

# Scholarly editing with TEI-XML

Brown Bag meeting at MPIWG

Klaus Thoden

April 12, 2016

# Introduction to XML

- ▶ Markup language, adding metainformation to a text
  - ▶ To enrich it
  - ▶ To make information explicit, in a meaningful and reusable way
- ▶ Similar to HTML, but stricter
- ▶ XML is in a lot of places, even in Microsoft Word documents

## Basic syntax: just a few rules

- ▶ Angled brackets
- ▶ Tree-like structure
- ▶ Case sensitivity
- ▶ Example

```
<text> <p type="foo">An example</p> </text>
```

# The Text Encoding Initiative

- ▶ Started in 1987
- ▶ Encoding standard, institution, community
- ▶ Benefits: data exchange, interoperability, preservation
- ▶ Guidelines for nearly every aspect you might want to encode:  
around 500 tags
- ▶ <http://www.tei-c.org/Guidelines/P5/>

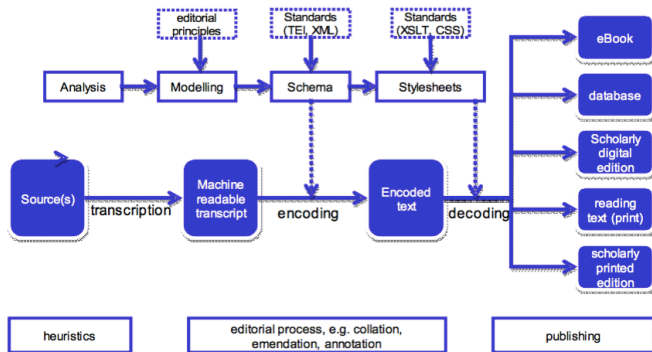
# XML Editors

- ▶ XML-aware editors
- ▶ Might even contain built-in support for TEI: select schemas, validation support, rendering
- ▶ TextGrid, oXygen, nxml-mode in Emacs, jEdit and TextWrangler to a lesser extent

# A model of the text

- ▶ TEI is big, and user decides what she wants to encode
- ▶ A model of the text is created
- ▶ The goal is the markup, not an 'edition'

# Editorial work



(Quelle: Rehbein/Fritze 2012)

# Autopsy of a TEI file

- ▶ File
  - ▶ Head
    - ▶ Contains metadata (author, title, date)
    - ▶ Provenance of source
    - ▶ Documentation of work on the source
  - ▶ Body
    - ▶ Text structure (paragraphs, sections, images)
    - ▶ Citations, emphases
    - ▶ corrections, normalizations
    - ▶ Dates, Names, Places
    - ▶ Annotations, editors' comments



# Auxiliaries

- ▶ Schema files (the model, grammar)
- ▶ Formatters (the processors)
  - ▶ CSS: set of rules to display XML code in the browser – but doesn't alter the structure
  - ▶ XSL: XML-aware programming language to transform XML code into other formats or change the structure (add, modify)

## Useful links

- ▶ Tutorial on editing digital texts with TEI (in German):  
<https://de.dariah.eu/tei-tutorial>
- ▶ Introduction to XML: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/SG.html>
- ▶ Schema customization: <http://www.tei-c.org/Roma/>
- ▶ Document conversion: <http://www.tei-c.org/oxgarage/>
- ▶ TextGrid: <http://www.textgrid.de>
- ▶ Fun stuff: dramavis + Play(s)