

TEI Modul 2: core mit Übungen

Rückgreifend auf Folien von S. Dumont

Markus Schnöpf



Leopoldina
Nationale Akademie
der Wissenschaften



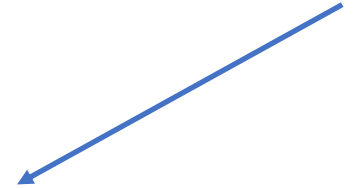
Wo stehen wir?

- Wir befinden uns hier:

<https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

Text Body

- + 1 [The TEI Infrastructure](#)
- + 2 [The TEI Header](#)
- + 3 [Elements Available in All TEI Documents](#)
- + 4 [Default Text Structure](#)
- + 5 [Characters, Glyphs, and Writing Modes](#)
- + 6 [Verse](#)
- + 7 [Performance Texts](#)
- + 8 [Transcriptions of Speech](#)
- + 9 [Dictionaries](#)
- + 10 [Manuscript Description](#)
- + 11 [Representation of Primary Sources](#)
- + 12 [Critical Apparatus](#)
- + 13 [Names, Dates, People, and Places](#)
- + 14 [Tables, Formulæ, Graphics and Notated Music](#)
- + 15 [Language Corpora](#)
- + 16 [Linking, Segmentation, and Alignment](#)
- + 17 [Simple Analytic Mechanisms](#)
- + 18 [Feature Structures](#)
- + 19 [Graphs, Networks, and Trees](#)
- + 20 [Non-hierarchical Structures](#)
- + 21 [Certainty, Precision, and Responsibility](#)
- + 22 [Documentation Elements](#)
- + 23 [Using the TEI](#)



Zu behandelnde Elemente

- Pb, lb, p, l, choice, corr, del, facs etc.
- Grundhandwerkszeug des digitalen Editors
- Fragen an sich selbst:
 - Zu welchem Zweck möchte ich welche Phänomene auszeichnen?
 - Ist Punctuation und Wortebene wichtig in meinem Editionsprojekt?
 - Müssen Sätze ausgezeichnet werden?
 - Habe ich Gedichte in meinem Korpus?
 - Wie kann ich möglichst sparsam auszeichnen, ohne meine Editionsziele aus den Augen zu verlieren?

Grundelemente TEI-Text

- Absatz: **<p>** (paragraph)
- Folio- und Seitenbeginn: **<pb n="1v" />**
- Zeilenbeginn: **<lb/>**
- Liste:
 <list>
 <item></item>
 </list>
- Tabelle: **<table>**
- Anmerkung, Fußnote etc.: **<note>**
- Zeichnungen, Bild etc.: **<figure>**
- Überschrift: **<head>**
- Semantische Abschnitte: **<div>**
- Element **<hi>** (highlighted)

Attribut: @rendition

- Attribut @rendition
 - @rendition mit (innerhalb der Edition) standardisierten Werten (z.B. „#u“ für Unterstreichung), die in <rendition> dokumentiert sind.
 - **<tagsDecl>**
<!-- define italic style using CSS, selecting it as default for emph and hi elements -->
<rendition xml:id="IT" scheme="css"
selector="emph hi">font-style: italic;</rendition>
</tagsDecl>
 - Meistens kann **<hi>** und **@rendition** kombiniert werden:
 - Kursivierung: „eine **<hi rendition="#IT">höfliche</hi>** Entschuldigung“
 - Hochstellungen
 - Gesperrter Druck
 - Wechsel (von Kurrent) zur Antiquaschrift

Des Autors Textmodifikationen

- **Hinzufügung:** `<add>` (addition)
`<add place="above">demohngeachtet</add>`
- **Streichung:** `` (deletion)
letztere `<del rendition="#s">jetztblös`
- **Ersetzung**
`<subst>`
 `<del rendition="#s">pl`
 `<add place="inline">pulsirendes</add>`
`</subst>`

Des Herausgebers Textmodifikationen

- **Ergänzungen durch den Herausgeber:**
`<supplied>`fehlendes Wort`</supplied>`
- **Ausschreibungen durch den Herausgeber:**
 - `<ex>` (editorial expansion)
Sch`<ex>`illing`</ex>`
- oder als `<choice>` kodieren:
`<choice>`
 `<abbr>`a.`</abbr>`
 `<expan>`an`</expan>`
`</choice>`
`<choice>`
 `<abbr>`d.`</abbr>`
 `<expan>`der`</expan>`
`</choice>`

Wahlfreiheiten

- **Auflösung von Abkürzungen:**
`<choice>`
 `<abbr>Sch</abbr>`
 `<expn>Schilling</expn>`
`</choice>`
- **Normalisierungen von Schreibweisen:**
`<choice>`
 `<orig>Etat</orig>`
 `<reg>État</reg>`
`</choice>`
- **Korrekturen durch den Herausgeber:**
`<choice>`
 `<sic>I</sic>`
 `<corr cert="high">The</corr>`
`</choice>`

Weitere Textphänomene

- **Unsichere Lesung:** **<unclear>**
Enthält ein Wort, Satz oder Textstelle, die nicht mit Sicherheit transkribiert werden kann, weil sie schwer lesbar ist
- **Lücke:** **<gap>**
Gibt eine Lücke wieder, die entweder im Manuskript so vorhanden ist (z.B. schadhafte Stelle) oder bei der Transkription durch den Herausgeber entsteht, weil er Text weglässt (z.B. weil dieser für die Edition nicht relevant ist)
- **Schadhafte Stelle:** **<damage>**
Schaden am Manuskript

Streichungen

- Herausgeber streicht überflüssiges Wort
- Option 1:
 `<surplus>` mit `</surplus>`
 → Vorteil: kurz und präzise
- Option 2:
 `<choice>`
 `<sic>` mit `</sic>`
 `<corr type="deleted"/>`
 `</choice>`
 → Vorteil: kein neues Element notwendig, Tagging zum Korrigieren kann hier “nachgenutzt” werden
- Beide Optionen sind in der TEI möglich und valide

Stellenkommentare

- Für Anmerkungen aller Art, auch Stellenkommentare durch den Herausgeber: **<note>**
- Prinzipiell werden Anmerkungen in TEI-XML an der Stelle in den Fließtext eingefügt, an dem sie tatsächlich (z.B. per Einweisungszeichen) oder vom Editor vermutet eingefügt werden
- In der Darstellung der digitalen Edition kann der Stellenkommentar dann z.B. am Rand des Textes angezeigt werden oder als tooltip.

- Häufig ist es erwünscht, die Textstelle, auf den sich ein Kommentar bezieht, anzugeben, weil:
 - dadurch der Kommentar eindeutig einer Textstelle zugeordnet wird
 - man im Web ein oder mehrere Wörter bzw. ganze Sätze als Linkfläche hat, falls die Stellenkommentare nur bei Bedarf eingeblendet werden sollen.
 - Man eine Druckausgabe mit LaTeX plant
- Problem: Um Bezugstext auszuzeichnen, gibt es in der TEI keine empfohlene Methode

Stellenkommentare mit Bezugstext

- Option 1:

```
<ref xml:id="xyz" corresp="abc">[Bezugstext]</ref>  
<note xml:id="abc" corresp="xyz">[Stellenkommentar]</note>
```

- Option 2:

```
<app>  
  <lem>[Bezugstext]</lem>  
  <note>[Stellenkommentar]</note>  
</app>
```

- Option 3:

```
<anchor xml:id="xyz"/>[Bezugstext]<note  
target="#xyz">[Stellenkommentar]</note>
```

Bilder - Faksimiles

- Faksimiles

```
<text>  
  <pb facs="page1.png"/>  
  <!-- text contained on page 1 is encoded here -->  
  <pb facs="page2.png"/>  
  <!-- text contained on page 2 is encoded here -->  
</text>
```

- Reine Faksimile-Edition

```
<facsimile>  
  <graphic url="page1.png"/>  
  <graphic url="page2.png"/>  
  <graphic url="page3.png"/>  
  <graphic url="page4.png"/>  
</facsimile>
```

Bilder –Illustrationen und Diagramme

- <figure> gruppiert Elemente, die grafische Informationen (Illustrationen, Formeln, Figuren) beinhalten oder repräsentieren
- <media> zeigt den Ort externer Medien wie Audio, Video etc..
- <graphic> zeigt den Ort einer Illustration oder Abbildung, die Teil des Textes ist, oder ein Bild dessen zur Verfügung stellt.