# Erste Vorlesung

Semistrukturierte Daten: Datensätze schauen ungefähr gleich aus (muss nicht sein)

Struktur ändert sich vielleicht

Schema ist nicht nötig

Daten müssen in einem gerichteten Graphen dargestellt werden

Beschreibung der Kanten, dann die Daten (Datentypen können vorkommen)

XML eschreibt Struktur und Semantik des Dokumenteninhalts

Man kann eine neue Sprache entwerfen, damit kann man Date beschreiben

(Gib mir zum Beispiel die Kurse von SE aus, geht in XML)

Datenaustausch + Semantik WEB

Wohlgeformtheit: Gewisse syntaktische Regeln müssen eingehalten werden, damit ein anderer es verwenden kann

Werkzeuge überprüfen die Wohlgeformtheit

z.B.: Es darf nur eine ROOT geben, es muss einen Start-Tag und einen End-Tag geben…

Überprüft NICHT, ob die ausgetauschten Daten einer gewissen Struktur entsprechen

Validität: Schaut, ob Daten einer gewissen Struktur entsprechen

Braucht ein Schema!

Formale Dokumentenmodelle: DTD oder XML Schema

Parser ist eingebaut, nimmt ein Dokument (vielleicht auch ein Schema) und sagt, ob es Wohlgeformt und/oder Valide ist

XML unterstützt Datenaustausch

Multi-Delivery: Daten werden abgelegt, darüber baut man ein Stylesheet und aus diesen Daten macht man ein bestimmtes Layout

Intelligente Suche

Es gibt verflucht viele Dokumentenmodelle bei XML

Don’t even try listing them

Element- und Attribute beginnt mit einem Buchstaben oder Underscore, kein Leerzeichen, Sonderzeichen, kleinGroß erlaubt

Keine Attribute mit demselben Namen!

Reihenfolge der Attribute ist nicht bestimmt!

Prolog Flashback!

Standalone=Yes  
Dokument-Definition ist darin vorhanden

Intere und Externe DTD  
Bei großen DTDs ist es klüger, diese Extern einzubinden