

## XML in der Übersicht

XML steht für eXtensible Markup Language und es handelt sich um einen Metastandard für Dateiformate. Der Ausdruck extensible drückt es schon aus, XML ist erweiterbar. Darüberhinaus ist XML ein übergeordneter Standard, der es ermöglicht, übergeordnete Markup-Sprachen zu definieren.

Beispiele:

- RSS – Datenbereitstellung von Informationen in WWW (<http://de.wikipedia.org/wiki/RSS>)
- SVG – Repräsentation von Vectorgrafiken (<http://de.wikipedia.org/wiki/SVG>)
- XHTML – XML-Version von html (<http://de.wikipedia.org/wiki/XHTML>)
- XI – zentrale Schnittstelle zum SAP System  
(<http://www.sap.com/germany/solutions/netweaver/components/exchangeinfrastructure/>)

XML Dokumente bestehen aus Text, aus ASCII-Text. Dieser Text wird durch das Markup-Prinzip realisiert. Er enthält **Elemente** (auch Markups oder Tags) und **Attribute**. Jedes Markup besteht entweder aus einem öffnenden und schließenden Element (Start- und Endtag) oder kann in Kurzschreibweise erscheinen. Innerhalb dieser Tags kann es Text, weitere Tags oder Beides geben. Jeder Tag kann mit Attributen erweitert werden. Das Attribut besteht aus Namen und Wert.

```
<?xml version='1.0' encoding='UTF-8'?>
<!DOCTYPE message SYSTEM "Adresse oder .dtd">
<!-- Willkommen in der XML-Message -->
<message>
  <product>
    <name>Test</name>
    <inhalt>
      <zutat klasse='Elementgruppe'>Wasser</zutat>
      <menge>1cl</menge>
    </inhalt>
    <inhalt>
      <zutat>Feuer</zutat>
      <menge>2ml</menge>
    </inhalt>
    <inhalt>
      <zutat>Erde</zutat>
      <menge>3dl</menge>
    </inhalt>
    <inhalt>
      <zutat>Sauerstoff</zutat>
      <menge />
    </inhalt>
    <anleitung>Produktdefinition ...
  </anleitung>
  </ product >
</message>
<!-- Innerhalb der <message> koennen weitere <product> folgen -->
```

Für ein XML gibt es Regeln, diese sind im Dokument „[Die generelle Struktur eines XML.doc](#)“ zu finden.

## Vorteile:

XML ist ein **offener, lizenzfreier** Standard, der gut software- und entwicklungsseitig **gut unterstützt** wird. Die Popularität erklärt sich auch ein wenig darüber, dass populäre Unternehmen an diesem Standard mitarbeiten, z.B. hat Intel XML in einem Router integriert, IBM und Oracle haben ihre Datenbanken XML-fähig gemacht, Microsoft hat XML in seine Produkte integriert, ohne eigene Erweiterungen, und auch SAP definiert seine zentrale Schnittstelle „XI“ mit Hilfe von XML. Darüber hinaus ist es ein sehr flexibles Austauschformat. Es kann über die unterschiedlichsten Kanäle verteilt werden (z.B. eMail, FTP, HTTP, CD-ROM). In einem XML kann in bestimmten Grenzen eine **freie Modellierung** von Informationen stattfinden, ähnlich den Datenbanken. Damit können Daten **strukturiert** gespeichert werden. Sie sind dadurch durch andere Applikation lesbar. Das XML-Format ist offen, dadurch kann es **problemlos an andere Systeme** angebunden werden. Da es eine text- und dateibasierende Speicherung ist, ermöglicht XML die Verbreitung von Informationen über jedes Netz, **plattformunabhängig**. Ebenso ist es, mit etwas Mühe, **lesbar**. Die **Struktur des XML ist vom Layout getrennt**, insofern das für die Anwendung überhaupt benötigt wird.

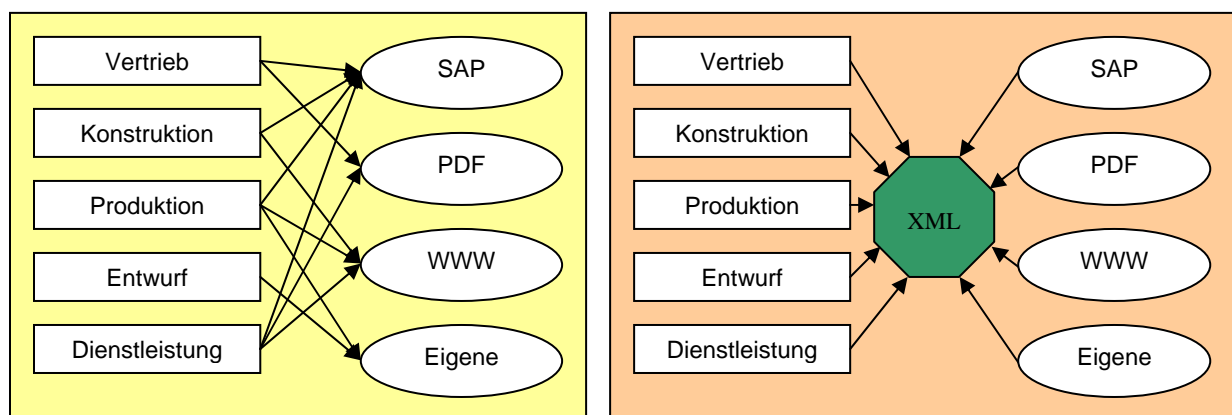
Ein weiterer Vorteil ist die Verwendung von Unicode, XML lässt sich internationalisieren, arbeitet also mit **beliebigen Zeichensätzen**. Für die Übersetzung in andere Sprachen wird kein spezielles Interface benötigt, es ist vergleichsweise einfach.

### Vorteil des XML gegenüber ASCII

Der kleinste gemeinsame Nenner eines Formates zum Austausch ist im besten Fall ASCII. Damit könnte man schon Daten recht gut austauschen. Leider bedeutet das aber auch Mehrarbeit, da es jeweils für das System aufgearbeitet werden muss. Die Übergabe von Sonderzeichen ist ebenso schlecht möglich. Mit XML würde ein eindeutiger Standard vorliegen und die Aufbereitung entfällt und der Unicode sorgt für die Verwendung aller Zeichen.

### Vorteile im Unternehmen

Ein System kann Daten, die im XML-Format vorliegen schnell erweitert werden. Beispielsweise können so Kunden und Lieferanten von Firmen schnell die Informationen erhalten, die sonst nur über eine Datenbankanwendung und deren Sicherheitskonzept erreichbar sind.



Das XML-Format ist offen, dadurch kann es problemlos an andere Systeme angebunden werden.

Für die langfristige Dateiablage ist XML ebenso geeignet und XML kann einfach in andere Dateiformate gewandelt werden.

## Die XML-Familie:

XML selbst wird durch weitere Standards beschrieben und enthält folgende Mitglieder.

- DTD – beschreibt die Struktur des XML-Dokumentes
- Schema – verbesserte Strukturbeschreibung auf Basis des XML-Syntax
- Namespaces – Elemente werden in einem Dokument unterschieden
- XSLT – ermöglicht die Umwandlung von XML-Dokumenten in andere XML-Strukturen
- XSL:FO – für die Dokumentenformatierung (Druckvorlagen)
- XQuery – Abfragesprache für XML-basierende Datenbanken
- Xpath – Adressierung definierter Teile von XML-Instanzen
- XLink, XPointer – Abfragesprache um auf Teile von XML-Dokumenten zuzugreifen

## Nachteile:

XML wird im Textformat gespeichert und benötigt dadurch einen höheren Speicherbedarf und die Verarbeitung ist damit auch langsamer. Speicherplatz wird aber immer günstiger. Die Vorteile des Textformates sollten diesen Nachteil überwiegen. Zusätzlich können Textformate sehr gut komprimiert werden, z.B. ZIP oder GZ.

## Weiterführende Informationen:

SELFHTML: Einführung in XML - <http://de.selfhtml.org/xml/intro.htm>

W3C – XML in 10 points - <http://www.w3c.de/Misc/XML-in-10-points.html>

XML im Unternehmen von Frank Bitzer, ISBN: 3898422887