

# DAIMLERCHRYSLER

## **eXtensible Markup Language** **Die neue Sprache(n) des Web**

Mario Jeckle  
Forschungszentrum Ulm  
FT3/EK  
[mario.jeckle@daimlerchrysler.com](mailto:mario.jeckle@daimlerchrysler.com)  
[www.jeckle.de](http://www.jeckle.de)

# Inhaltsübersicht

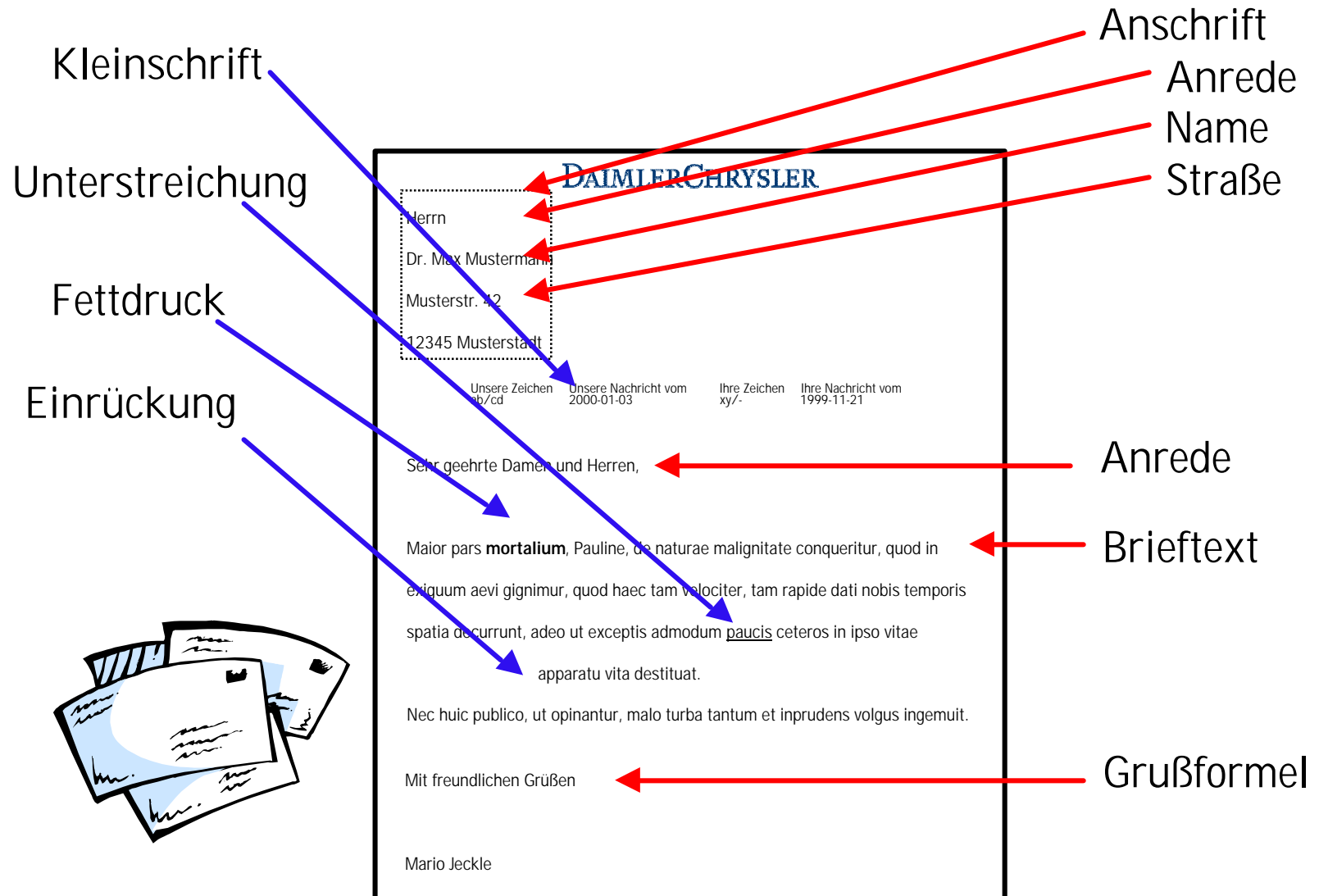
- Einführung und Grundlagen
  - Auszeichnungssprachen
  - XML und HTML
  - XML -- strukturierte Daten als Text
- XML-Technologie
- XML Sprachfamilie
  - XHTML
  - XMI
  - SVG
  - MathML
- Praktischer Einsatz
  - Tools
  - Schemadesign
  - Größenaspekte
  - einige Anwendungsbereiche

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Auszeichnungssprachen...

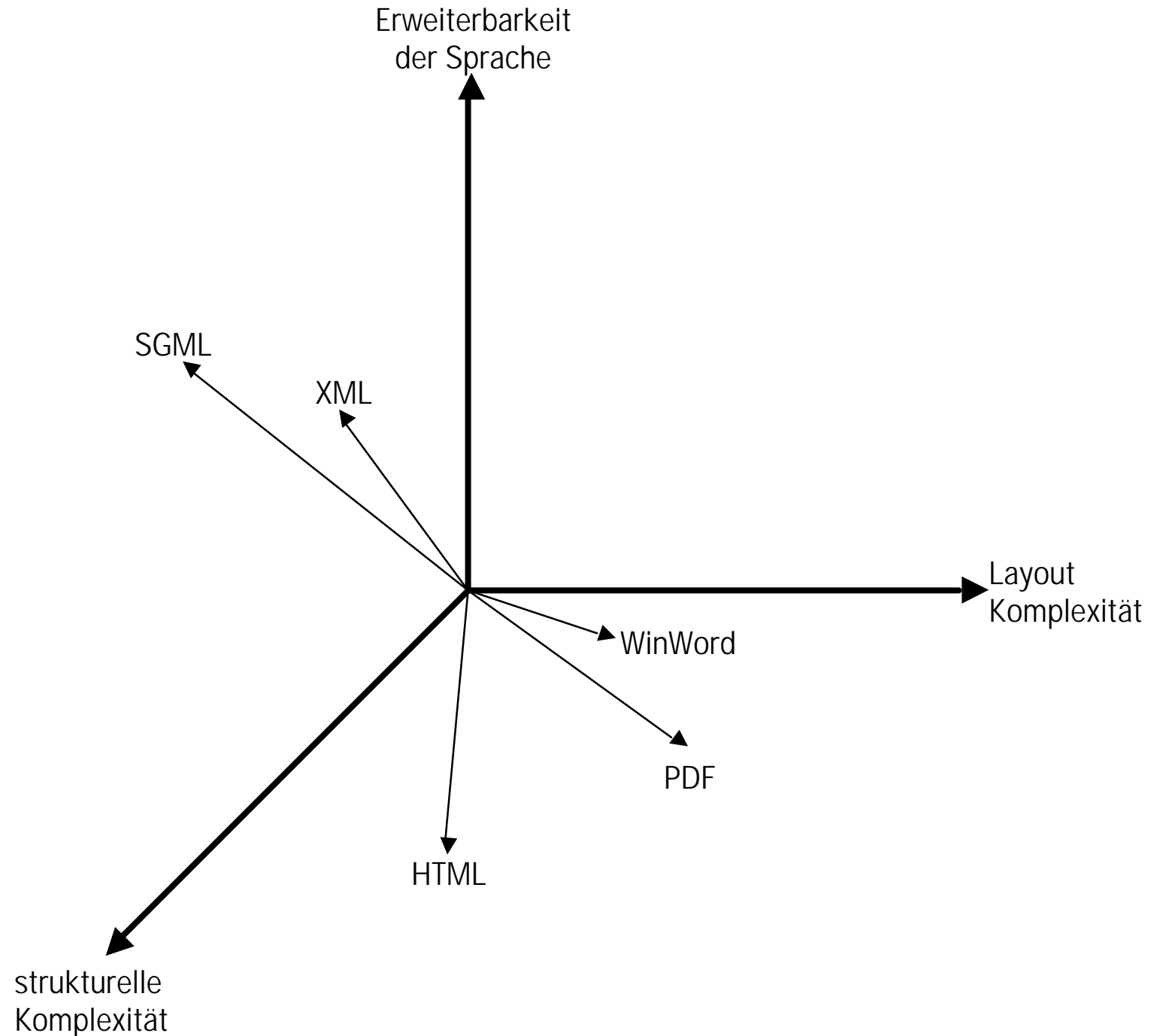
*presentational markup*

*editorial markup*



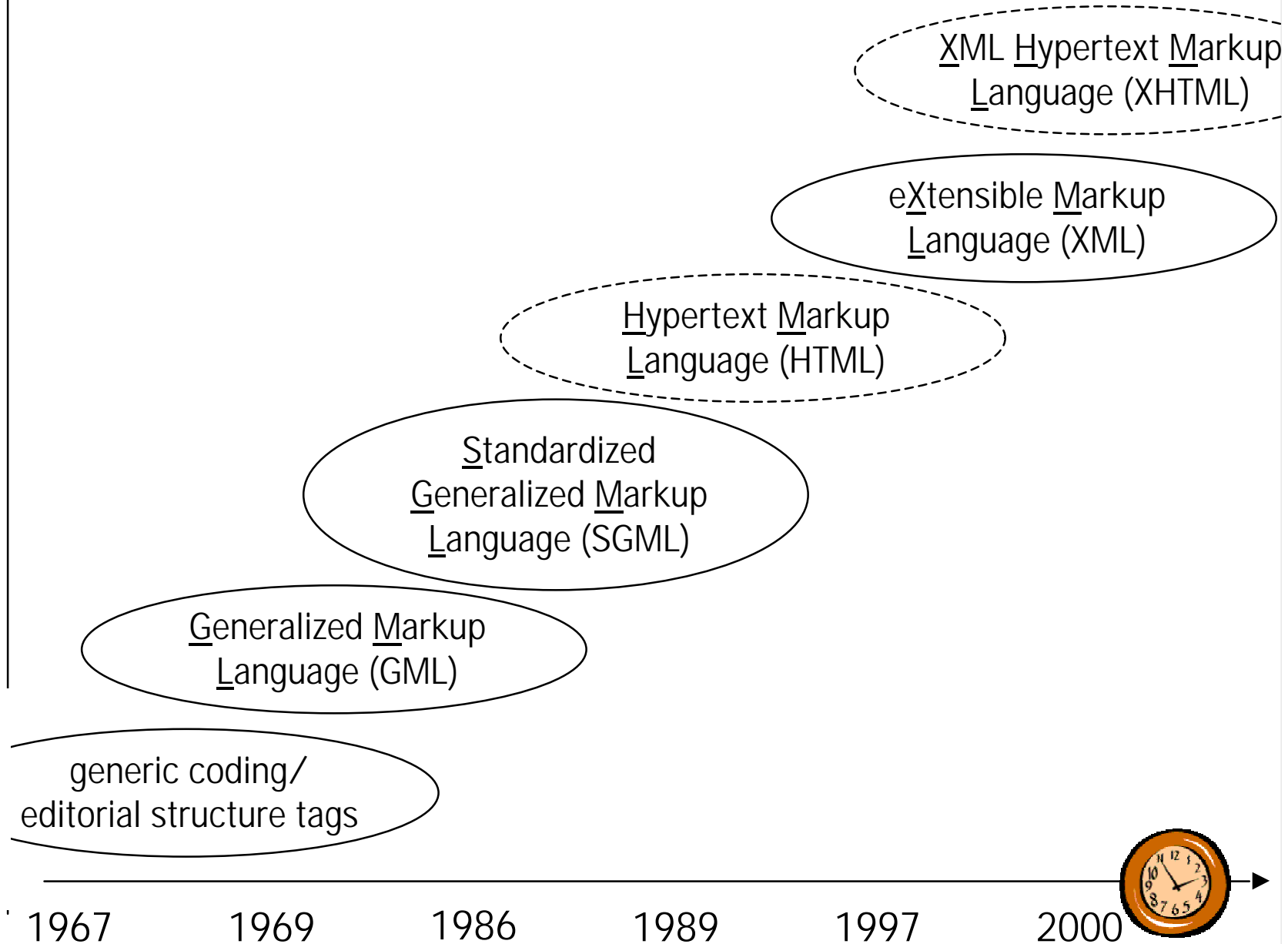
- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# World of documents...



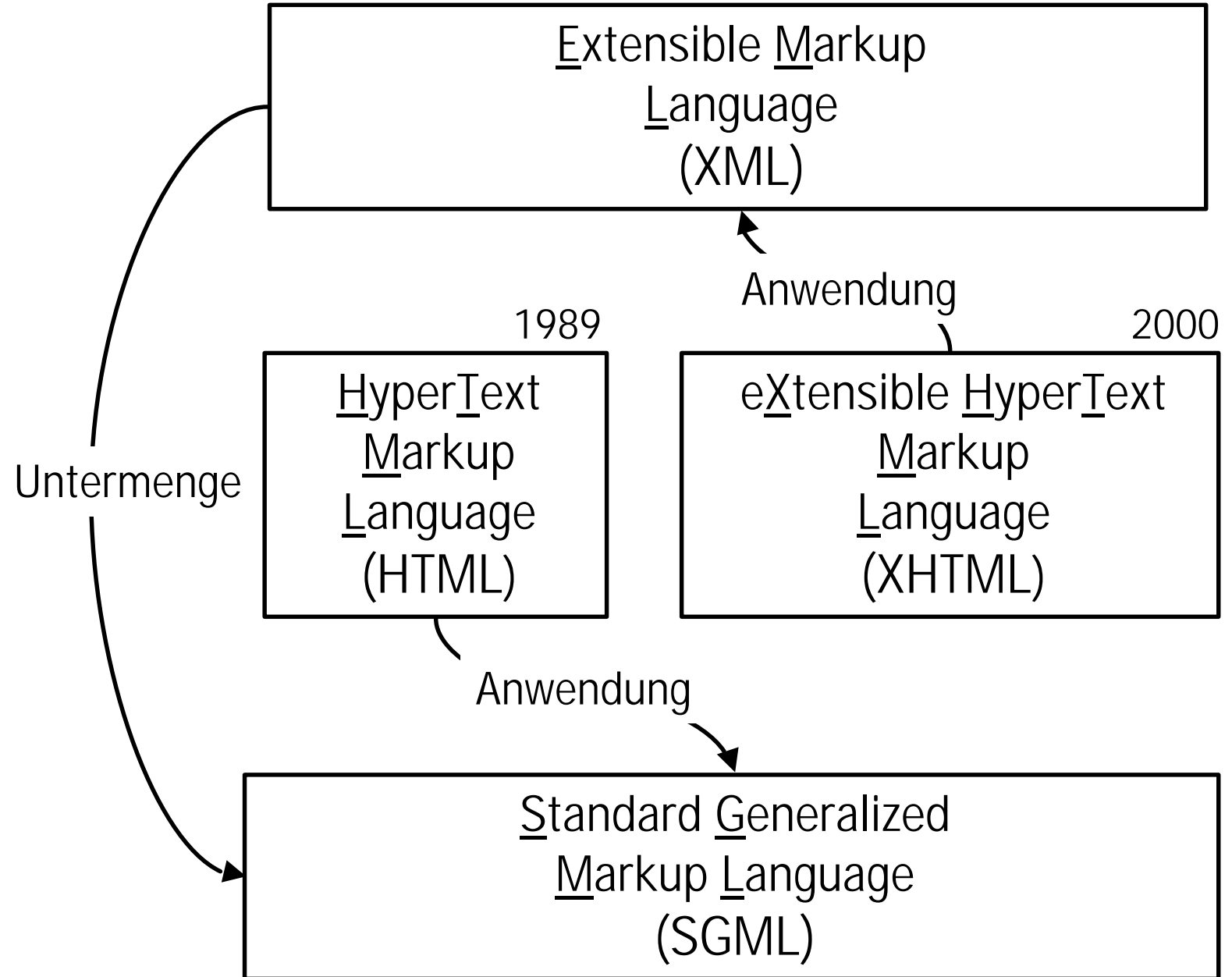
- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Historische Entwicklung der XML



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# HTML im Umfeld SGML/XML



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

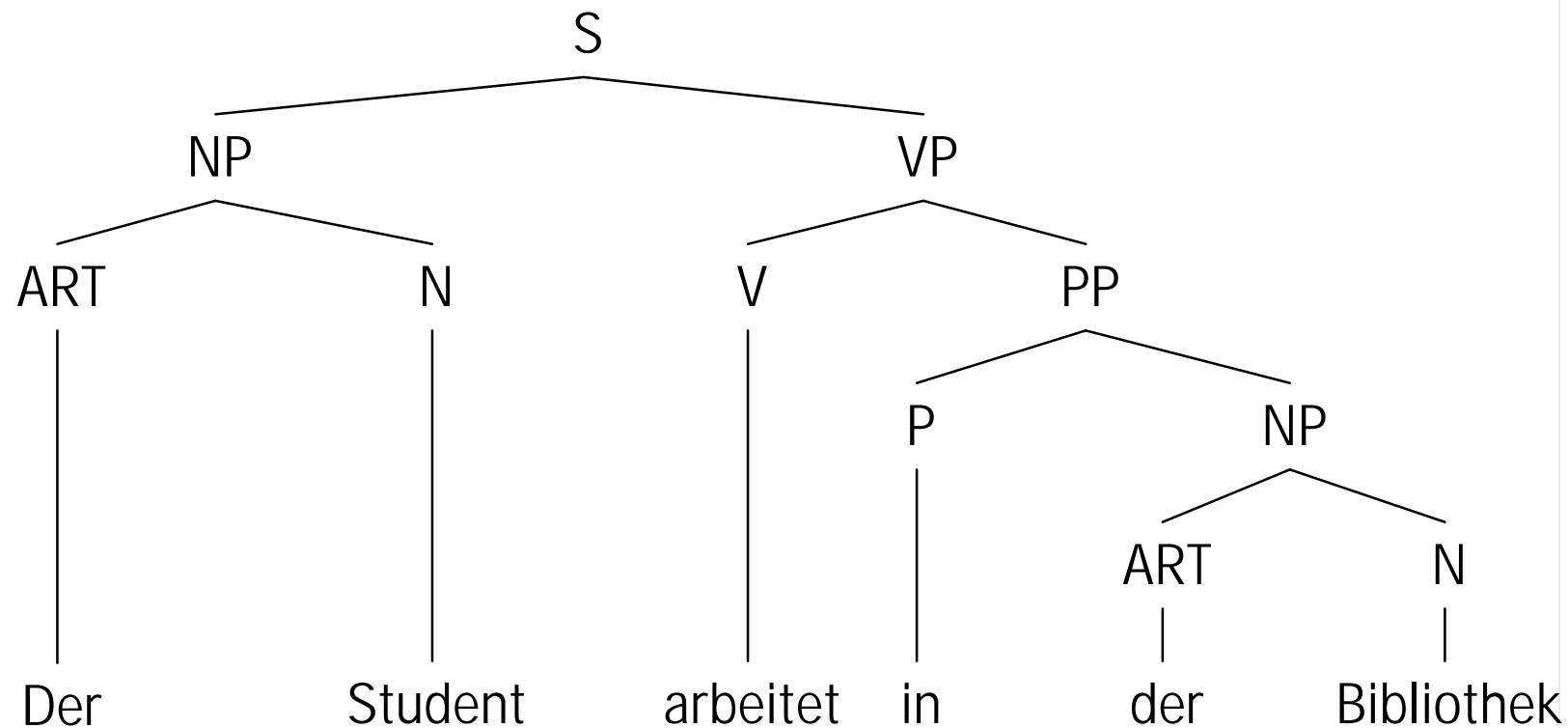
## Entwicklungsziele der XML

1. Einfache Nutzung im Internet.
2. Unterstützung eines breiten Anwendungsspektrums.
3. SGML Kompatibilität.
4. Einfache Applikationsentwicklung.
5. Minimierung optionaler Sprachmerkmale; Idealerweise gleich Null.
6. Lesbarkeit (*human readability*).
7. Kompakte Spezifikation.
8. Formaler und präziser Sprachentwurf.
9. Leichte Dokumenterstellung.
10. Nicht notwendigerweise knappes Markup.

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Konstituenten-Analyse

**Prinzip:** Der Satz (S) wird in die Kategorie Nominalphrase (NP) und Verbalphrase (VP) geteilt. Die Nominalphrase wird aufgespalten in Artikel (ART) und Nomen (N).



## Äquivalente Klammerdarstellung:

[S[NP[ART **Der**]ART[N **Student**]N]NP[VP[V **arbeitet**]V[PP[P **in**]P  
[NP[ART **der**]ART [N **Bibliothek**]N]NP]PP]VP]S

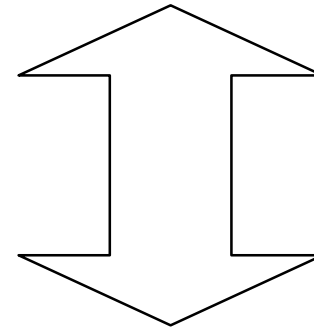


## eXtensible Markup Language

die neue Sprache des Web

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

[S[NP[ART **Der**]ART[N **Student**]N]NP[VP[V **arbeitet**]V[PP[P **in**]P]  
[NP[ART **der**]ART[N **Bibliothek**]N]NP]PP]VP]S

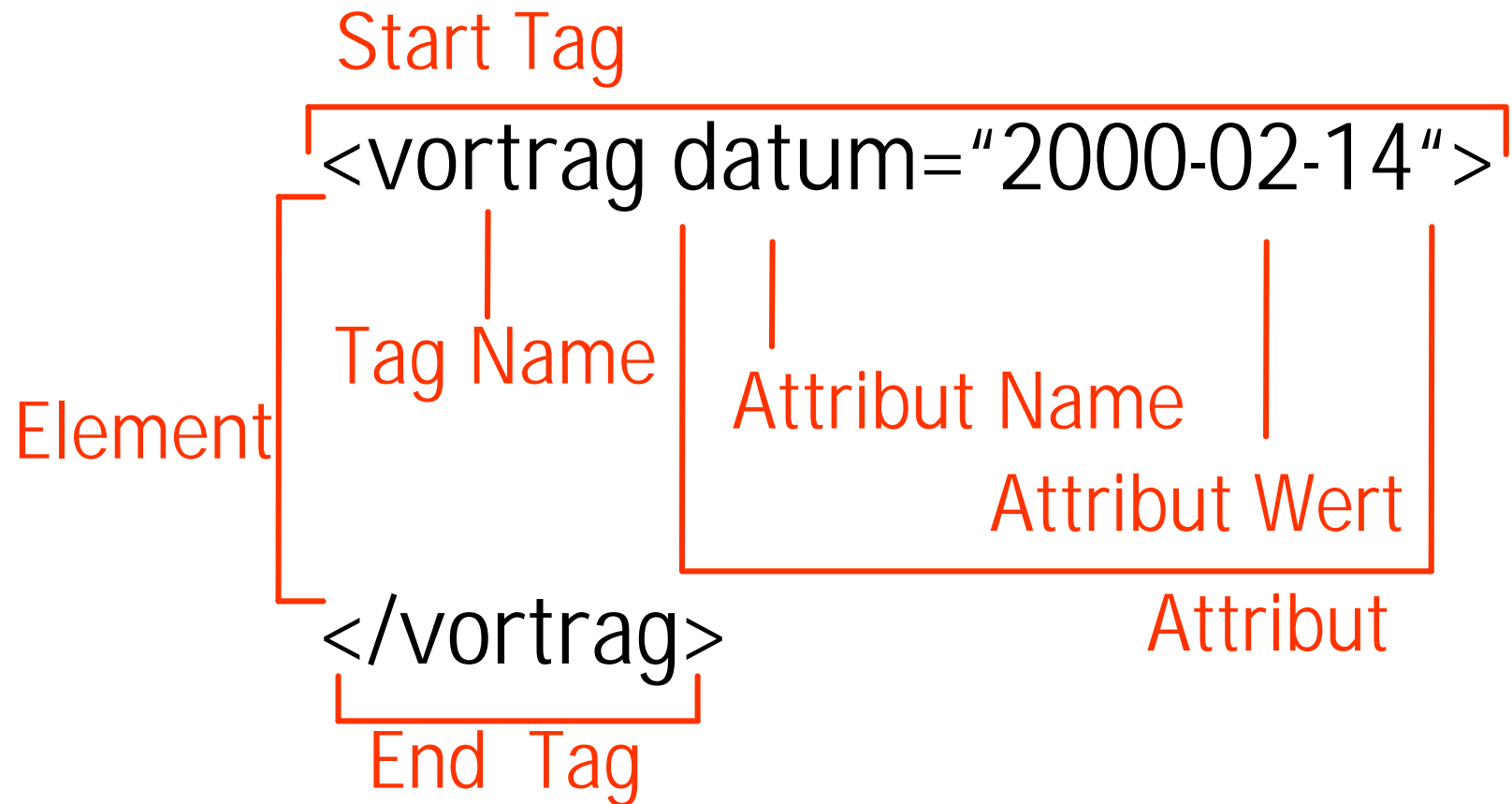


```
<S>  
  <NP>  
    <ART>Der</ART>  
    <N>Student</N>  
  </NP>  
  <VP>  
    <V>arbeitet</V>  
    <PP>
```

```
<P>in</P>  
  <NP>  
    <ART>der</ART>  
    <N>Bibliothek</N>  
  </NP>  
</PP>  
</VS>  
</S>
```

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# XML-Terminologie



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# XML-Grundbegriffe

## XML Dokument

```
<?xml version = "1.0"?>  
<!DOCTYPE vortrag SYSTEM "vortrag.dtd">  
<vortrag datum="2000-02-14">  
    <abstract>  
        Die eXtensible Markup ....  
    </abstract>  
</vortrag>
```

## XML Document Type Definition (DTD)

```
<?xml version="1.0" encoding="UTF-8" ?>  
<!ELEMENT vortrag (abstract)>  
<!ATTLIST vortrag  
    datum CDATA #REQUIRED>  
<!ELEMENT abstract ANY>
```

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Dokumentvalidierung

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
  xml:lang="en" lang="en">
  <head>
    <title>MyPage</title>
  </head>
  <body>
    content
  </body>
</html>
```

XML Dokument

```
<ENTITY % HTMLat1 PUBLIC
  "-//W3C//ENTITIES Latin 1 for XHTML//EN"
  "xhtml-lat1.ent">
%HTMLat1;

<ENTITY % HTMLsymbol PUBLIC
  "-//W3C//ENTITIES Symbols for XHTML//EN"
  "xhtml-symbol.ent">
%HTMLsymbol;

<ENTITY % HTMLspecial PUBLIC
  "-//W3C//ENTITIES Special for XHTML//EN"
  "xhtml-special.ent">
%HTMLspecial;

<ENTITY % ContentType "CDATA">
<!-- media type, as per [RFC2045] -->

<ENTITY % ContentTypes "CDATA">
<!-- comma-separated list of media types, as per [RFC2045] -->
```

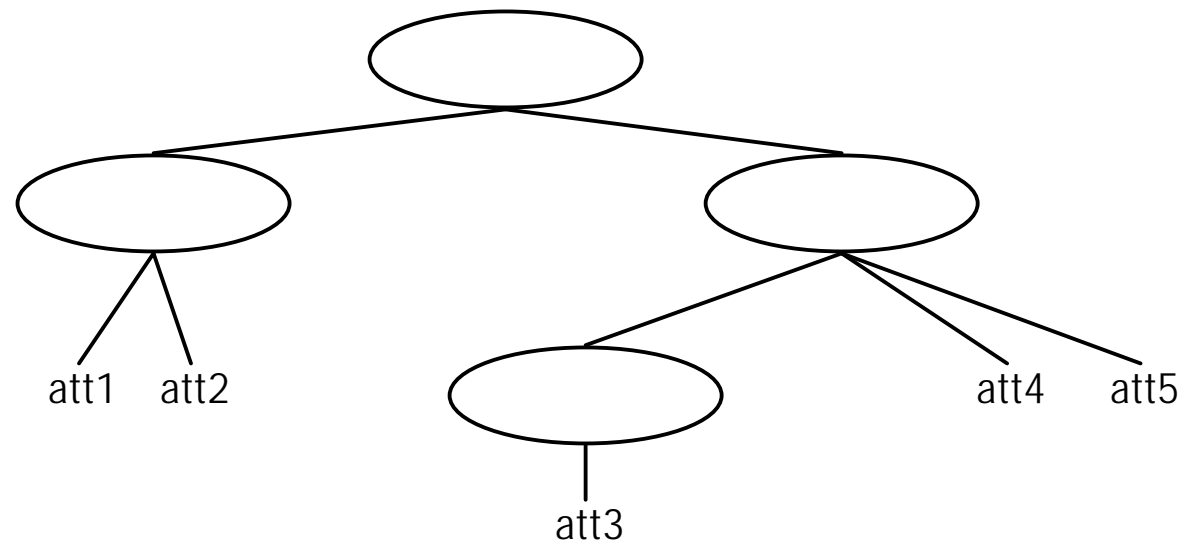
DTD des Dokuments



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

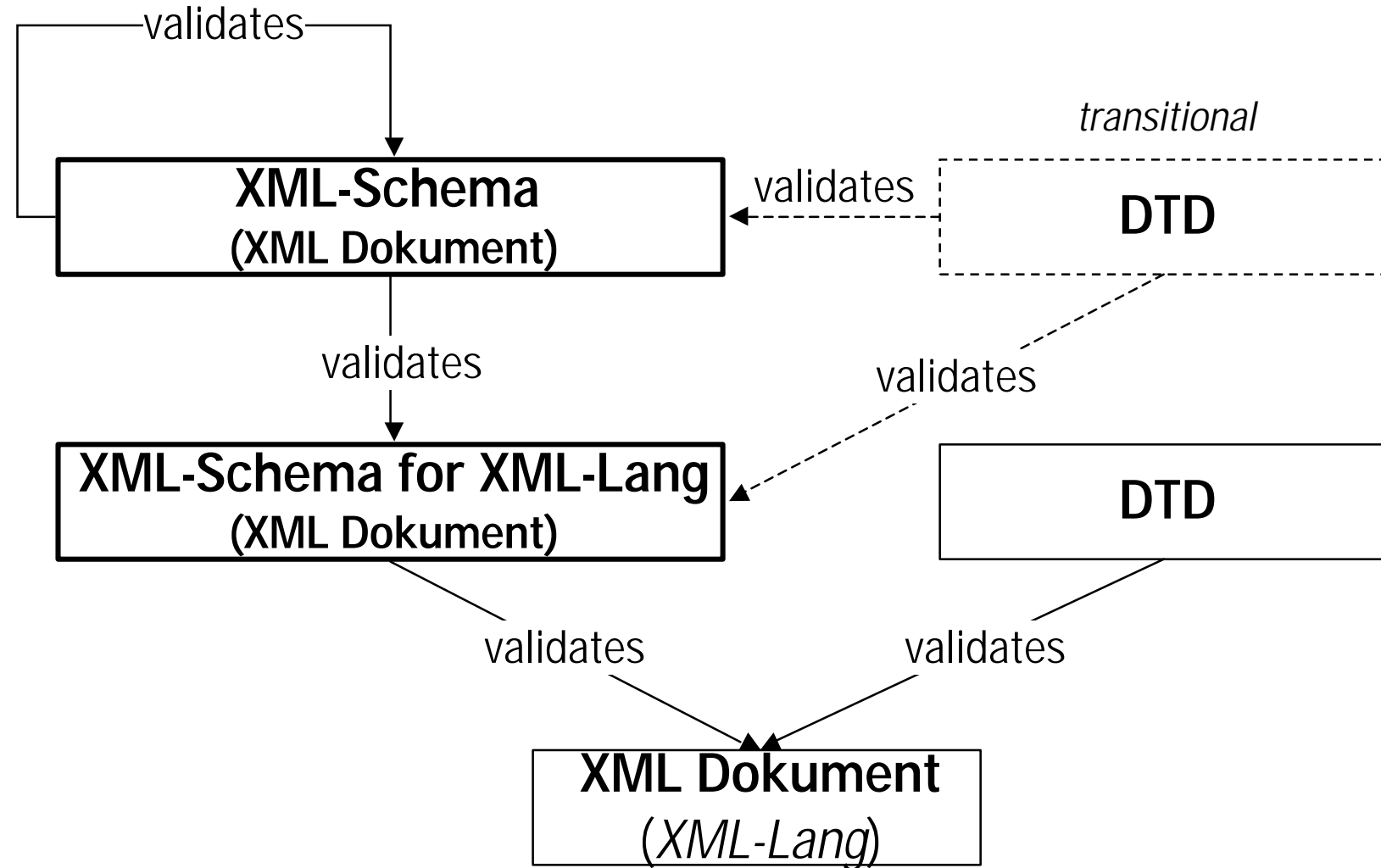
## DTD-Mechanismus in XML v1.0

- streng hierarchisch
- *ELEMENTs* als innere Knoten
- *ATTLISTs* zur Attributierung der Knoten
- Keine Datentypen (abgesehen von CHAR-Data)
- Rudimentärer Referenzierungsmechanismus
- Selektionstyp (Enumeration)
- Vorgabewerte
- DTD ist nicht XML!



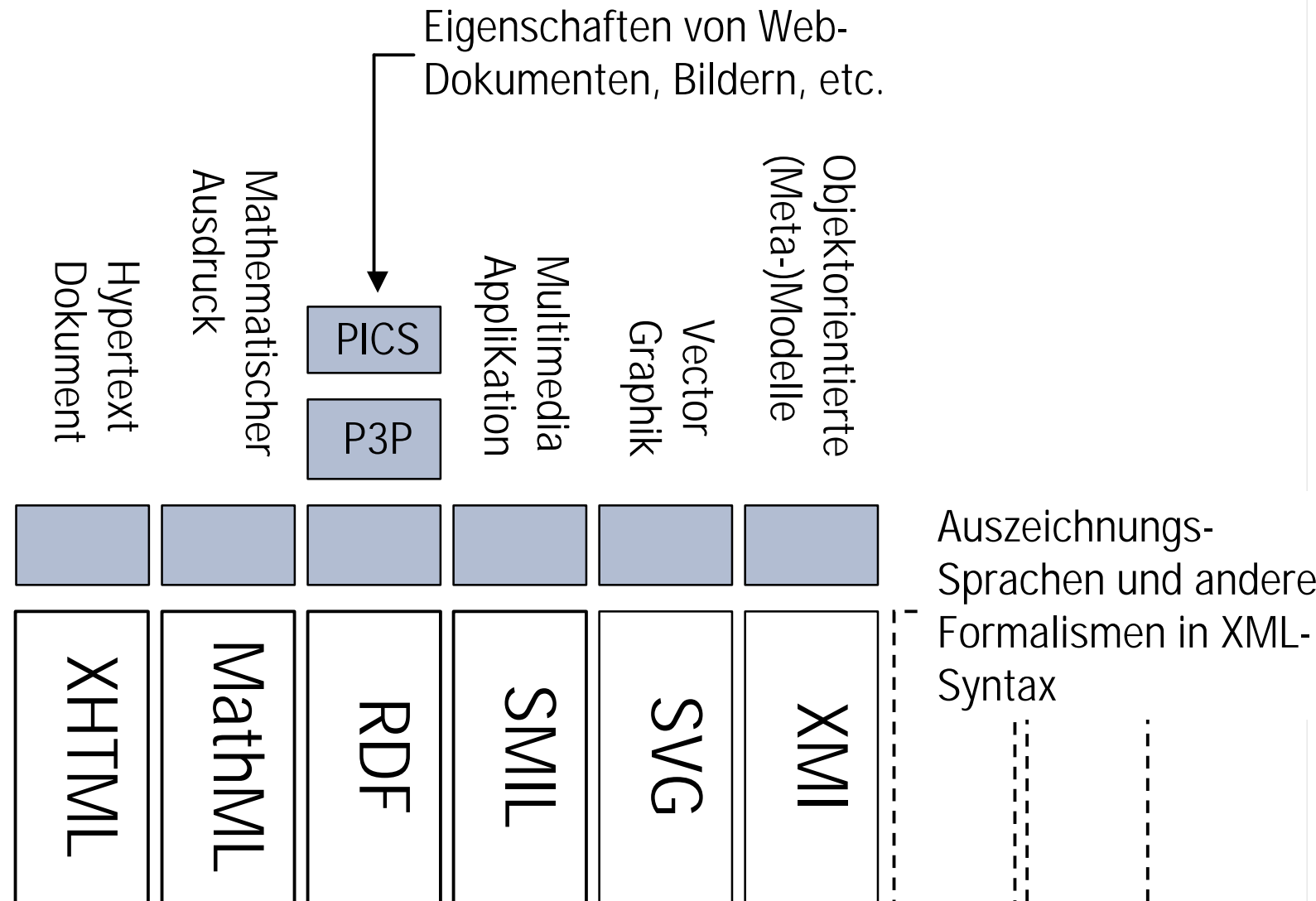
- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Zukünftige Spracharchitektur



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Familie der XML-Sprachen



eXtensible Markup Language (XML)

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# eXtensible HyperText Markup Language (XHTML)

- Keine überlappenden Elemente

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

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

- 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
```



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# eXtensible HyperText Markup Language (XHTML)

- Attributminimierung

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

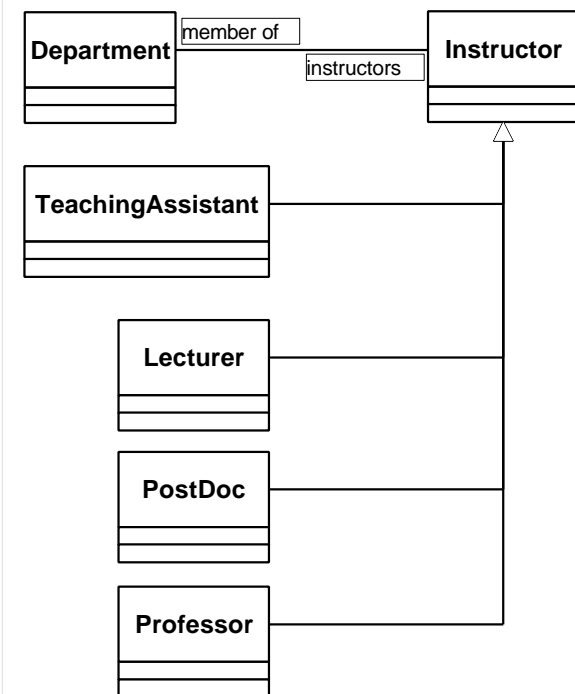
```
<dl compact>
```

- Leere Elemente

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

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

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz



# XML Metadata Interchange (XMI)

```

<XMI version="1.1" xmlns:UML="org.omg/UML1.3">
  <XMI.header>
    <XMI.model xmi.name="Department" href="Department.xml"/>
    <XMI.metamodel xmi.name="UML" href="UML.xml"/>
  </XMI.header>
  <XMI.content>
    <UML:Class name="Department" xmi.id="Department"/>
    <UML:Class name="Instructor" xmi.id="Instructor"/>
    <UML:Class name="Professor" xmi.id="Professor"
      generalization="Instructor"/>
    <UML:Class name="Postdoc" xmi.id="Postdoc" generalization="Instructor"/>
    <UML:Class name="Lecturer" xmi.id="Lecturer" generalization="Instructor"/>
    <UML:Class name="TeachingAssistant" xmi.id="TeachingAssistant"
      generalization="Instructor"/>
    <UML:Association>
      <UML:Association.connection>
        <UML:AssociationEnd name="instructors" type="Instructor"/>
        <UML:AssociationEnd name="memberOf" type="Department"/>
      </UML:Association.connection>
    </UML:Association>
  </XMI.content>
</XMI>
    
```

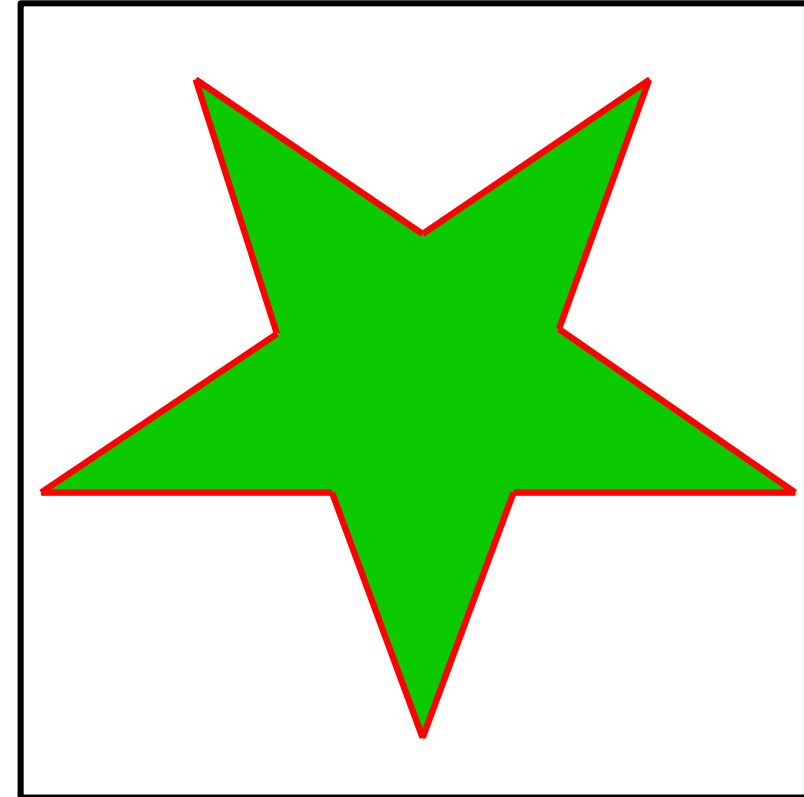
- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Scalable Vector Graphics (SVG)

```
<v:shape style="top: 0; left: 0;
width: 175; height: 175"
stroke="true" strokecolor="red"
strokeweight="2" fill="true"
fillcolor="green"
coordorigin="0 0"
coordsize="175 175">
```

```
<v:path v="m 8,65
l 72,65,92,11,112,65,174,65,
122,100,142,155,92,121,
42,155,60,100 x e"/>
```

```
</v:shape>
```



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Mathematical Markup Language (MathML)

```

<mrow>
  <mi>x</mi>
  <mo>=</mo>
  <mfrac>
    <mrow>
      <mrow>
        <mo>-</mo>
        <mi>b</mi>
      </mrow>
      <mo>&PlusMinus;</mo>
    </mrow>
    <msqrt>
      <mrow>
        <msup>
          <mi>b</mi>
          <mn>2</mn>
        </msup>
        <mo>-</mo>
      </mrow>
    </msqrt>
  </mfrac>
</mrow>

```

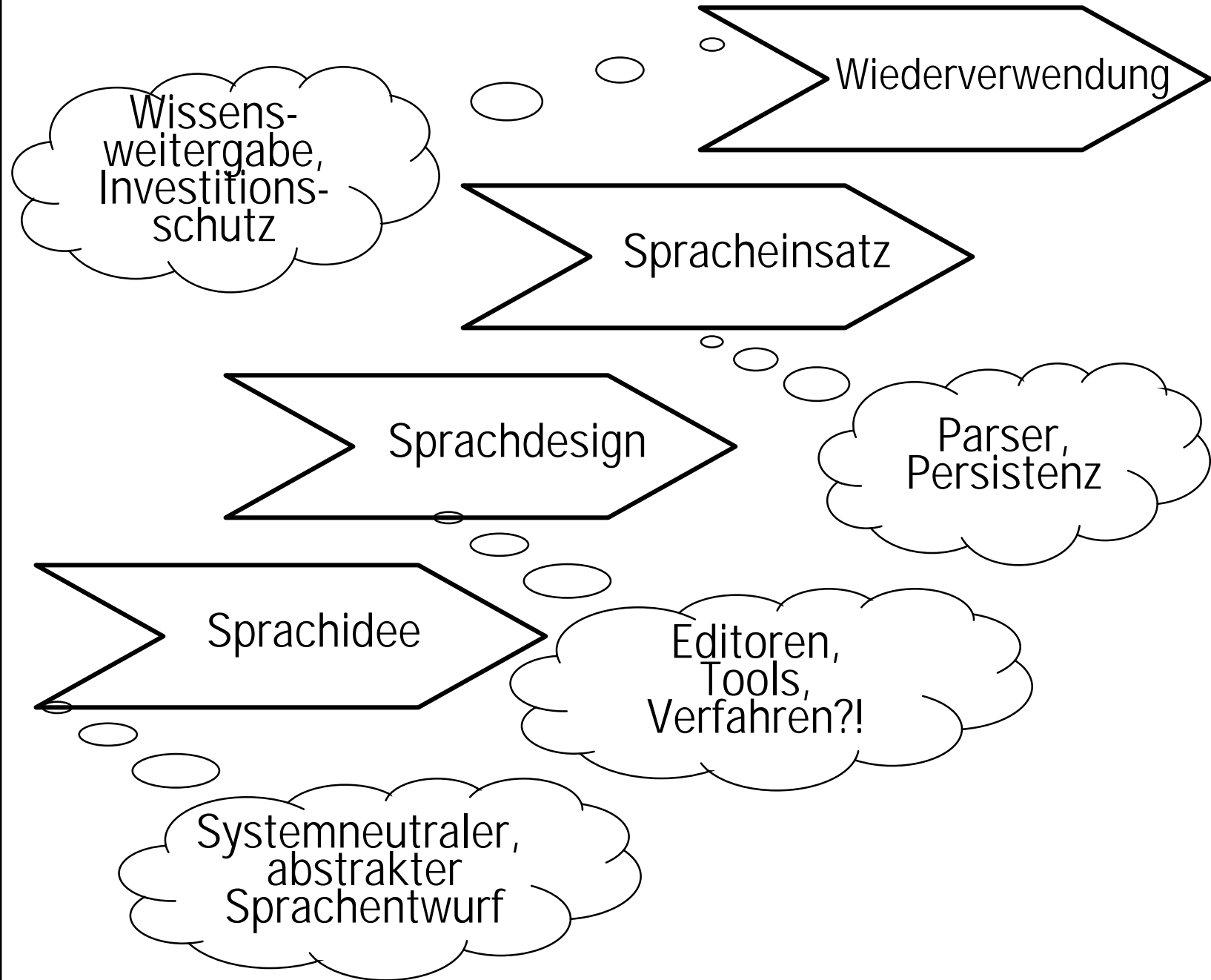
```

<mrow>
  <mn>4</mn>
  <mo>&InvisibleTimes;</mo>
  <mi>a</mi>
  <mo>&InvisibleTimes;</mo>
  <mi>c</mi>
</mrow>
</msqrt>
</mrow>
<mrow>
  <mn>2</mn>
  <mo>&InvisibleTimes;</mo>
  <mi>a</mi>
</mrow>
</mfrac>
</mrow>

```

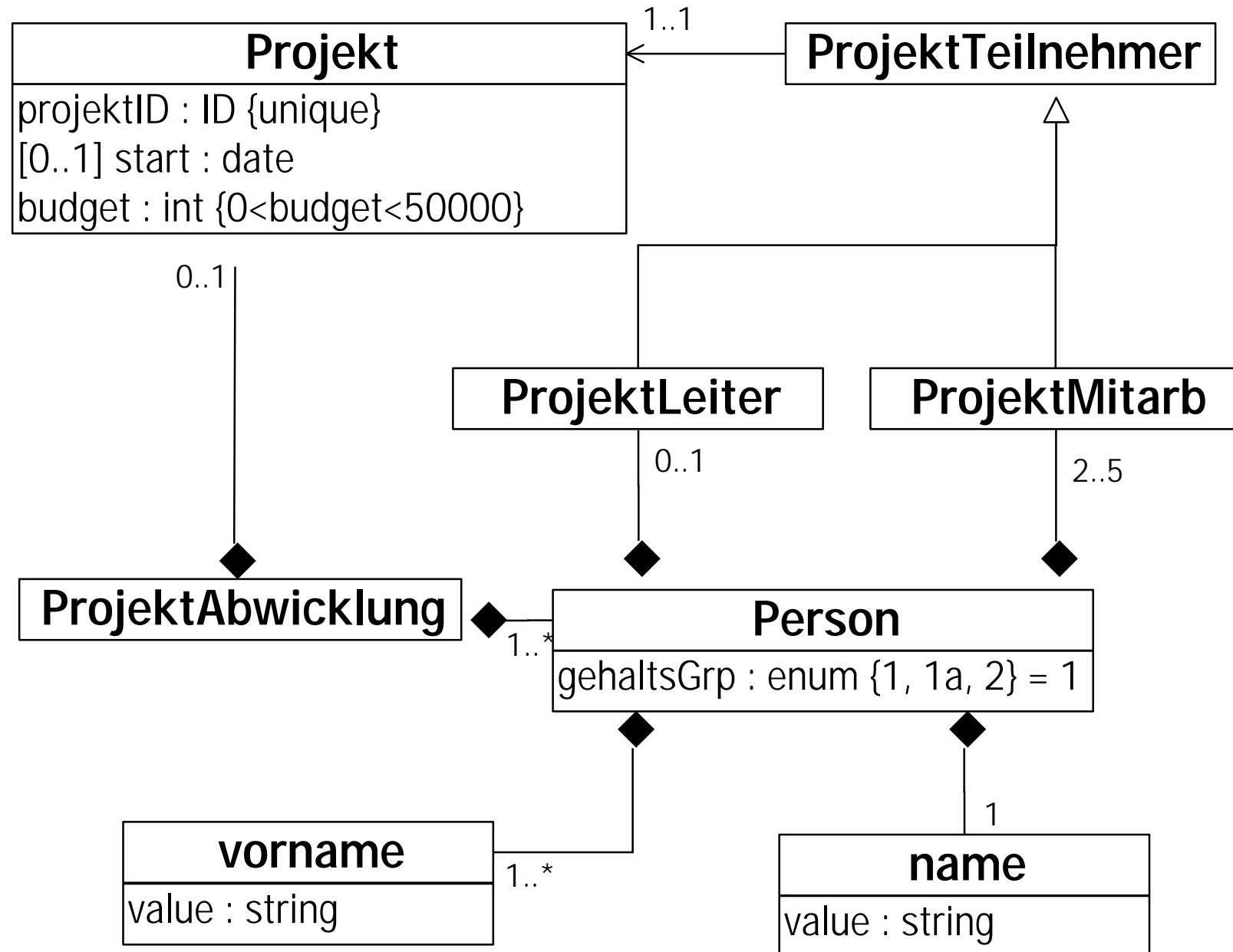
$$\text{"Mitternachtsformel"} x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz



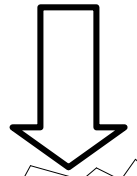
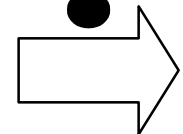
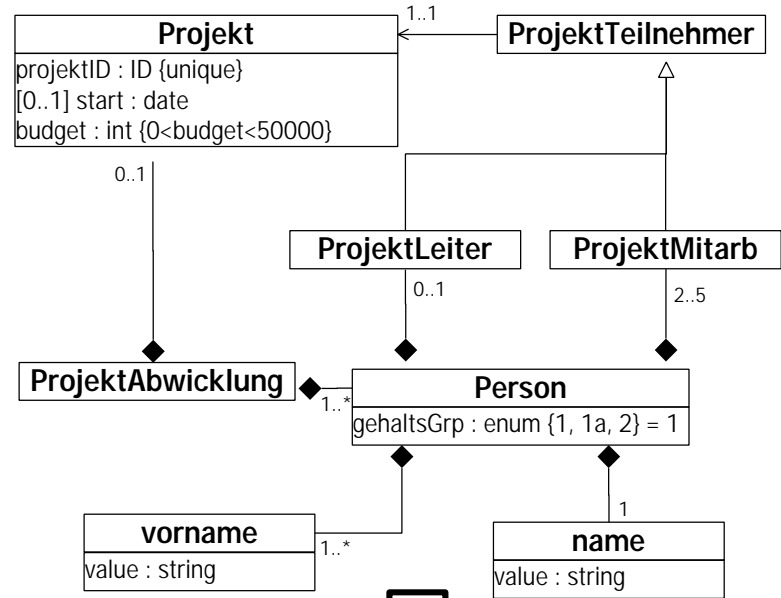
- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

# Sprachidee und -design



- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz

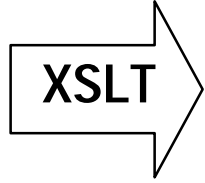
# Spachidee- und design



```

<?xml version = "1.0"?>
<!DOCTYPE XMI SYSTEM "uml.dtd">
<XMI xmi.version="1.0">
  <XMI.header>
    <XMI.metamodel xmi.name="uml" xmi.version="1.3" />
  </XMI.header>
  <XMI.content>
    <Model_Management.Model xmi.id="i00000001">
      <Foundation.Core.ModelElement.name>....</>
      <Foundation.Core.ModelElement.visibility xmi.value="public"/>
      <Foundation.Core.GeneralizableElement.isRoot xmi.value="true"/>
      <Foundation.Core.GeneralizableElement.isLeaf xmi.value="true"/>
      <Foundation.Core.GeneralizableElement.isAbstract xmi.value="false"/>
    </Model_Management.Model>
  </XMI.content>
</XMI>
  
```

XMI-Dokument



```

<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE schema SYSTEM "schema1.dtd">
<schema targetNamespace="http://purl.org/metadata/dublin_core"
  version="1.0" xmlns="http://www.w3.org/1999/XMLSchema">
  <datatype name="budgetType" source="integer"
    <minExclusive value="0"/>
    <maxExclusive value="50000"/>
  </datatype>
  <datatype name="gehaltsGrpType" source="string">
    <enumeration value="1"/>
    <enumeration value="1a"/>
    <enumeration value="2"/>
  </datatype>
</schema>
  
```

eXtensible Markup Language

die neue Sprache des Web

- ▶ Einführung
- ▶ Technologie
- ▶ Sprachfamilie
- ▶ praktischer Einsatz
- ▶ Referenzen

[www.jeckle.de](http://www.jeckle.de)

[www.xml.com](http://www.xml.com)

[www.xmlhack.com](http://www.xmlhack.com)

XML-Home:

[www.w3.org/TR/REC-xml](http://www.w3.org/TR/REC-xml)

[www.mintert.com/xml/trans](http://www.mintert.com/xml/trans)

Parser:

[java.sun.com/xml](http://java.sun.com/xml) (*Java Project X*)

[xml.apache.org](http://xml.apache.org) (*Xerces/IBM's XML4J*)