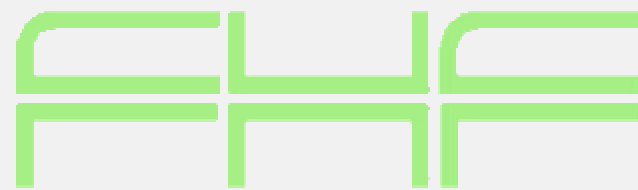


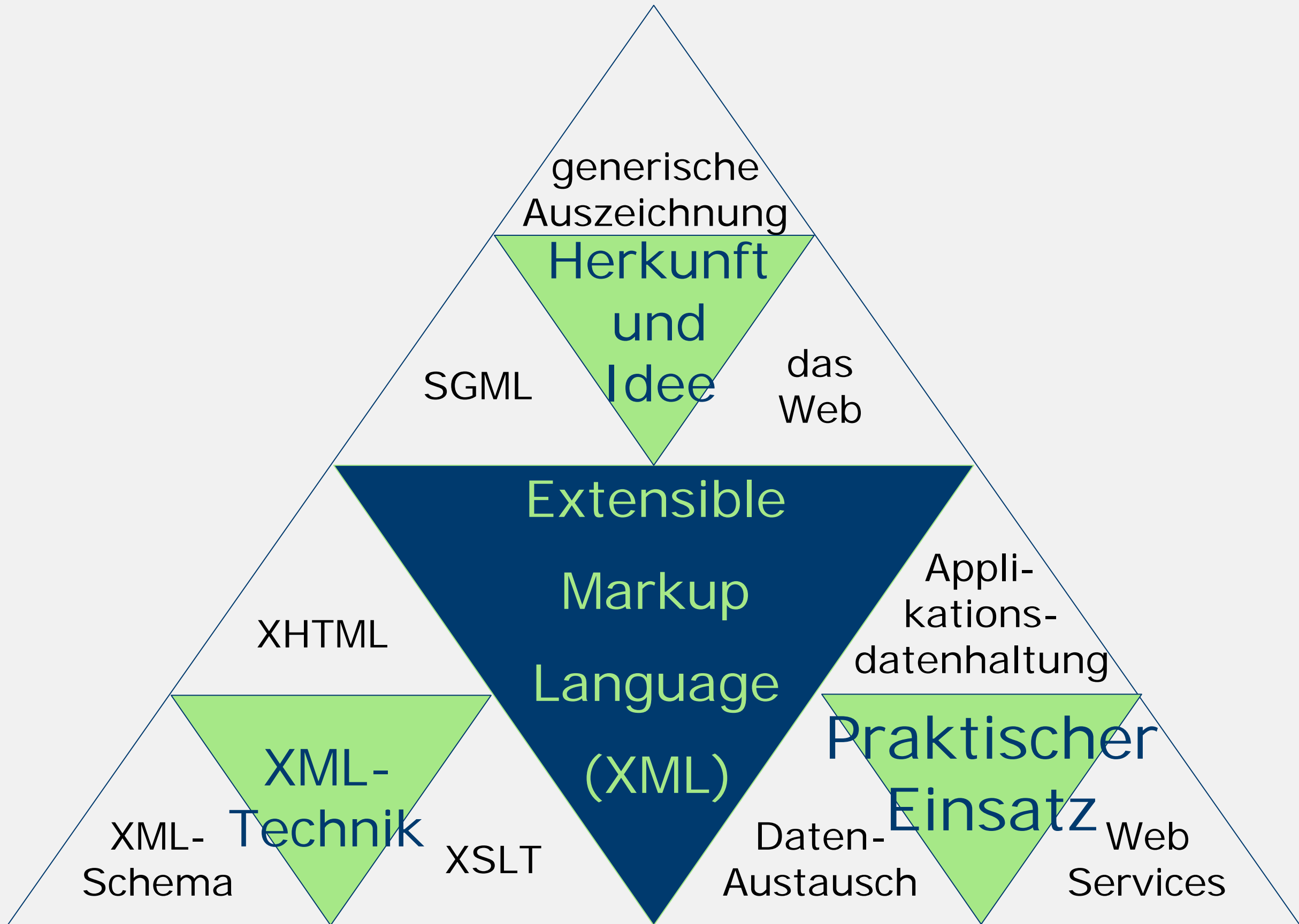
# Die Extensible Markup Language (XML)

Idee und praktische Anwendung

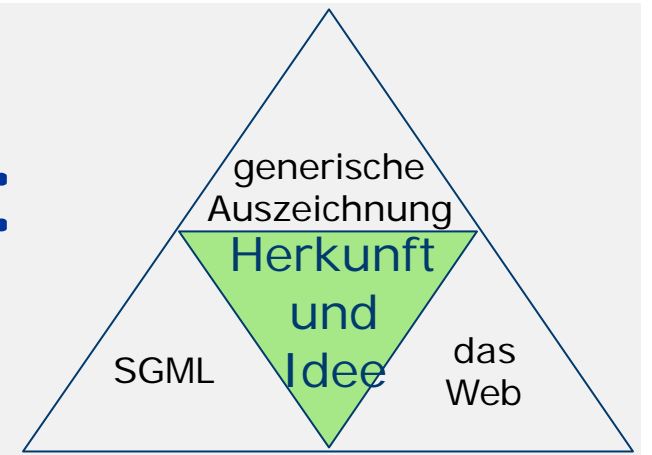


Prof. Mario Jeckle

Fachhochschule Furtwangen  
mario@jeckle.de  
<http://www.jeckle.de>



# Motivation und Einführung: Auszeichnungssprachen



## Präsentations Eigenschaften

## Editorielle Eigenschaften

Kleinschreibung

Unterstreichnung

Fettdruck

Einrückung

Anschrift

Anrede

Brieftext

Grußformel

**AHF**

Herrn  
Dr. Max Mustermann  
Musterstr. 42  
12345 Musterstadt

Unsere Zeichen ab/cd    Unsere Nachricht vom 2003-07-09    Ihre Zeichen xy/-    Ihre Nachricht vom 2003-07-01

**Sehr geehrte Damen und Herren,**

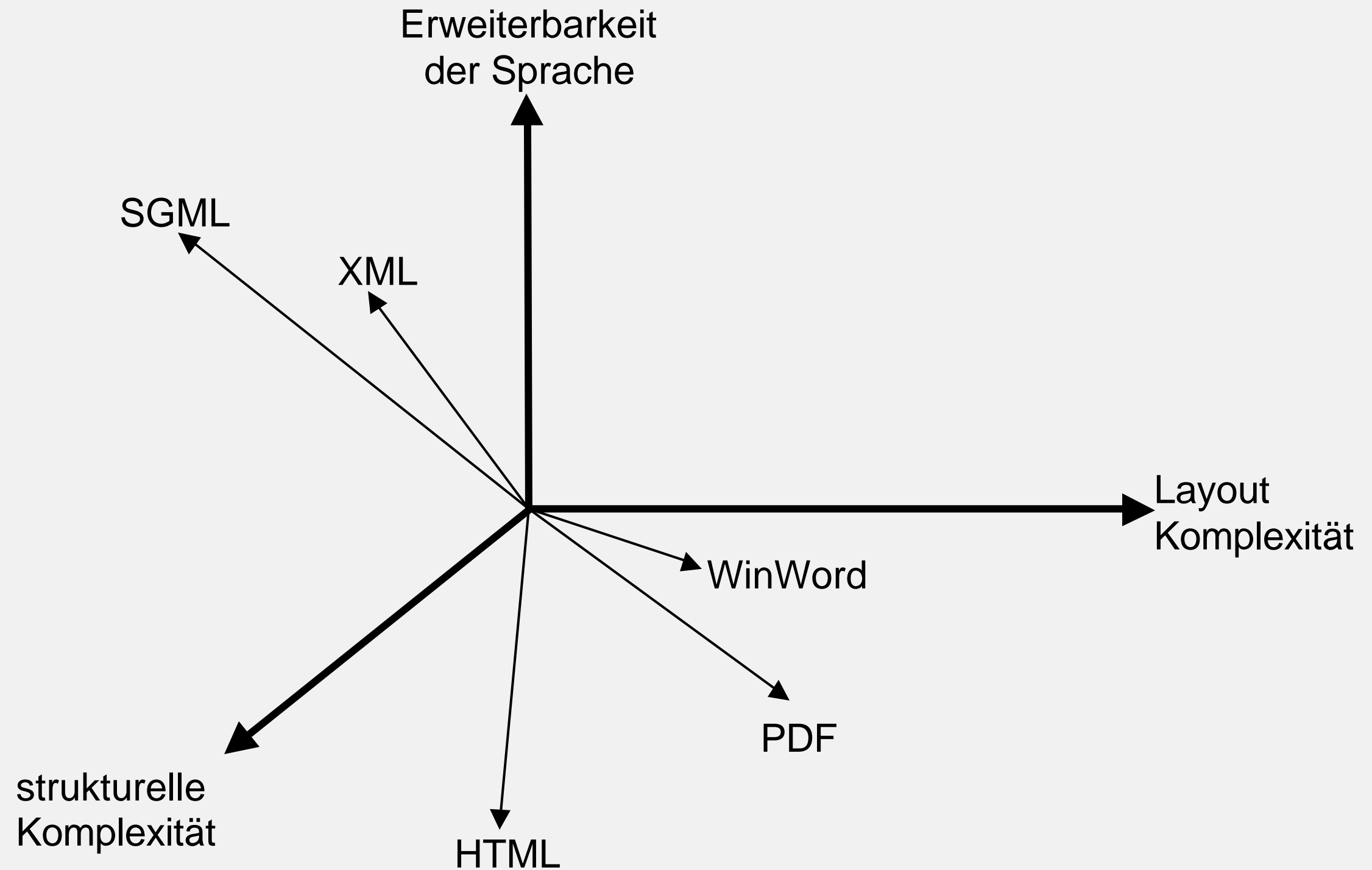
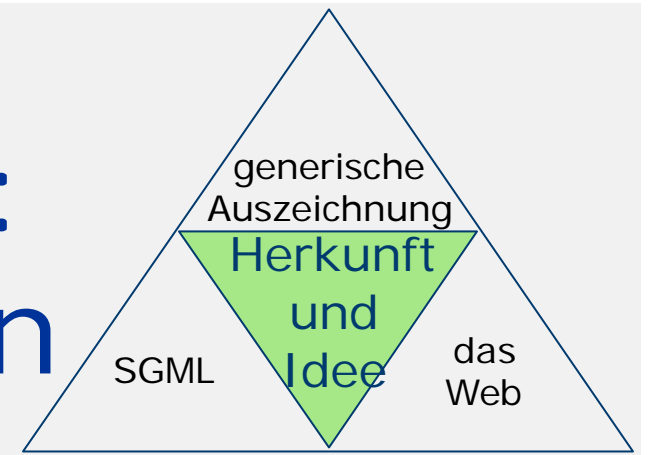
Maior pars mortalium, Pauline, de naturae malignitate conqueritur, quod in exiguum aevi gignimur, quod haec tam velociter, tam rapide dati nobis temporis spatia **decurrunt**, adeo ut exceptis admodum paucis ceteros in ipso vitae apparatu vita destituat.

Nec huic publico, ut opinantur, malo turba tantum et imprudens volgus ingemuit.

**Mit freundlichen Grüßen**

Mario Jeckle

# Motivation und Einführung: Dokument(atons-)sprachen



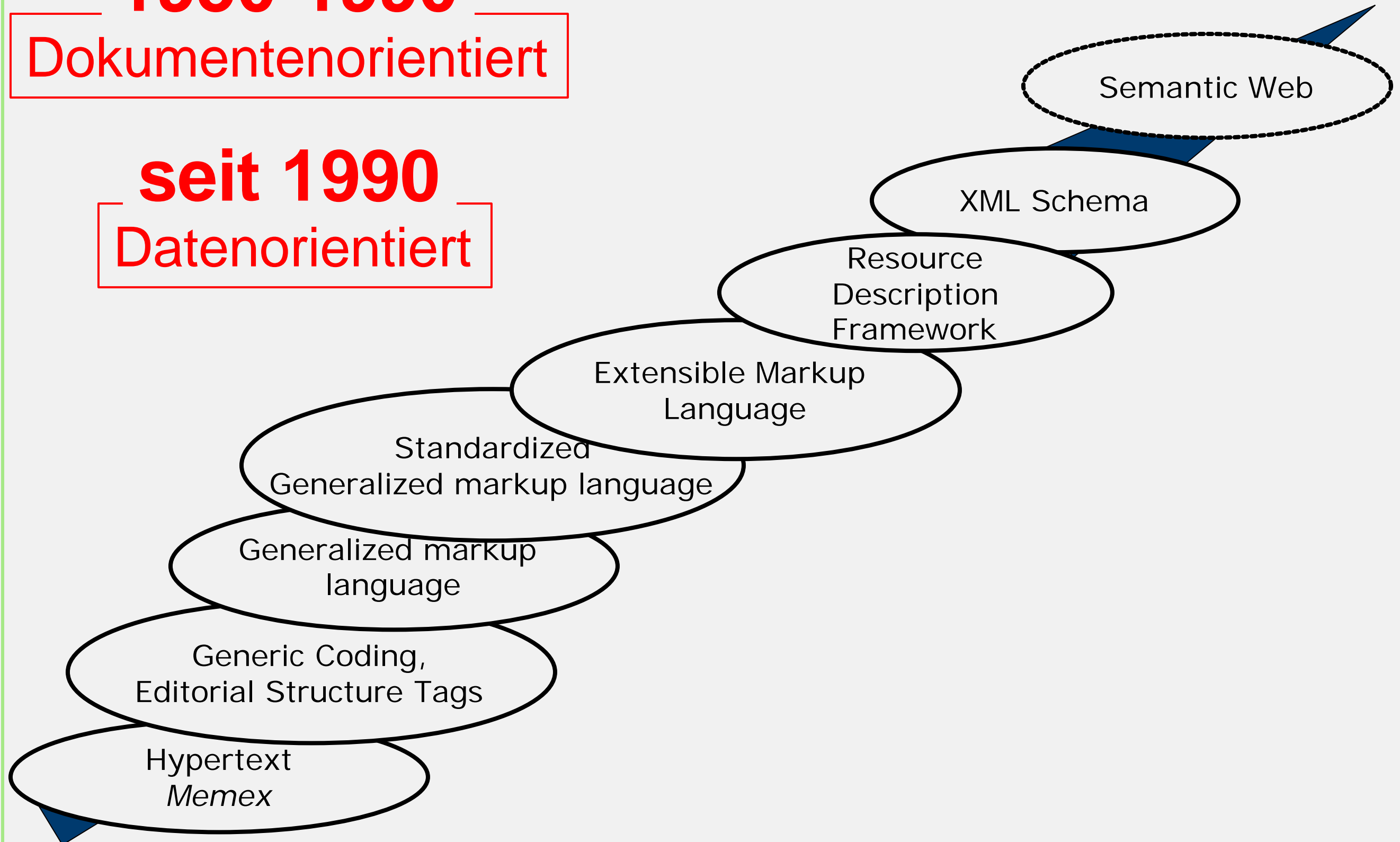
# Motivation und Einführung: Historische Entwicklung hin zu XML

**1950-1990**

**Dokumentenorientiert**

**seit 1990**

**Datenorientiert**

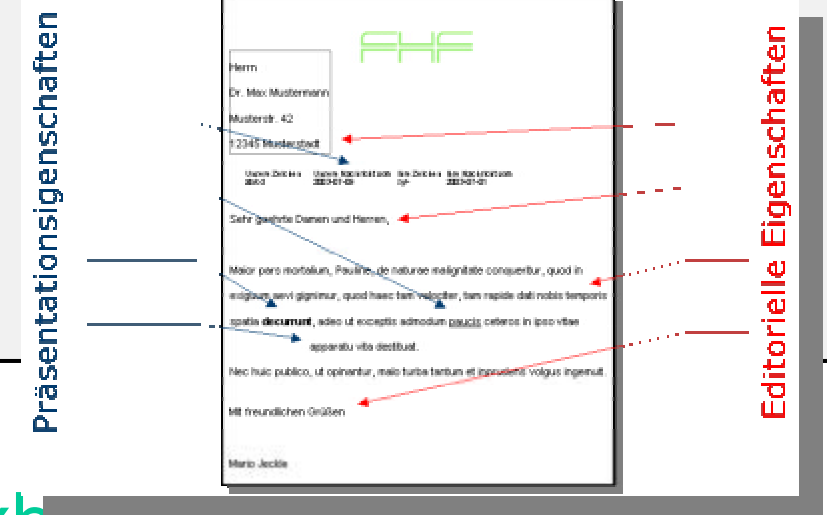


# Motivation und Einführung: Ein erstes XML-Dokument

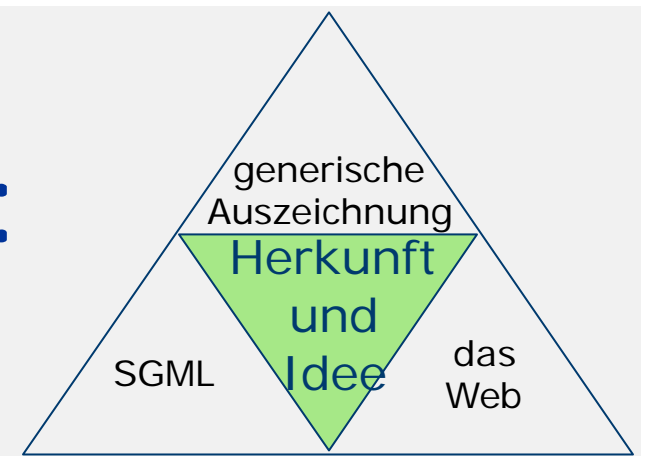
```

<?xml version="1.0" encoding="UTF-8"?>
<Vortrag xmlns:xhtml="http://www.w3.org/1999/xhtml" >
  <Titel>Die Extensible Markup Language (XML)</Titel>
  <Untertitel>Idee und praktische Anwendung</Untertitel>
  <Abstract>Seit ihrer Veröffentlichung durch das
  <xhtml:a href="http://www.w3.org">World Wide Web Konsortium
  </xhtml:a> ist der <xhtml:b>Extensible Markup Language</xhtml:b>...
  </Abstract>
  <Veranstaltung datum="2003-07-09">
    <Name>GI Regionalgruppe Südbaden</Name>
  </Veranstaltung>
  <Referent>
    <Name>Mario Jeckle</Name>
    <Institution>Fachhochschule Furtwangen</Institution>
    <URL>http://www.jeckle.de</URL>
    <E-Mail>mario@jeckle.de</E-Mail>
  </Referent>
</Vortrag>

```

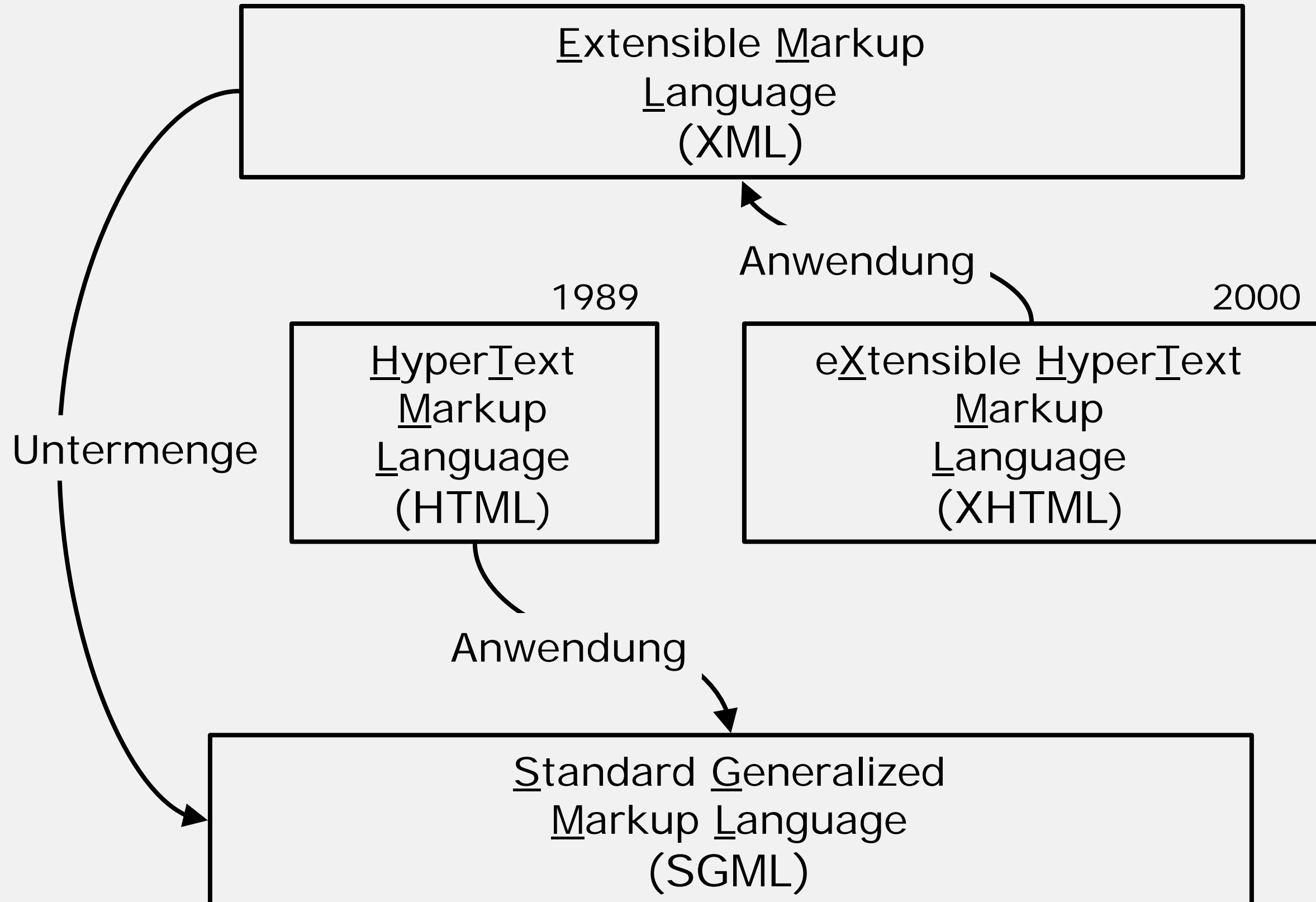
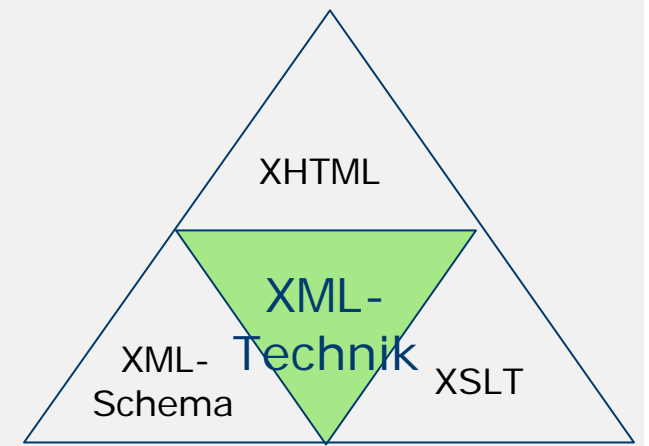


# Motivation und Einführung: XML in 10 Punkten

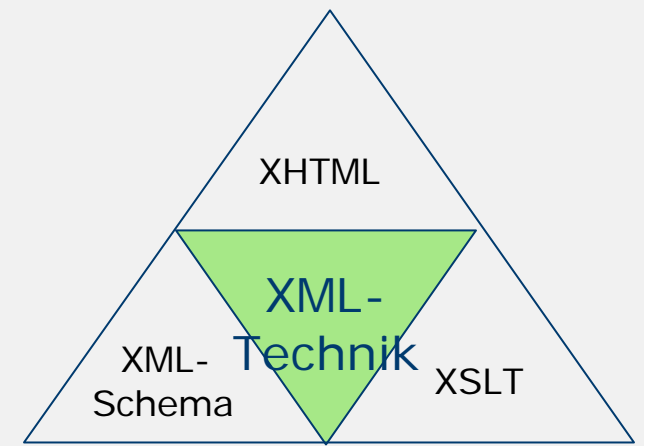


- I. XML steht für strukturierte Daten
- II. XML sieht ein wenig wie HTML aus
- III. XML ist Text, aber nicht zum Lesen
- IV. XML ist vom Design her ausführlich
- V. XML ist eine Familie von Techniken
- VI. XML ist neu, aber nicht so neu
- VII. XML überführt HTML in XHTML
- VIII. XML ist modular
- IX. XML ist die Basis für RDF und das Semantic Web
- X. XML ist lizenzfrei, plattformunabhängig und gut unterstützt

# XML-Technik: (X)HTML



# XML-Technik: (X)HTML



- Keine überlappenden Elemente

```
<p>here is an emphasized <em>paragraph</em>.</p>
```

```
<p>here is an emphasized <em>paragraph</p>.</em>
```

- Korrekte Terminierung

```
<p>here is a paragraph.</p>
<p>here is another paragraph.</p>
```

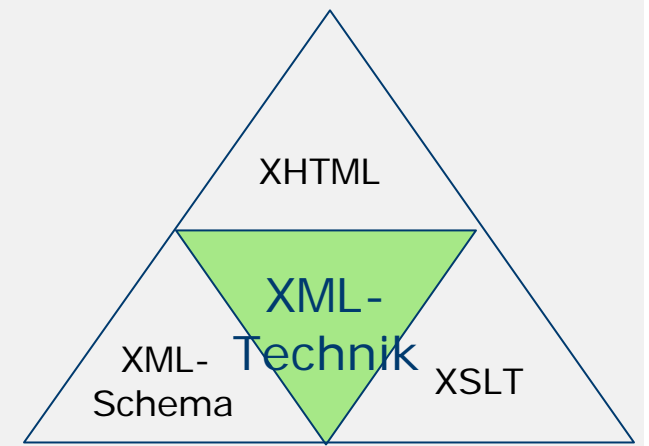
```
<p>here is a paragraph.
<p>here is another paragraph.
```

- Attributdarstellung

```
table rows="3"
```

```
table rows=3
```

# XML-Technik: (X)HTML



- Attributminimierung

```
<dl compact="compact">
```

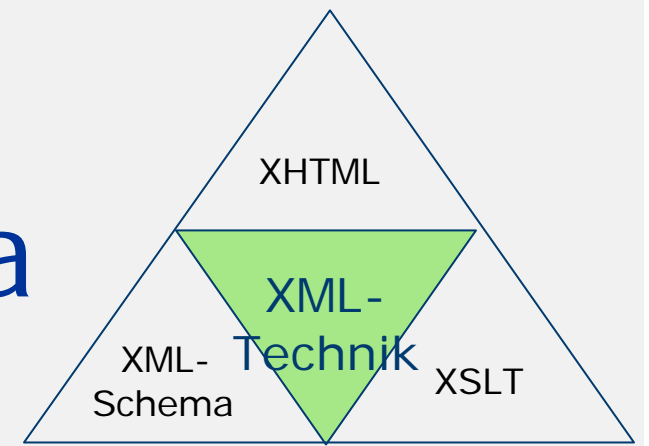
```
<dl compact>
```

- Leere Elemente

```
<br/><hr/>
```

```
<br><hr>
```

# XML-Technik: XML-Schema



```

<?xml version="1.0" encoding="UTF-8"?>
<Vortrag xmlns:xhtml="http://www.w3.org/1999/xhtml">
  <Titel>Die Extensible Markup Language (XML)</Titel>
  <Untertitel>Idee und praktische Anwendung</Untertitel>
  <Abstract>Seit ihrer Veröffentlichung durch das
  <xhtml:a href="http://www.w3.org">World Wide Web Konsortium
  </xhtml:a> ist der <xhtml:b>Extensible Markup Language</xhtml:b>...
  </Abstract>
  <Veranstaltung datum="2003-07-09">
    <Name>GI Regionalgruppe Südbaden</Name>
  </Veranstaltung>
  <Referent>
    <Name>Mario Jeckle</Name>
    <Institution>Fachhochschule Furtwangen</Institution>
    <URL>http://www.jeckle.de</URL>
    <E-Mail>mario@jeckle.de</E-Mail>
  </Referent>
</Vortrag>

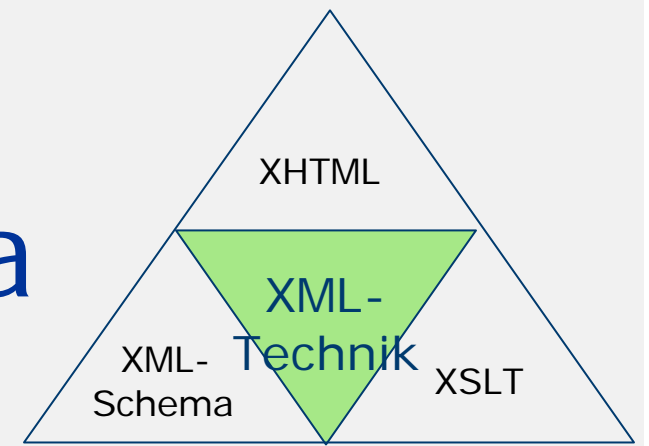
```

**Text**

**Datum**

**URI**

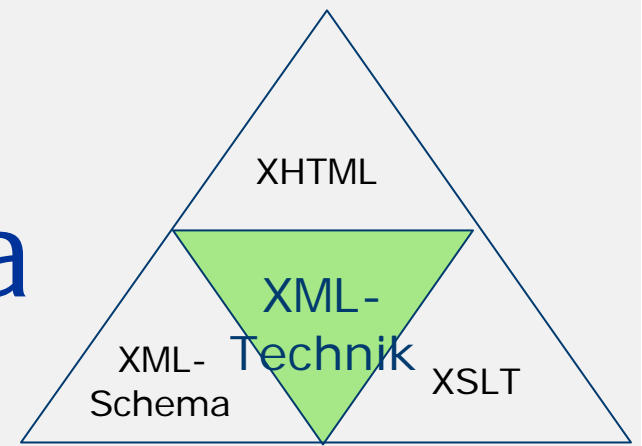
# XML-Technik: XML-Schema



XML Schema Part 2 definiert ein an die gängigen Programmier- und Datenbanksprachen angelehntes Typsystem zur Validierung von XML-Dokumenten und zur plattformunabhängigen lexikalischen Repräsentation typisierter Inhalte.

Typ	XML-Schema	Microsoft CLS	Java	SQL92/SQL3
$-2^7 \leq t \leq 2^7$	byte	-	byte	tinyint
$0 \leq t \leq 2^8 - 1$	unsignedByte	byte	-	
$-2^{15} \leq t \leq 2^{15} - 1$	short	short	short	smallint
$-2^{31} \leq t \leq 2^{31} - 1$	int	int	int	int, integer
$-2^{63} \leq t \leq 2^{63} - 1$	long	long	long	byte
Unicode-Zeichen(kette)	string	string	String	
IEEE754	float double	float double	float double	float, real, double
{true,false,0,1}	boolean	bool	boolean	boolean
$\mathbb{N}$	decimal	decimal	BigDecimal	numeric(p,s), dec, decimal(p,s)

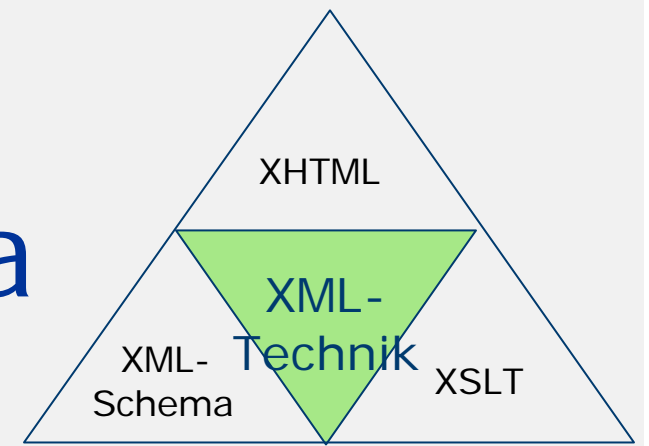
# XML-Technik: XML-Schema



```

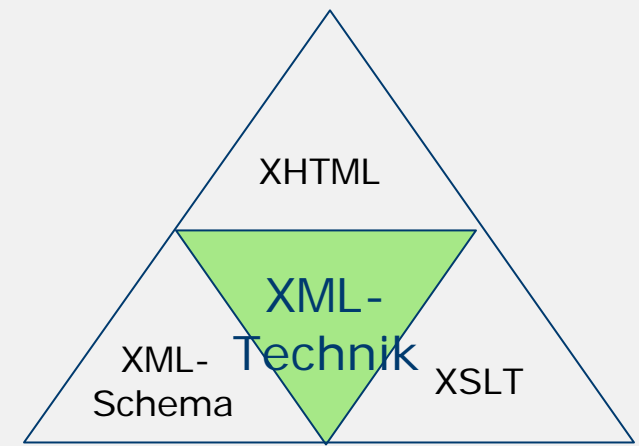
<?xml version="1.0" encoding="UTF-8"?>
<Vortrag xmlns:xhtml="http://www.w3.org/1999/xhtml">
  <Titel>Die Extensible Markup Language (XML)</Titel>
  <Untertitel>Idee und praktische Anwendung</Untertitel>
  <Abstract>Seit ihrer Veröffentlichung durch das
  <xhtml:a href="http://www.w3.org">World Wide Web Konsortium
  </xhtml:a> ist der <xhtml:b>Extensible Marku
  </xhtml:b>...
  </Abstract>
  <Veranstaltung datum="2003-07-09">
    <Titel>
    <Untertitel>
    <Abstract>
    <Veranstaltung
      <Name>Mario Jeckle</Name>
      <Institution>Fachhochschule Furtwang
      <URL>http://www.jeckle.de</URL>
      <E-Mail>mario@jeckle.de</E-Mail>
    </Referent>
  </Vortrag>
  
```

# XML-Technik: XML-Schema



- Grammatik für beliebige XML-Vokabulare
- Part 1 beschreibt Strukturen und Inhaltseinschränkungen
- Part 2 definiert Datentypdefinition für Schema Part 1 und weitere XML-Vokabulare
- Signifikante Erweiterung der DTD-Mächtigkeit, wird diese langfristig ersetzen
- Ist eine XML-Sprache
- Integriert die wichtigsten konkurrierenden Vorgängeransätze
- Seit 2001-05-02 W3C Recommendation
- Basis aller W3C-Standards der zweiten Generation (XPath v2.0, XSLT v2.0, XHTML v2.0, SOAP v1.2, ...)
- Werkzeugunterstützung verfügbar
- Erster Schritt der Schema-Bestrebungen, weitere werden folgen ...

# XML-Technik: XSLT

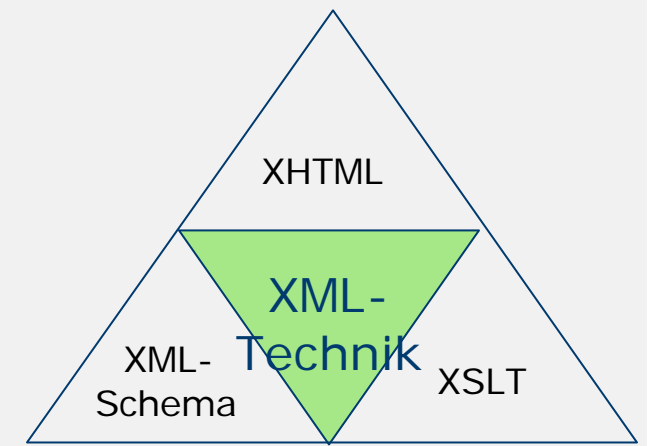


- Problem: Menge existierender XML-Vokabulare
- Ursache: XML Vokabularerstellung ist einfach ...
- Lösung:
  - Standardisierung
  - Transformationen

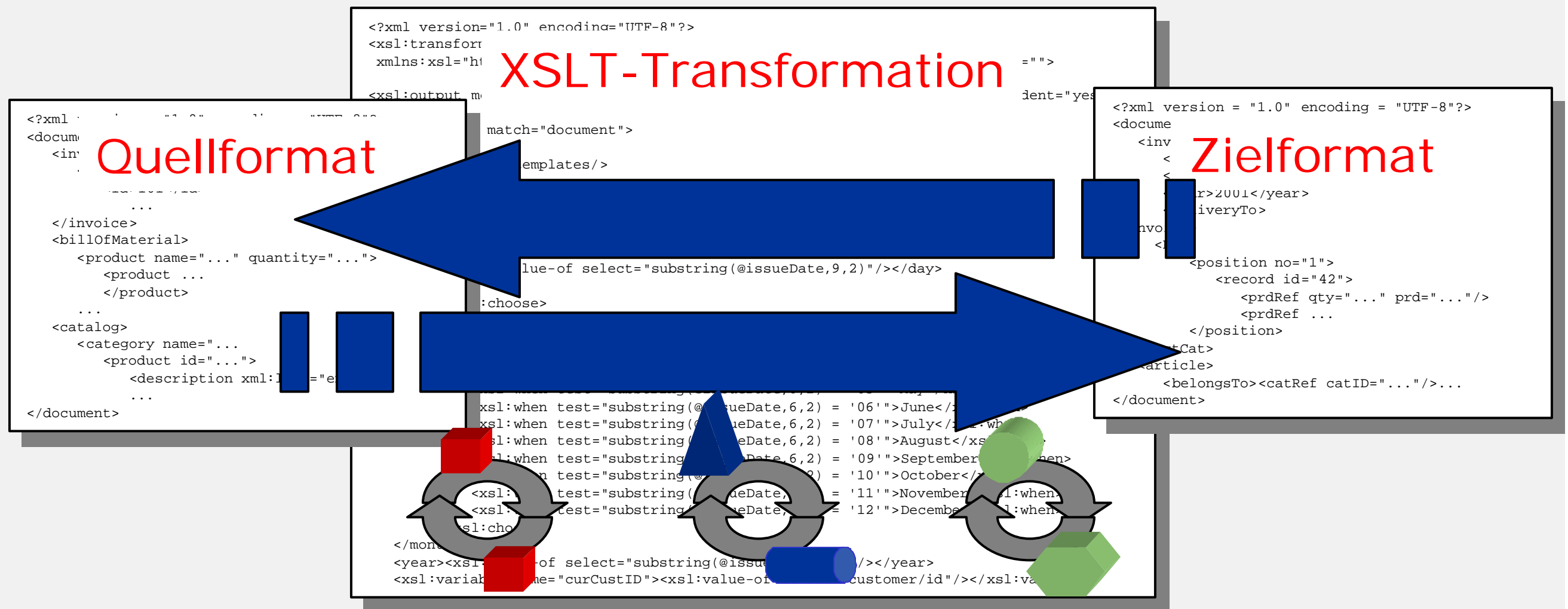
```
<?xml version = "1.0" encoding = "UTF-8"?>
<document>
  <invoice issueDate="2001-08-23">
    <customer>
      <id>101</id>
      ...
    </invoice>
    <billOfMaterial>
      <product name="..." quantity="...">
        <product ...
        </product>
        ...
      </billOfMaterial>
    <catalog>
      <catalog name="...">
        <catalog id="...">
          <description xml:lang="en">
            ...
          </description>
        </catalog>
      </catalog>
    </document>
```

```
<?xml version = "1.0" encoding = "UTF-8"?>
<document>
  <invoice>
    <day>23</day>
    <month>August</month>
    <year>2001</year>
    <deliveryTo>
      ...
    </invoice>
    <BOM>
      <position no="1">
        <record id="42">
          <prdRef qty="..." prd="..." />
          <prdRef ...
        </record>
      </position>
    </BOM>
    <productCat>
      <article>
        <belongsTo><catRef ca
      </article>
    </productCat>
  </document>
```

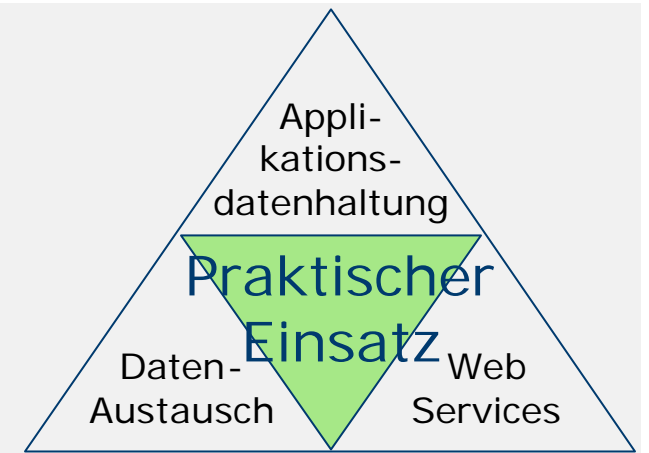
# XML-Technik: XSLT



- Wandlung zwischen verschiedenen XML-Vokabularen mit der W3C-standardisierten Sprache XSL Transformations

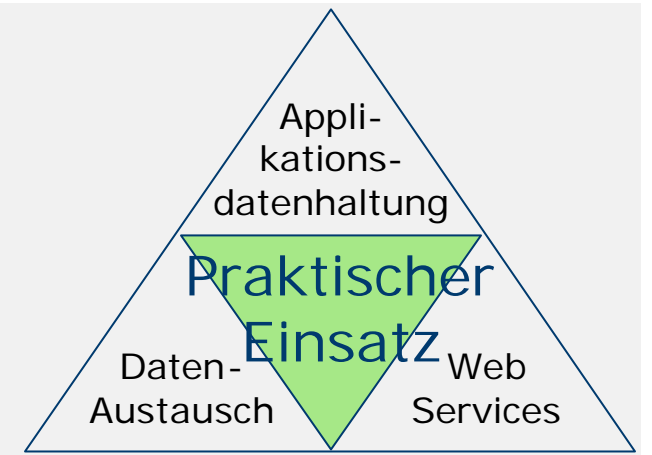


# Praktischer Einsatz



- XML bietet sich an zur Speicherung von ...
  - Konfigurationsdaten
  - Export-/Importdaten
  - Übermittlung von Dokumenten und entfernten Funktionsaufrufen
- XML-Technik bietet:
  - generische programmiersprachenunabhängige Verarbeitungsschnittstellen (Parser und Speichermodell)
  - Sprach- und applikationsunabhängiges Typsystem (XML Schema Part 2)
  - Encodierungs- und Aufrufkonventionen (Web Services)

# Praktischer Einsatz: Applikationsdatenhaltung



Java Data Binding (JAXB, Castor)

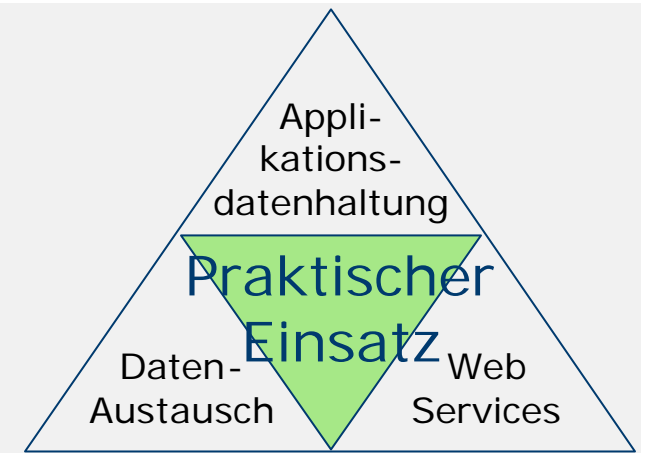
Problem-  
spezifität  
und  
Lösungs-  
optimier-  
ung

Document Object Model (DOM)

Abstraktion  
und  
Generizität

Simple API for XML (SAX)

# Praktischer Einsatz: Applikationsdatenhaltung



## Java Data Binding (JAXB, Castor)

- (aus DTD/XSD) generierte Speicherdarstellung
- Java-Speichersicht
- Generierte Zugriffsmethoden

## Document Object Model (DOM)

- Abstrahierte Speicherdarstellung
- Baumartige Sichtweise
- Objektorientierte Schnittstelle

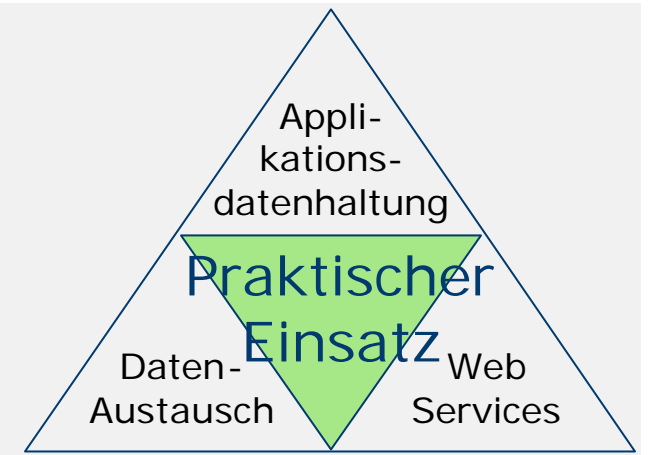
## Simple API for XML (SAX)

- Lineare Verarbeitung
- Ereignisbasiert
- Programmierparadigmenneutral

Abstraktion  
und  
Generizität

Problem-  
spezifität  
und  
Lösungs-  
optimier-  
ung

# Praktischer Einsatz: Applikationsdatenhaltung



- Simple API for XML (SAX)
  - Entwickelt durch Mitglieder der Mailingliste xml-dev auf Initiative von David Megginson
  - Zusammenstellung sprachunabhängiger Schnittstellen
  - Leichtgewichtig:
    - einfach (in eigenen Programmen) einzusetzen
    - einfach (in beliebige Programmiersprachen) umzusetzen
  - Serieller linearer ereignisbasierter Parsingvorgang
  - Definiert und erfordert keine Speicherstrukturen
  - SAX ist eine Schnittstelle für Parser, kein Parser!
  - Laufzeitverhalten skaliert linear mit Dokumentgröße

## Java Data Binding (JAXB, Castor)

- (aus DTD/XSD) generierte Speicherdarstellung
- Java-Speichersicht
- Generierte Zugriffsmethoden

## Document Object Model (DOM)

- Abstrahierte Speicherdarstellung
- Baumartige Sichtweise
- Objektorientierte Schnittstelle

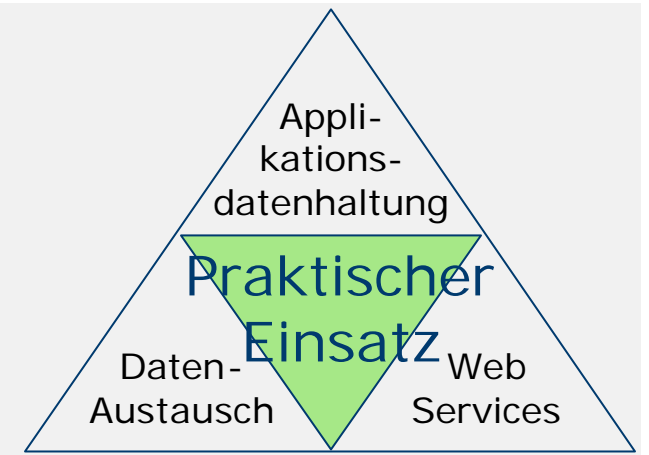
## Simple API for XML (SAX)

- Lineare Verarbeitung
- Ereignisbasiert
- Programmierparadigmenneutral

Abstraktion  
und  
Generizität

Problem-  
spezifität  
und  
Lösungs-  
optimier-  
ung

# Praktischer Einsatz: Applikationsdatenhaltung



- Document Object Model (DOM)
  - W3C standardisierte API für HTML, XML, CSS, etc.
  - Definiert programmiersprachenunabhängige baumartige objektorientierte Speicherstruktur für XML mit Operationen zur Navigation und Manipulation
  - Speicheraufwand zur Verwaltung des gesamten XML-Dokuments im Hauptspeicher teilweise sehr hoch
  - Mit *JDOM* existiert eine für Java optimierte moderne Implementierung
  - Verfügbar: Diverse Implementierungen: Java 1.4, JDOM, DOM4J, ...

## Java Data Binding (JAXB, Castor)

- (aus DTD/XSD) generierte Speicherdarstellung
- Java-Speichersicht
- Generierte Zugriffsmethoden

## Document Object Model (DOM)

- Abstrahierte Speicherdarstellung
- Baumartige Sichtweise
- Objektorientierte Schnittstelle

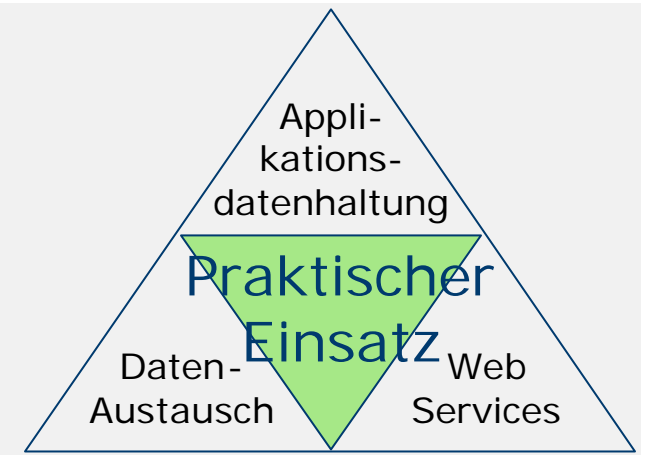
## Simple API for XML (SAX)

- Lineare Verarbeitung
- Ereignisbasiert
- Programmierparadigmenneutral

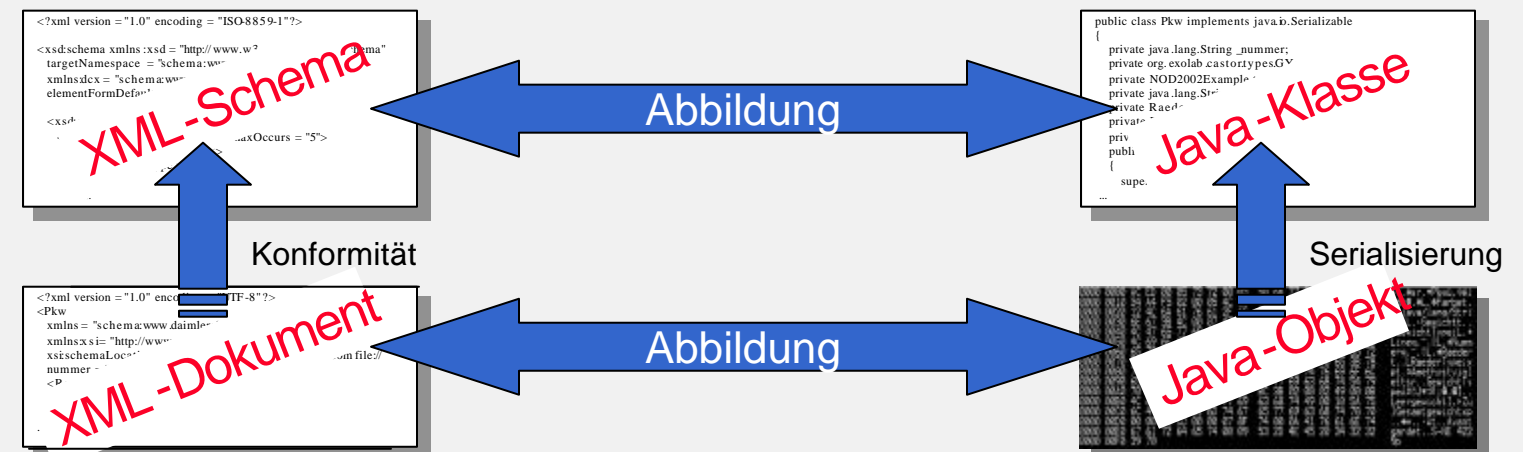
Abstraktion  
und  
Generizität

Problem-  
spezifität  
und  
Lösungs-  
optimierung

# Praktischer Einsatz: Applikationsdatenhaltung



- Java Data Binding
  - Grundidee: Verringerung des *Impedance Mismatch* zwischen Java und XML
  - Technisch: Generierung von Java-Klassendefinitionen aus Grammatik (DTD/XSD) und optional zusätzlich einer Abbildungsdefinition



**Java Data Binding (JAXB, Castor)**

- (aus DTD/XSD) generierte Speicherdarstellung
- Java-Speichersicht
- Generierte Zugriffsmethoden

**Document Object Model (DOM)**

- Abstrahierte Speicherdarstellung
- Baumartige Sichtweise
- Objektorientierte Schnittstelle

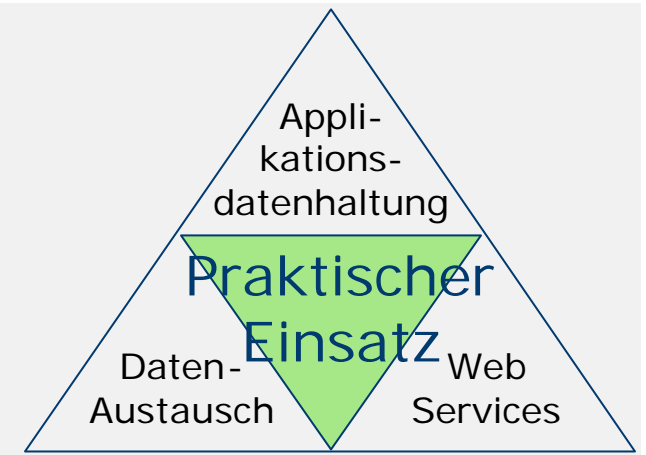
**Simple API for XML (SAX)**

- Lineare Verarbeitung
- Ereignisbasiert
- Programmierparadigmenneutral

Problem-  
spezifität  
und  
Lösungs-  
optimierung

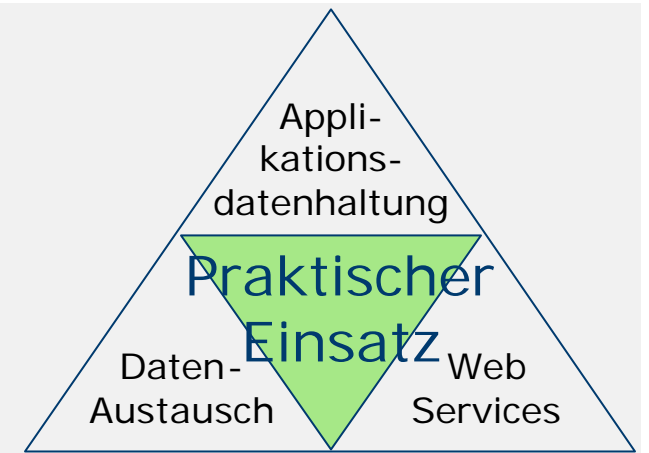
Abstraktion  
und  
Generizität

# Praktischer Einsatz: Datenaustausch



- Ablage von applikationsinternen Daten im XML-Format ...
  - Datenversand an Clients (z.B. WML, XHTML, RSS)
  - Langzeitspeicherung/Archivierung
  - Kopplungsmodell-neutraler Datentransfer (A2A)
  - Erleichterung der Systemintegration (B2Bi, EAI)
  - Unterstützung flexibler ad hoc Kollaboration (B2B)

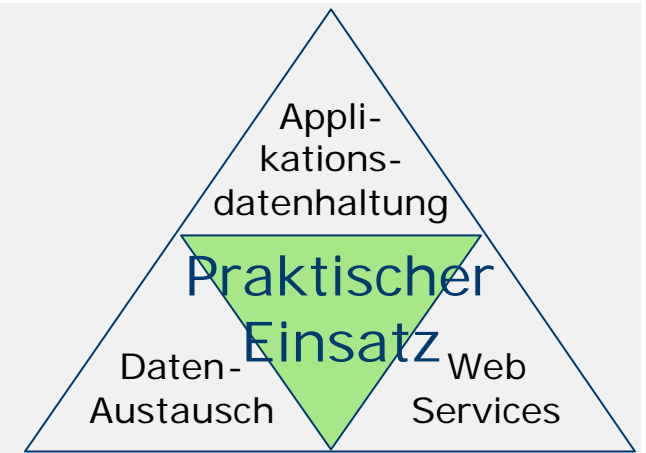
# Praktischer Einsatz: Web Services



## Was ist ein Web Service?

- Intel: Web Services sind eine Technik zur Kommunikation zwischen heterogenen Geräten.
- Microsoft: Über das Internet auslieferbare Softwarelösung auf der Basis von XML.
- SUN: Applikationsfunktionalität die über das Web zugreifbar wird.
- ... sowie ca. 50 weitere Definitionen

# Praktischer Einsatz: Web Services



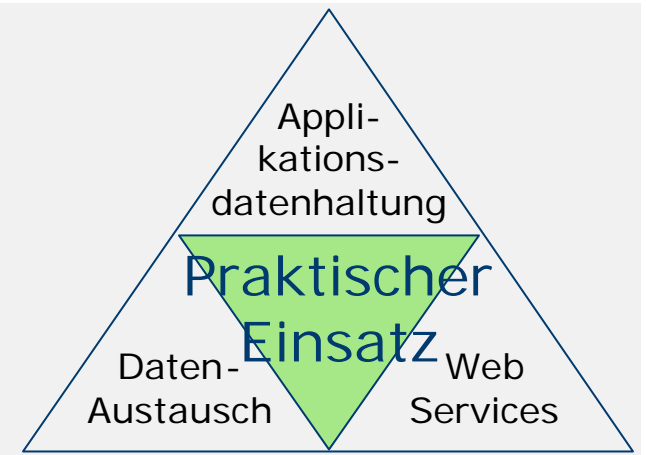
## Was ist ein Web Service?

- Definitionsdimensionen
  - Strikt technisch
    - Gefahr des vorzeitigen Veraltens
  - Strikt funktional
    - Gefahr der mangelnden (technischen)

Trennschärfe

- Einbezug des kulturellen Kontexts:
  - US-Amerikanische Definitionen: eher funktional
  - Europäische Definitionen: eher technisch

# Praktischer Einsatz: Web Services



Prozeßbeschreibung

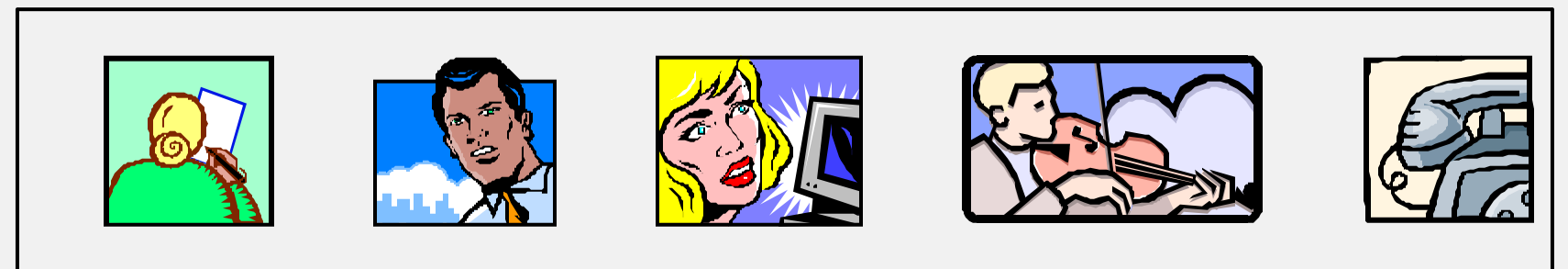
Verzeichnisdienst

Beschreibung

Datentransport

Inhaltsdarstellung

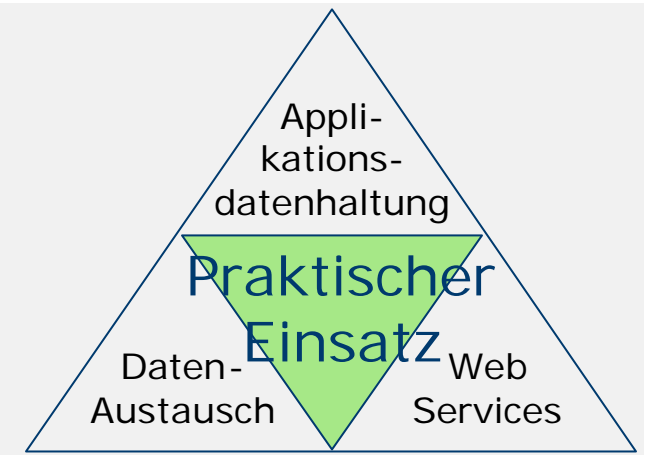
Infrastruktur



Web

Dienstleistung

# Praktischer Einsatz: Web Services



Prozeßbeschreibung

Verzeichnisdienst

Beschreibung

Datentransport

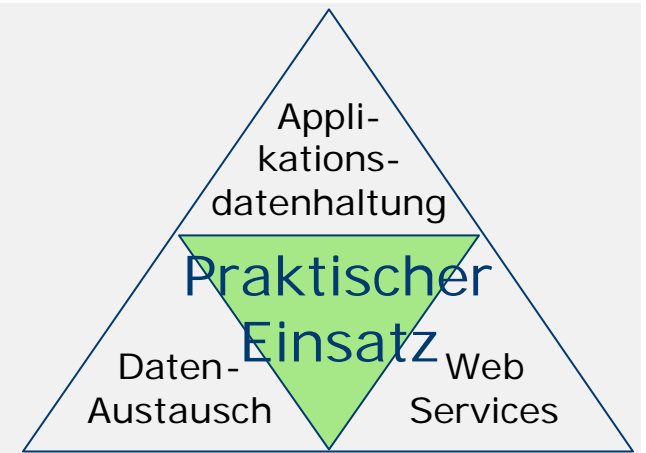
Inhaltsdarstellung

Infrastruktur

basaler  
Web Service



# Praktischer Einsatz: Web Services



Prozeßbeschreibung

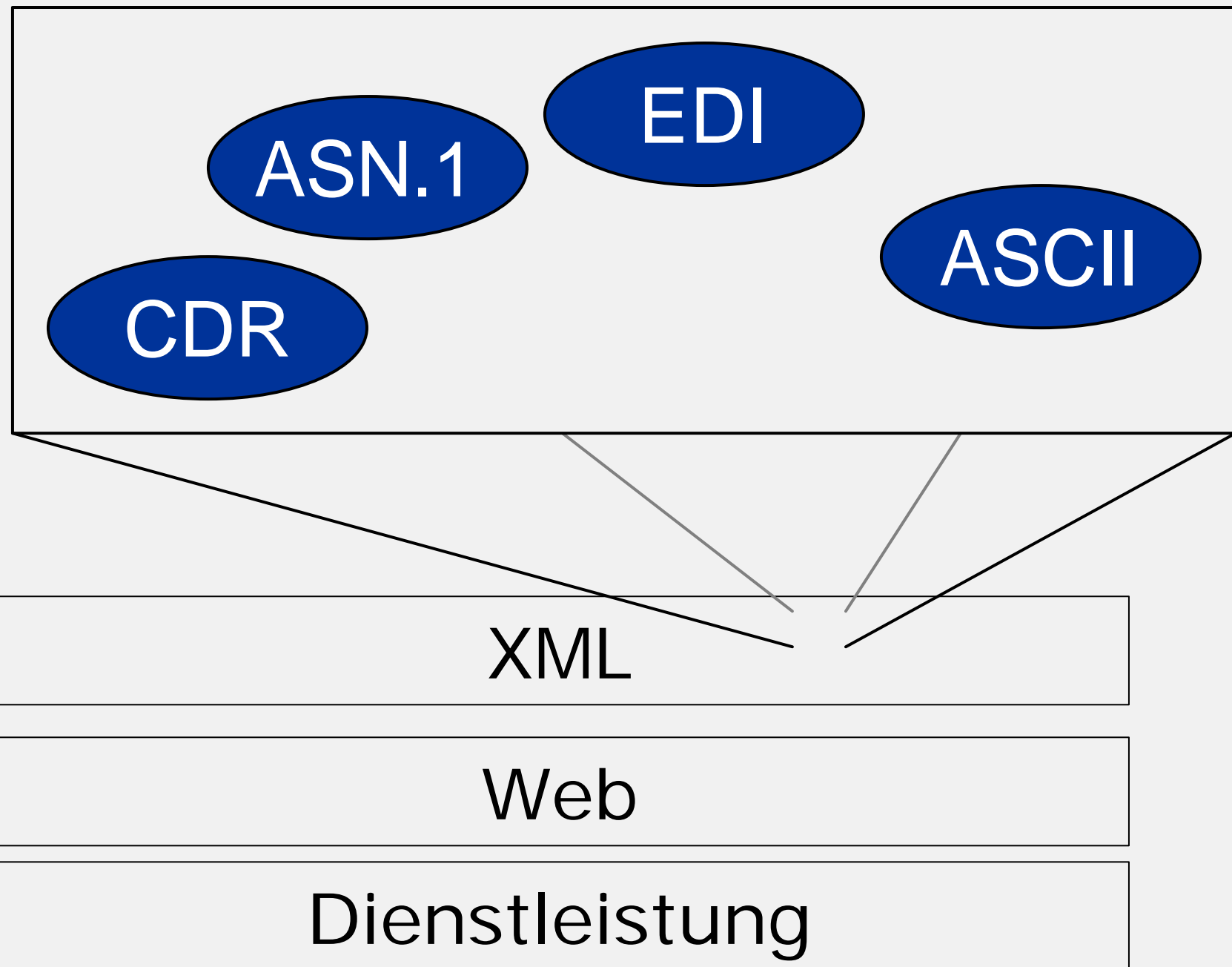
Verzeichnisdienst

Beschreibung

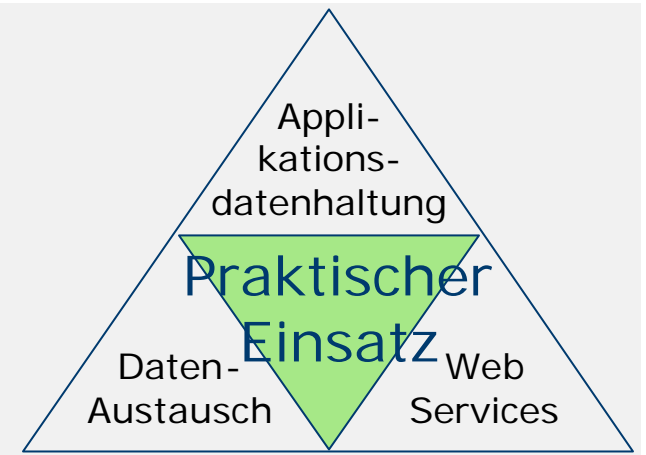
Datentransport

Inhaltsdarstellung

Infrastruktur



# Praktischer Einsatz: Web Services



Prozeßbeschreibung

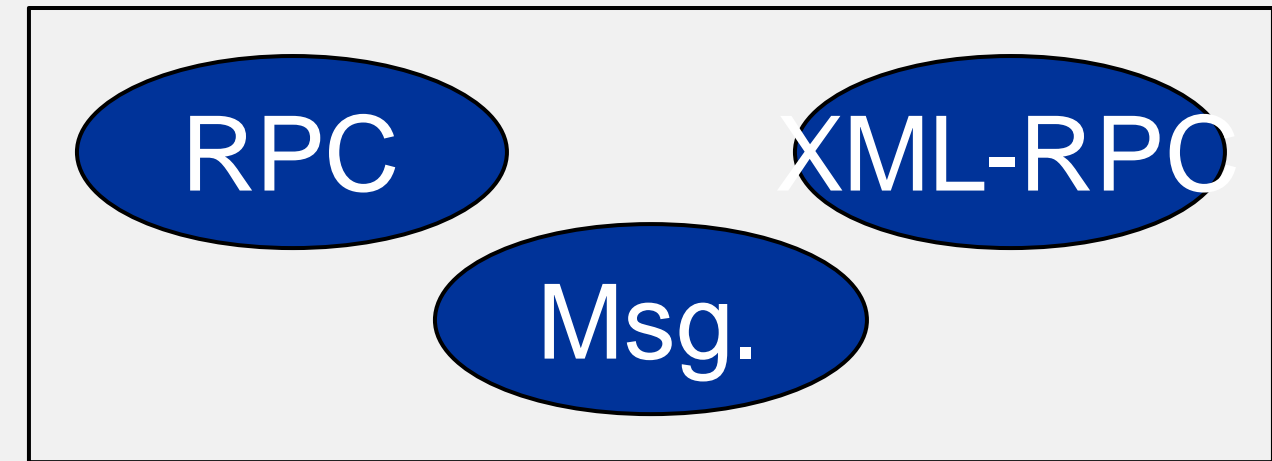
Verzeichnisdienst

Beschreibung

Datentransport

Inhaltsdarstellung

Infrastruktur



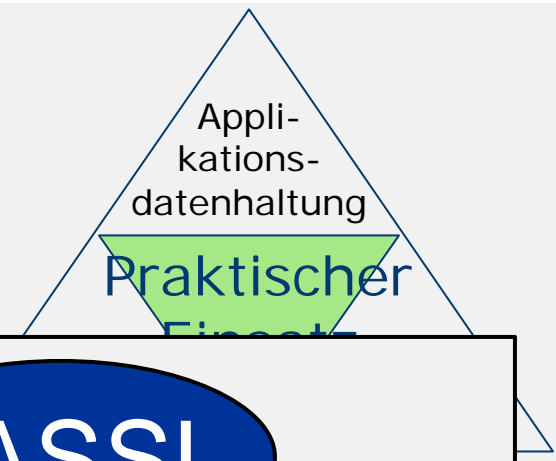
SOAP

XML

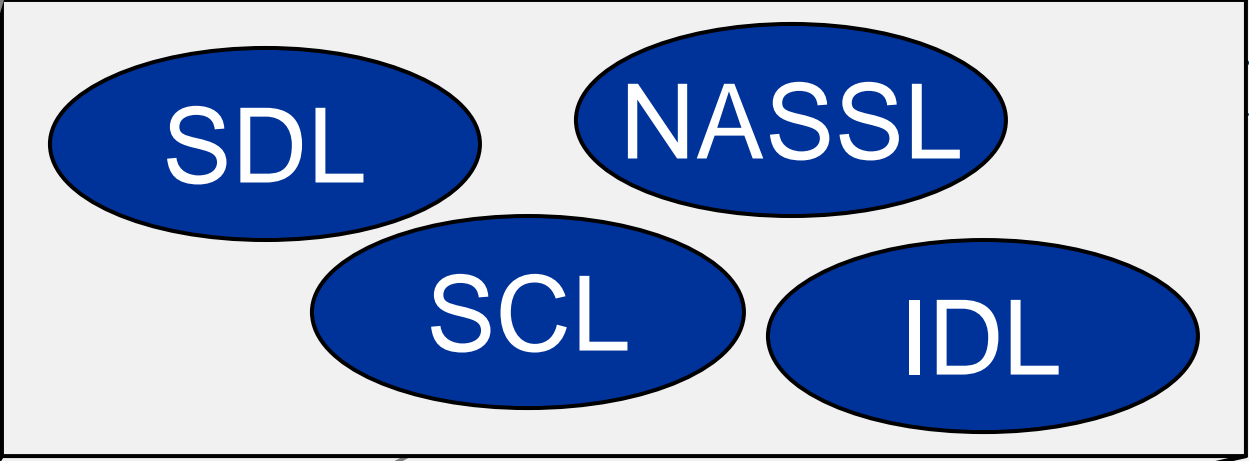
Web

Dienstleistung

# Praktischer Einsatz: Web Services



Prozeßbeschreibung



Verzeichnisdienst

Beschreibung



Datentransport



Inhaltsdarstellung



Infrastruktur



# Praktischer Einsatz: Web Services



Prozeßbeschreibung

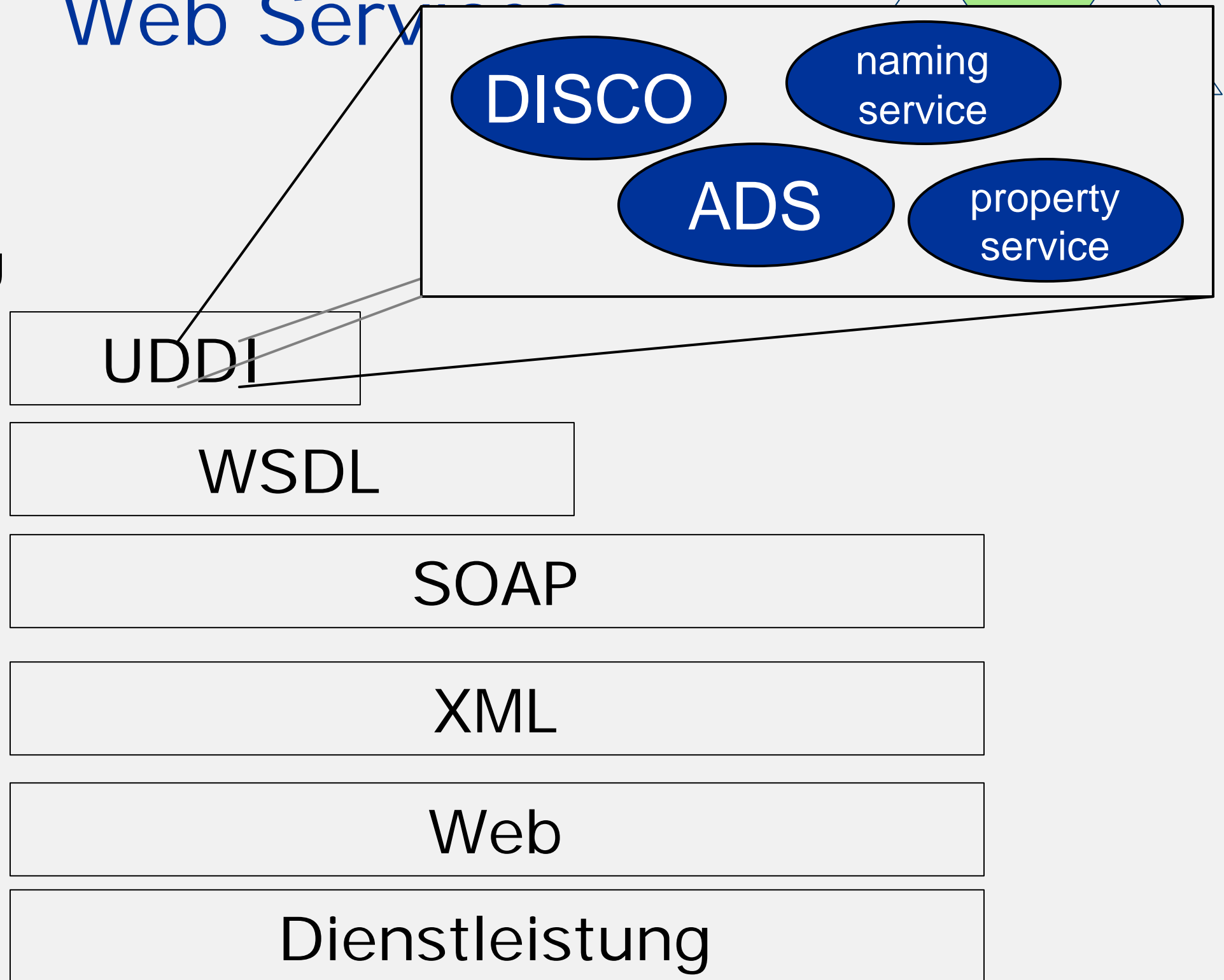
Verzeichnisdienst

Beschreibung

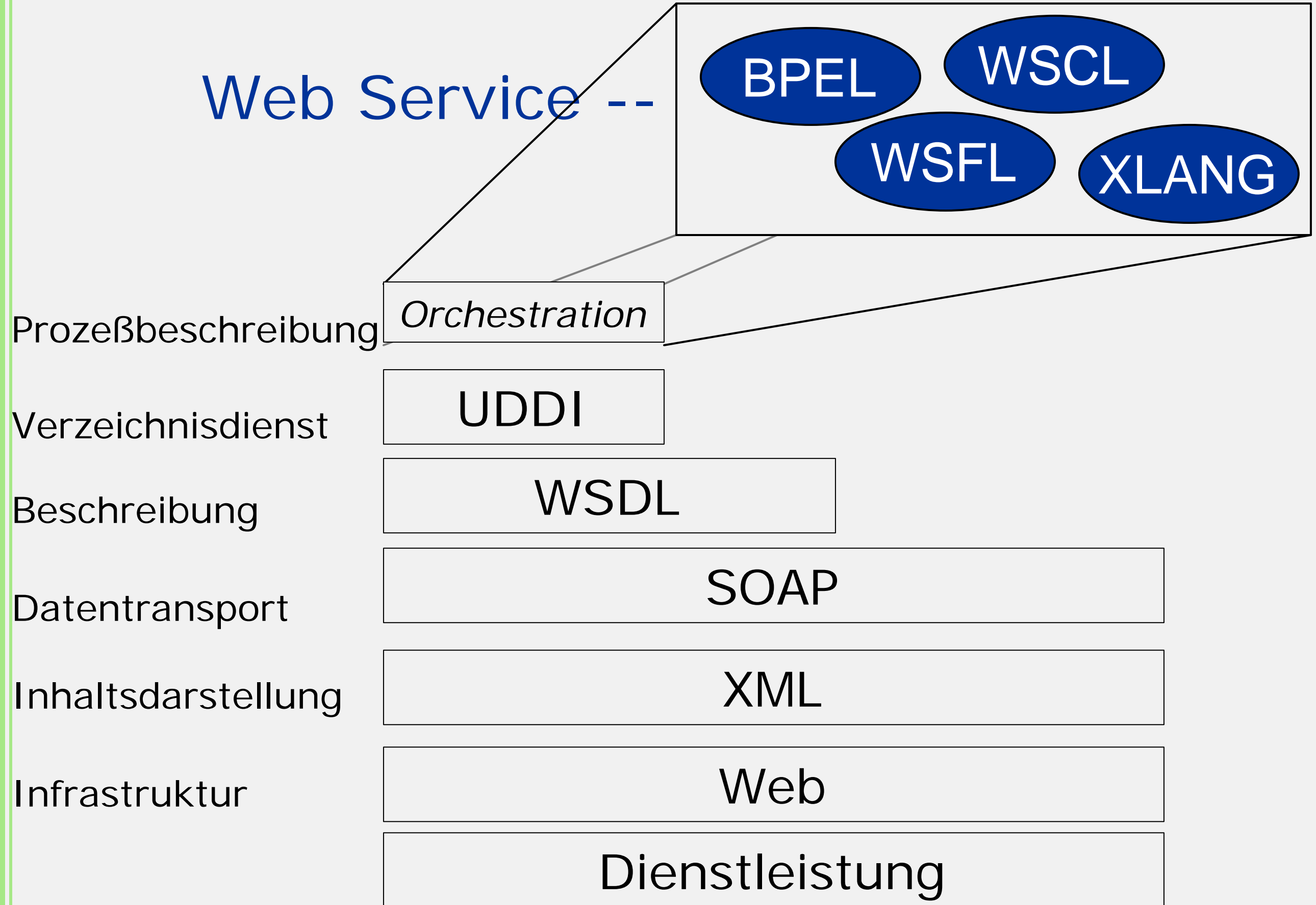
Datentransport

Inhaltsdarstellung

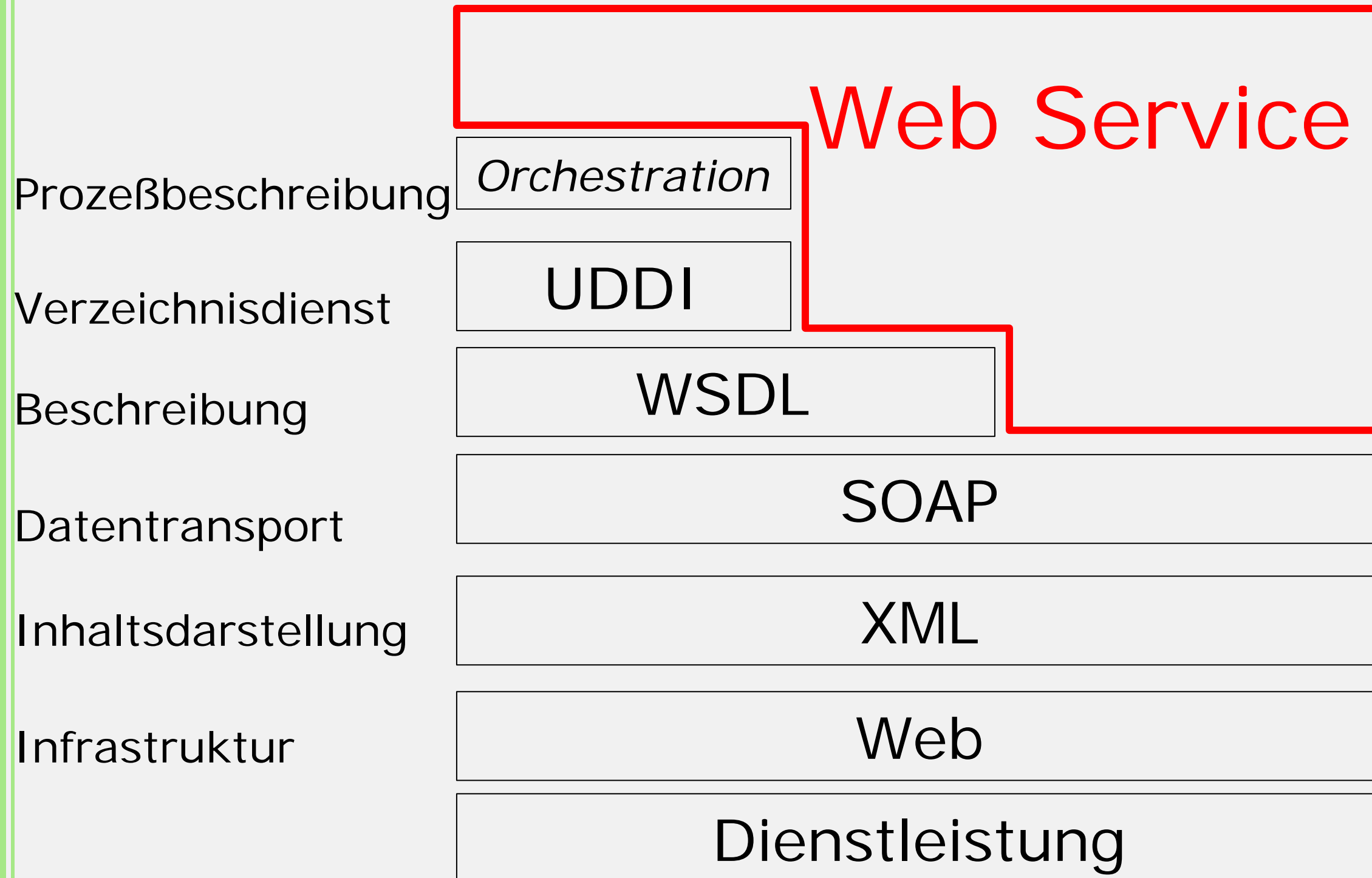
Infrastruktur



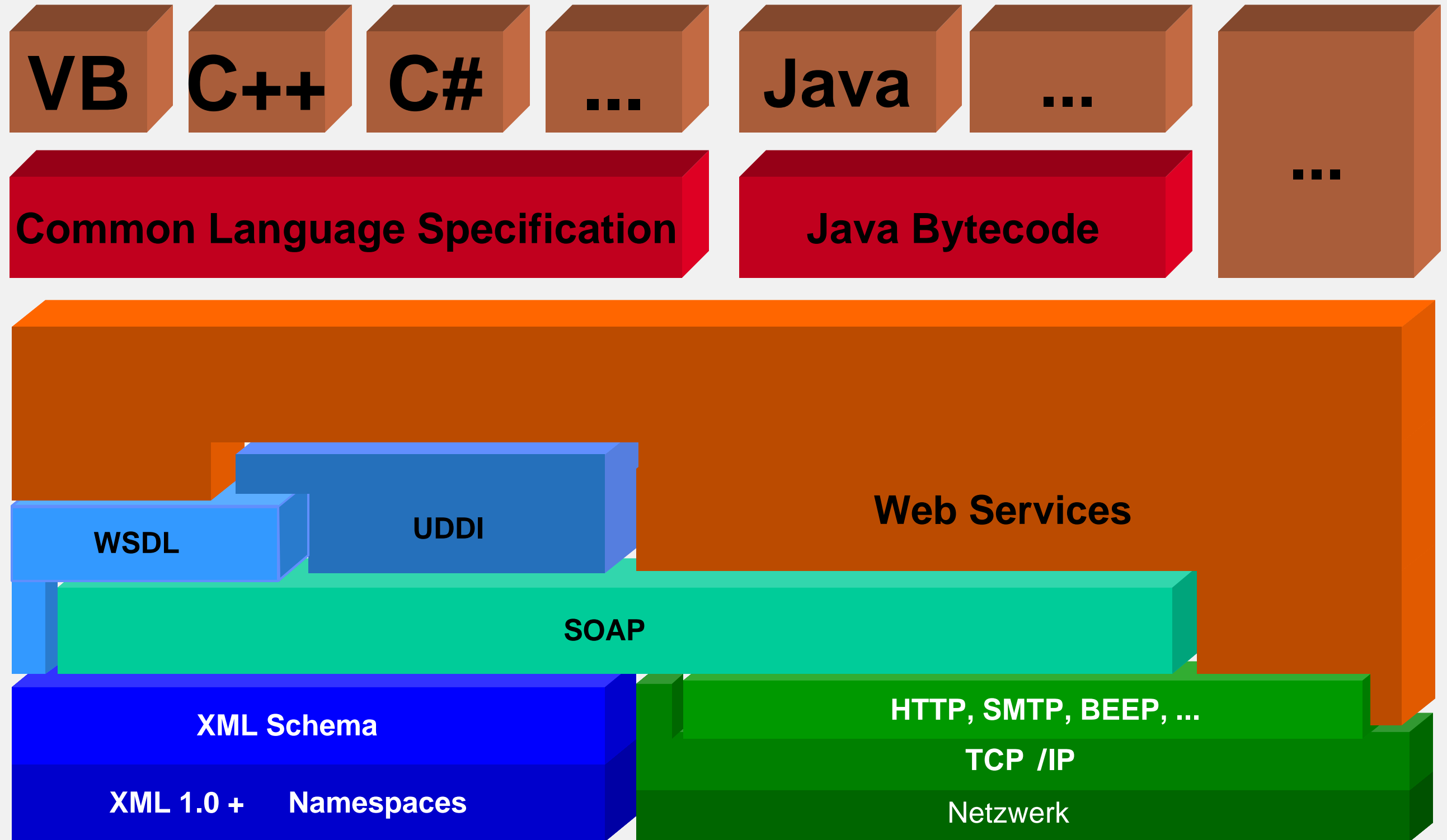
# Web Service --



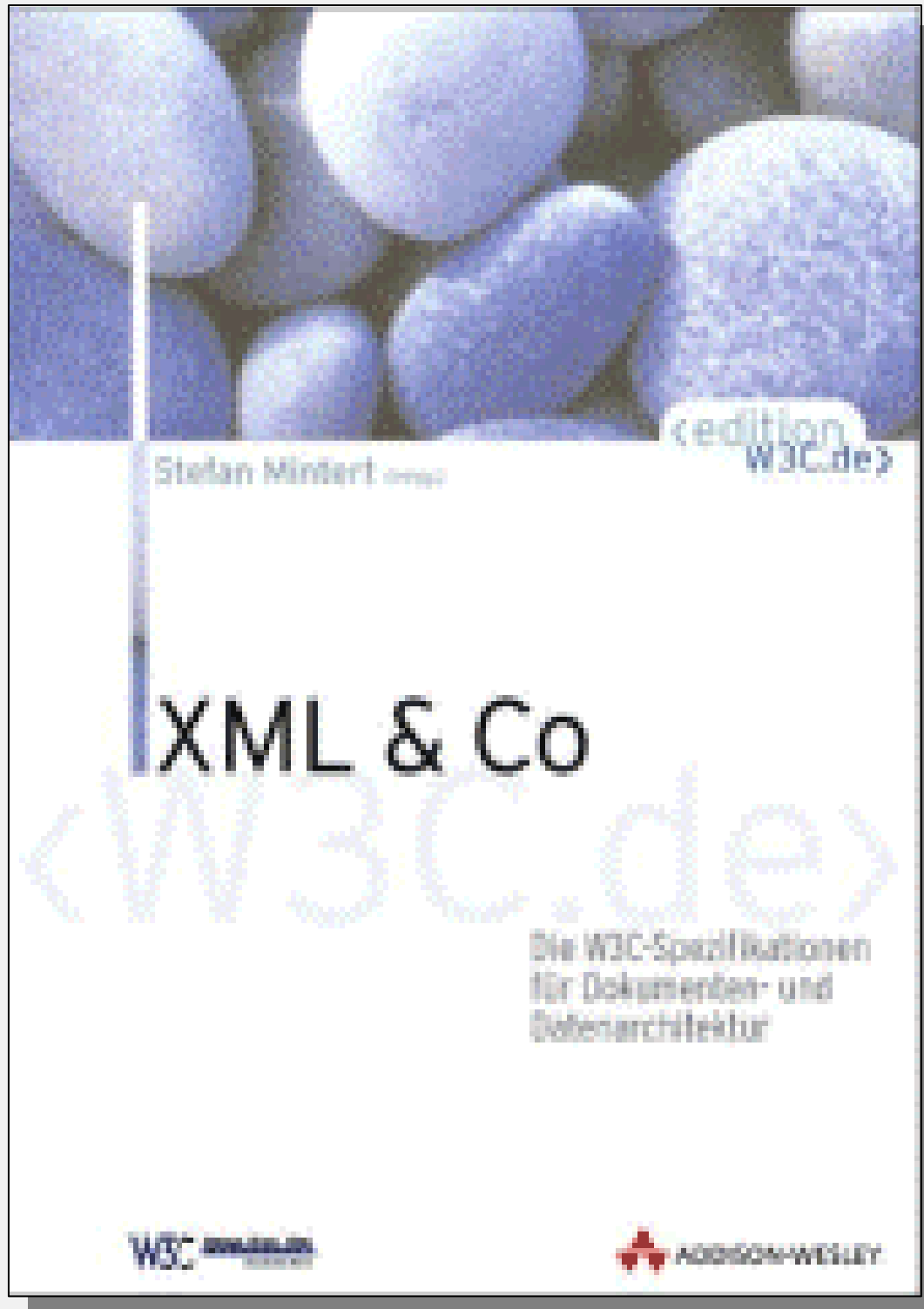
# Web Service -- Was ist das?



# Web Service -- Was ist das? Technische Sicht



# < Werbung >



Kommentierte Übersetzungen der relevanten XML-Basispezifikationen des W3C:

- XML v1.0 2<sup>nd</sup> edition
- XML Schema
- XMLBase
- XML Namensräume
- XSLT
- XPath
- InfoSet
- XLink

# < /Werbung >

# Escaping the Acronym Soup: XML Acronym Demystifier

The screenshot shows the website 'The XML Acronym Demystifier'. At the top, there's a title 'The XML Acronym Demystifier' and a search bar. Below the search bar, there's a list of XML acronyms and their full names, such as 'W3C Extensible Markup Language', 'Telecommunications Interchange Markup (T1M)', 'Scheme Markup Language (SML)', etc. A central 3D barrel graphic contains the text 'XML ACRONYM DEMYSTIFIER'. The Gartner logo is visible in the bottom right corner of the screenshot.

xml-acronym-demystifier.org