

# XSLT Workshop: Synopsen mit XSLT

Mathias Göbel und Philipp Vanscheidt  
TextGrid-Nutzertreffen, Freiburg, 2014

# Agenda

## 1. Was können wir mit XSLT alles machen?

### 1. Beispiele:

1. Webseiten
2. Textprozessierung
3. Datenbankabfragen
4. ...

## 2. Theorie

1. XML
2. XPath + Übung
3. XSLT
4. Wie verhalten sich beide zueinander?

## 3. XSLT im TextGrid Lab

1. *processing instruction*
2. Transformieren
3. Was nun?

## 4. Hands on!

1. Briefbeispiel
2. Synopse Bild-Text

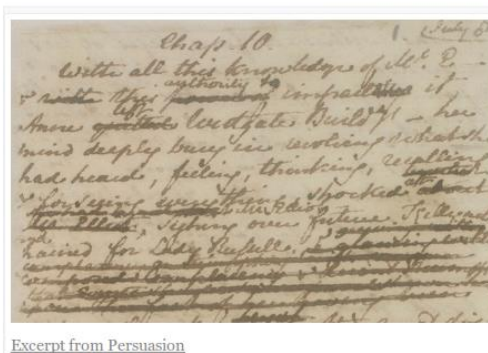
- Webseiten erstellen
- Text oder Textteile extrahieren
- Datenbankabfragen erstellen
- nentral nefual sträwkcür txeT
- XPath-Funktionen anwenden
- Dokumente lokal schreiben
- ...

# Beispiele: Webseiten mit XSLT



## Home

The **Jane Austen's Fiction Manuscripts Digital Edition** gathers together in the virtual space of the web some 1100 pages of fiction written in Jane Austen's own hand. Through digital reunification, it is now possible to access, read, and compare high quality images of original manuscripts whose material forms are scattered around the world in libraries and private collections. Unlike the famous printed novels, all published in a short span between 1811 and 1818, these manuscripts trace Jane Austen's development as a writer from childhood to the year of her death; that is, from 1787 (aged 11 or 12) to 1817 (aged 41). Not only do they provide a unique visual record of her imagination from her teenage experiments to her last unfinished writings, these pages represent one of the earliest collections of creative writings in the author's hand to survive for a British novelist.



Excerpt from Persuasion

and physical description. Click on → [Manuscripts](#) in the main menu.

The main resources in the digital edition are the manuscripts themselves, accessed through the manuscript menu and grouped chronologically and according to their material states as fair copies (the → [Juvenilia](#) and → [Lady Susan](#)), working drafts (→ [The Watsons](#), → [Persuasion](#), → [Sanditon](#)), and further materials. Each manuscript can be opened in a variety of ways: as facsimile pages which can be magnified using 'zoomify'; as transcribed text set side by side with the original manuscript page; or through a Headnote that provides details of the manuscript's history



**TextGrid**

Virtuelle Forschungsumgebung  
für die Geisteswissenschaften

- TEI Stylesheets (XSLT-Sammlung zur Aufbereitung von TEI-Dokumenten)
  - Anwendungsbeispiel:

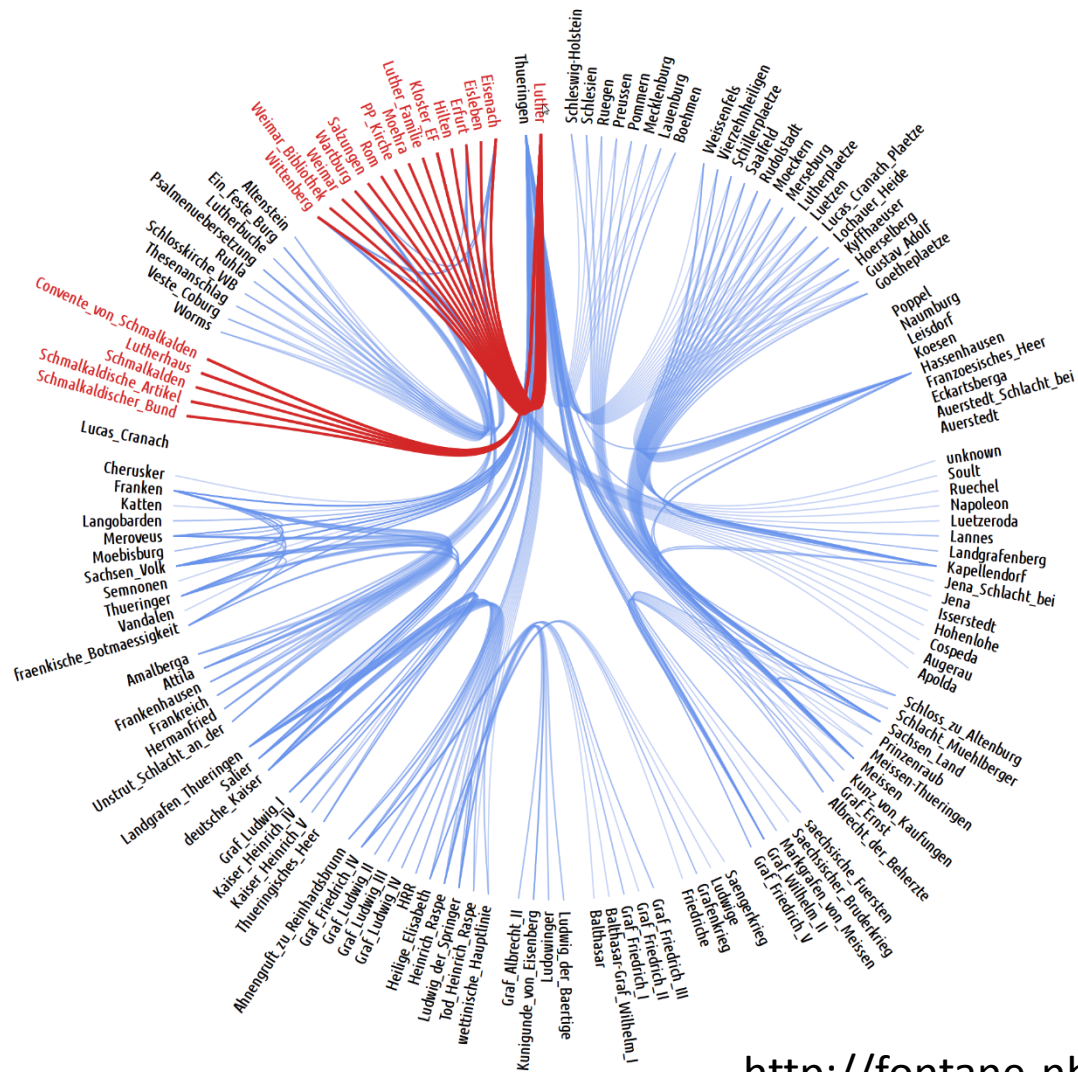


## Beispiele: Text extrahieren



- GND:
  - Geburtsdaten abfragen:
    - `document(concat('http://d-nb.info/', $gndid, '/about/rdf'))//gnd:dateOfBirth`
- OSM:
  - Geokoordinaten abfragen:
    - `document(concat('http://nominatim.openstreetmap.org/search/', $place_name, '?format=xml'))`

## Beispiele: Netzwerk



<http://fontane-nb.dariah.eu/tei-conf/>

# Agenda

## 1. Was können wir mit XSLT alles machen?

### 1. Beispiele:

1. Webseiten
2. Textprozessierung
3. Datenbankabfragen
4. ...

## 2. Theorie

1. XML
2. XPath + Übung
3. XSLT
4. Wie verhalten sich beide zueinander?

## 3. XSLT im TextGrid Lab

1. *processing instruction*
2. Transformieren
3. Was nun?

## 4. Hands on!

1. Briefbeispiel
2. Synopse Bild-Text

- Eine Technologie-Familie
  - XML
    - TEI
    - XHTML
    - SVG
  - XPath
  - XQuery
  - XPointer
  - Xlink
  - XSL(T)
  - ...

**Extensible Markup Language**

```
<?xml version="1.0"?>
<quiz>
  <frage>
    Wer war der fünfte
    deutsche Bundespräsident?
  </frage>
  <antwort>
    Karl Carstens
  </antwort>
  <!-- Anmerkung: Wir
    brauchen mehr Fragen -->
</quiz>
```

**XML**

Dateiendung: .xml

MIME-Type: application/xml, text/xml  
(deprecated)

Magische Zahl: 3C 3F 78 6D 6C hex  
<?xml

Entwickelt von: [World Wide Web Consortium](#)

Art: [Auszeichnungssprache](#)

Erweitert von: [SGML](#)

Erweitert zu: [XHTML](#), [RSS](#), [Atom](#)

Website: [1.0 \(Fifth Edition\)](#) [1.1 \(Second Edition\)](#)

<http://de.wikipedia.org/wiki/XML>

## XML...

- ...strukturiert Daten, die man in Form von Zeichen festhalten kann
- ...ist erweiterbar
- ...muss wohlgeformt sein
- ...kann validiert werden
- ...ist unabhängig von Computerprogrammen, Herstellern, Betriebssystemen
- ...ist frei verfügbar!**

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und
  legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

# XSL

eXtensible Stylesheet Language

XSL-FO

XSLT

(XPath)



- Dient der Adressierung von Informationen in XML-Dokumenten
  - Bietet Funktionen, zu deren Weiterverarbeitung
  - Beherrscht math. Grundfunktionen
- 
- (kann auch RegEx: match, replace, tokenize, ...)

## /sp

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

# XPath: Eine Übung

/\*

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

# XPath: Eine Übung

/\*

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp/speaker

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp/speaker

<sp>

<speaker>FAUST.</speaker>

<lg>

<!-- [...] -->

<l>Zum Augenblicke dürft' ich sagen:</l>

<l>Verweile doch, du bist so schön!</l>

<l>Es kann die Spur von meinen Erdetagen</l>

<l>Nicht in Äonen untergehn. -</l>

<l>Im Vorgefühl von solchem hohen Glück</l>

<l>Genieß' ich jetzt den höchsten Augenblick.</l>

</lg>

<lb />

<stage>

<hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>

</stage>

<lb />

<pb n="348" />

</sp>

## /sp/lg

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp/lg

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp/lg/1

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp/lg/l

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp/lg/l[3]

<sp>

```
<speaker>FAUST.</speaker>
<lg>
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>
  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>
</stage>
<lb />
<pb n="348" />
```

</sp>

## /sp/lg/l[3]

<sp>

<speaker>FAUST.</speaker>

<lg>

<!-- [...] -->

<l>Zum Augenblicke dürft' ich sagen:</l>

<l>Verweile doch, du bist so schön!</l>

<l>Es kann die Spur von meinen Erdetagen</l>

<l>Nicht in Äonen untergehn. -</l>

<l>Im Vorgefühl von solchem hohen Glück</l>

<l>Genieß' ich jetzt den höchsten Augenblick.</l>

</lg>

<lb />

<stage>

<hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>

</stage>

<lb />

<pb n="348" />

</sp>

## //1[3]

<sp>

<speaker>FAUST.</speaker>

<lg>

<!-- [...] -->

<l>Zum Augenblicke dürft' ich sagen:</l>

<l>Verweile doch, du bist so schön!</l>

<l>Es kann die Spur von meinen Erdetagen</l>

<l>Nicht in Äonen untergehn. -</l>

<l>Im Vorgefühl von solchem hohen Glück</l>

<l>Genieß' ich jetzt den höchsten Augenblick.</l>

</lg>

<lb />

<stage>

<hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den Boden.</hi>

</stage>

<lb />

<pb n="348" />

</sp>

## //1[3]/substring-after(. , 'meinen ')

<sp>

```
<speaker>FAUST.</speaker>
```

```
<lg>
```

```
<!-- [...] -->
```

```
<l>Zum Augenblicke dürft' ich sagen:</l>
```

```
<l>Verweile doch, du bist so schön!</l>
```

```
<l>Es kann die Spur von meinen Erdetagen</l>
```

```
<l>Nicht in Äonen untergehn. -</l>
```

```
<l>Im Vorgefühl von solchem hohen Glück</l>
```

```
<l>Genieß' ich jetzt den höchsten Augenblick.</l>
```

```
</lg>
```

```
<lb />
```

```
<stage>
```

```
<hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den  
Boden.</hi>
```

```
</stage>
```

```
<lb />
```

```
<pb n="348" />
```

</sp>

//1[3]/contains(., 'Erdetagen')  
/following::pb/number(@n) - 1

```
<!-- [...] -->
  <l>Zum Augenblicke dürft' ich sagen:</l>
  <l>Verweile doch, du bist so schön!</l>

  <l>Es kann die Spur von meinen Erdetagen</l>
  <l>Nicht in Äonen untergehn. -</l>
  <l>Im Vorgefühl von solchem hohen Glück</l>
  <l>Genieß' ich jetzt den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust sinkt zurück, die Lemuren fassen ihn auf und legen ihn auf den
Boden.</hi>
</stage>
<lb />

<pb n="348" />
```

//1[3]/contains(., 'Erdetagen')  
/following::pb/number(@n) - 1

```
<!-- [...] -->
<l>Zum Augenblicke dürft' ich sagen:</l>
<l>Verweile doch, du bist so schön!</l>

<l>Es kann die Spur von einem <b>Erde</b>tagen</l>
<l>Nicht in Äonen und Eonen. </l>
<l>Im Vorgefühl von einem hohen <b>Erde</b>k</l>
<l>Genieß' ich den höchsten Augenblick.</l>
</lg>
<lb />
<stage>
  <hi>Faust</hi> zurück, die Leuchten fassen ihn auf und legen ihn auf den
Boden.</hi>
</stage>
<lb />

<pb n="348" />
```

# XPath: Eine Übung

<sp>

```
<speaker>FAUST.</speaker>
```

```
<lg>
```

```
<!-- [...] -->
```

```
<l>Zum Augenblicke dürft' ich sagen:</l>
```

```
<l>Verweile doch, du bist so schön!</l>
```

```
<l>Es kann die Spur von meinen Erdentage</l>
```

```
<l>Nicht in Äonen untertehr -</l>
```

```
<l>Im Vorgefühl von solchen hohen Glück</l>
```

```
<l>Genieß' ich jetzt den höchsten Augenblick.</l>
```

```
</lg>
```

```
<lb />
```

```
<stage>
```

```
<hi>Faust sinkt zurück, die Lemuren fassen ihn auf und  
legen ihn auf den Boden.</hi>
```

```
</stage>
```

```
<lb />
```

```
<pb n="348" />
```

</sp>

**Wer spricht hier?**

# XPath: Eine Übung

<sp>

`<speaker>FAUST.</speaker>`

</sp>

`/sp/speaker`

`<speaker>FAUST.</speaker>`

`/sp/speaker/text()`

FAUST.

`/sp/speaker/substring-before(., '.')`

FAUST

`/sp/speaker/lower-case(substring-before(., '.'))`

faust

`/sp/speaker/concat(substring(., 1, 1),  
lower-case(substring-after(substring-  
before(., '.'), substring(., 1, 1))))`

Faust

*Wer spricht hier?*

<sp>

<speaker>FAUST.</speaker>

<lg>

<!-- [...] -->

<l>Zum Augenblicke dürft' ich sagen:</l>

<l>Verweile doch, du bist so schön!</l>

<l>Es kann die Spun von meinen Endetagen</l>

```
//l[contains(.,  
'Verweile doch, du bist so schön!'  
)]/preceding::speaker/concat(subst  
ring(., 1, 1), lower-  
case(substring-after(substring-  
before(. , '.'), substring(., 1,  
1))))
```

Faust

k.</l>

sen ihn auf und

</sp>



?



- XSLT schreibt man in XML
- Man benötigt ein Programm zum Transformieren = XSLT-Prozessor

„With XSLT you can **add/remove** elements and attributes to or from the output file. You can also **rearrange and sort** elements, perform **tests** and make decisions about which elements to hide and display, **and a lot more**.

XSLT uses XPath to find information in an XML document.“

([http://www.w3schools.com/xml/xml\\_xsl.asp](http://www.w3schools.com/xml/xml_xsl.asp), Hervorhebung von mir)

# XSLT in XML

```
<?xml version="1.0" encoding="UTF-8"?>  
<xsl:stylesheet version="2.0"  
    xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
```

```
</xsl:stylesheet>
```

```
<xsl:stylesheet version="2.0"  
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
```

```
  <xsl:template match="speaker">  
    <xsl:value-of select="."/>  
    <xsl:text>  
    </xsl:text>  
  </xsl:template>
```

```
  <xsl:template match="pb">  
    <xsl:value-of select="number(@n) -1"/>  
  </xsl:template>
```

```
  <xsl:template match="/">  
    <xsl:apply-templates/>  
  </xsl:template>  
  <xsl:template match="text()"/>  
</xsl:stylesheet>
```



- XSLT-Prozessor geht das Input-Dokument schrittweise durch
- Findet er ein Element, sucht er nach Anweisungen
  - Anweisung vorhanden
  - Keine Anweisung? Weitergehen...
- Trifft er ohne Anweisungen auf Text, wird dieser übernommen
- Reihenfolge der Anweisungen ist irrelevant
- Nur wohlgeformte Dokumente werden bearbeitet

- Processing Instruction noch vor dem Wurzelement
- Allgemein:

```
<?xml-stylesheet type="text/xsl" href="PFAD_ZUM_XSLT"?>
```

- Im TextGrid Laboratory:

```
<?xml-stylesheet type="text/xsl" href="textgrid:22dg1.0"?>
```

# Agenda

## 1. Was können wir mit XSLT alles machen?

### 1. Beispiele:

1. Webseiten
2. Textprozessierung
3. Datenbankabfragen
4. ...

## 2. Theorie

1. XML
2. XPath + Übung
3. XSLT
4. Wie verhalten sich beide zueinander?

## 3. XSLT im TextGrid Lab

1. *processing instruction*
2. Transformieren
3. Was nun?

## 4. Hands on!

1. Briefbeispiel
2. Synopse Bild-Text

- XSLT 2.0 – Prozessor (Saxon 9.5)
- TextGrid – XML-Editor
- Vorschauansicht
  - ALT+Bild↓+Bild↓
- Oxygen: „Play-Button“ 
  - Shift+ALT+T T

# XSLT im Lab

The screenshot displays the TextGridLab application window. The main editor shows an XSLT script for transforming an XML document. The script includes namespace declarations, a stylesheet reference, and a series of conditional checks for various XML elements and attributes. The transformation logic is implemented using XSLT 1.0 constructs like `<xsl:if>`, `<xsl:choose>`, and `<xsl:element>`.

On the left side, the 'Notizbuch C07' panel is visible, showing metadata fields such as 'Title(s)', 'Format', and 'Rights Holder(s)'. The 'Part of Edition(s)' section lists 'Notizbuch C07 (textgrid:164c9.0)'. Buttons for 'Save Metadata', 'Reload Metadata', and 'Copy TEI Header' are also present.

On the right side, the 'Gliederung' (Table of Contents) panel shows a hierarchical tree structure of the XML document, including elements like 'teiHeader', 'fileDesc', 'publicationStmt', 'sourceDesc', 'list', 'listPerson', 'listPlace', 'listOrg', 'msDesc', 'encodingDesc', 'profileDesc', and 'list'. The 'Results' panel is currently empty.

The status bar at the bottom indicates the current view is 'Entwurf' (Design) and shows the path 'TEI/teiHeader/fileDesc/sourceDesc/listPerson'. The user 'Mathias.Göbel@textgrid.de' is logged in.

# Hands on!



- Tutorials:
  - <http://www.w3schools.com/xsl/>
  - <http://de.selfhtml.org/xml/darstellung/index.htm>
- W3C-Standard:
  - <http://www.w3.org/standards/xml/transformation>
- Book:
  - Yes, there are many introductions to XSLT available in print.

?

