thomas Gottron  
Referent: Andreas Wolf

## **Ziele**

Ein Überblick über XML und die wichtigsten zu XML gehörenden Techniken geben.  
Einige ausgewählte XML Anwendungen vorstellen.

## **Einführung und Überblick**

XML steht für Extensible Markup Language und ist eine Meta-Auszeichnungssprache mit deren Hilfe man andere Auszeichnungssprachen definieren kann. In XML werden Inhalt und Layout eines Dokuments getrennt.

XML besteht aus einer ganzen Familie von Techniken wie z.B. XML, DTD, XML Schema, XSL, XPath, XML Namespaces, XLink, XPointer, und XQuery.

In DTDs und XML Schemata werden die *Grammatiken* der XML-Anwendungen definiert. Mit XSL kann man das Layout eines Dokuments definieren.

## **SGML**

Wie fast alles, hat XML natürlich einen Vorgänger, SGML(Standard Generalized Markup Language) und dieses natürlich auch wieder und so weiter.

1967 beschäftigte sich das Generic Coding Project der Graphic Communications Association mit der Trennung von Struktur und Layout bei Publizierungssystemen.

1969 wird von Charles Goldfarb, Edward Mosher und Raymond Lorie GML (Generalized Markup Language) für IBM entwickelt.

1978 Charles Goldfarb arbeitet im Auftrag des American National Standards Institute an einem Projekt für eine Textbeschreibungssprache nach dem Vorbild von GML

1986 SGML wird zum ISO-Standard erklärt ISO 8879

1996 DSSSL (Document Style Semantics and Specification Language) nach 10jähriger Entwicklungszeit wurde ein Standard für eine Style-Sprache geschaffen ISO 10179

## **DTD**

Die Dokumenttypdefinition (englisch Document Type Definition (DTD)) ist eine Deklaration in SGML- und XML-Dokumenten, die die Struktur eines solchen Dokuments festlegt.

## **XML Schema**

XML-Schema ist eine Empfehlung des W3C zum Definieren von XML-Dokumentstrukturen. Anders als bei den klassischen XML-DTDs wird die Struktur in Form eines XML-Dokuments beschrieben. Darüber hinaus wird eine große Anzahl von Datentypen unterstützt.

XML Schema soll in Zukunft einmal DTD ersetzen.

## **XSL**

XSL (Extensible Stylesheet Language) ist eine Familie von Sprachen zur Erzeugung von Layouts für XML-Dokumente. Der XSL-Dialekt XSLT kann auch zur Übersetzung/Transformation eines XML-Dialektes in ein anderes Format dienen (meistens in andere XML-Dialekte). Diese Layouts (auch Stylesheets genannt) können in die zu formatierenden XML-Dokumente eingebunden werden, wobei sich die Layouts speziellen Medien zuordnen lassen. So ist es möglich, ein Layout zum Drucken und ein Layout für die Darstellung am Computer zu verwenden.

Zu XSL gehören

- das eigentliche XSL (zur Unterscheidung genannt XSL-FO) (XSL Formatting Objects) für die Beschreibung eines Dokuments als Baum mit Formatierungsanweisungen und Stilangaben,
- XSLT (XSL Transformations) für die Transformation eines beliebigen XML-Dokuments in einen anderen Baum
- und indirekt auch XPath für die Adressierung von Baumbestandteilen.

### **XSL-FO**

XSL-FO (Extensible Stylesheet Language - Formatting Objects) ist eine XML-Anwendung, die beschreibt, wie Text, Bilder, Linien und andere grafische Elemente auf einer Seite angeordnet werden. Mit Hilfe von XSL-FO ist es möglich, qualitativ hochwertige Druckerzeugnisse entweder auf Papier oder auf dem Bildschirm zu erzeugen.

### **XSLT**

XSLT ist die Abkürzung für XSL Transformation. XSLT ist eine Sprache zur Transformation von XML-Dokumenten. Es baut auf die logische Baumstruktur eines XML-Dokumentes auf und erlaubt die Definition von Umwandlungsregeln. XSLT-Programme, sogenannte XSLT-Stylesheets, sind dabei ebenfalls nach den Regeln des XML-Standards aufgebaut (Namespace für XSLT [www.w3.org/1999/XSL/Transform](http://www.w3.org/1999/XSL/Transform)).

Spezielle XSLT-Prozessoren lesen XSLT-Stylesheets ein und transformieren eine XML-Eingabedatei nach den Stylesheet-Regeln in das gewünschte Ausgabeformat.

Die deklarative, funktional-applikative Sprache XSLT ist u. a. aus DSSSL hervorgegangen. Entwickelt wurde XSLT ursprünglich von James Clark (XSLT 1.0, XSLT 1.1), für die derzeitige Weiterentwicklung ist Michael Kay verantwortlich. Derzeit hat XSLT 2.0 den Status einer "Candidate Recommendation" beim W3C.

## **XML Anwendungen**

### **XHTML**

Die Extensible Hypertext Markup Language (XHTML) löst als Standard des W3C HTML als Textauszeichnungssprache für Webseiten im World Wide Web ab. Im Gegensatz zu seinem Vorgänger HTML, welcher mittels SGML definiert wurde, verwendet XHTML die strengere und einfachere zu parsende SGML-Teilmenge XML als Sprachgrundlage. XHTML-Dokumente entsprechen also den Syntaxregeln von XML.

### **SVG**

Scalable Vector Graphics (SVG) ist eine Sprache zur Beschreibung zweidimensionaler Vektorgrafiken in XML.

### **MathML**

Die Mathematical Markup Language (MathML) ist im Computer-Datenverkehr ein Dokumentenformat zur Darstellung von mathematischen Formeln und komplexen Ausdrücken. Der Standard MathML 2.0 wird durch eine Spezifikation des World Wide Web Consortium von 2001 festgelegt. Wie in allen XML-Sprachen (z. B. XHTML) werden in MathML die Inhalte eines Dokumentes in einer logischen Struktur unabhängig von ihrer graphischen Gestaltung abgelegt.