TEI & XML

An Introduction

What is TEI?

- TEI: Text Encoding Initiative

- Guidelines for information about storage, creation, transformation etc.

- Information type: texts, audios, pictures, videos.

- Encoding language: **E**xtensible **M**arkup **L**anguage (XML)

  - See https://tei-c.org for more information about TEI & XML

TEI Structure

- TEI file: header & text (body) →

- mother nodes + child nodes
  \<OpenTag\> element \</CloseTag\>

  \<TEI\>

  \<teiHeader\>\</teiHeader\>

  \<text\>\</text\>

  \</TEI\>

```xml
<TEI version="3.3.0" xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
        <fileDesc>
            <titleStmt>
                <title>The shortest TEI Document Imaginable</title>
            </titleStmt>
            <publicationStmt>
                <p>First published as part of TEI P2, this is the P5
                    version using a namespace.</p>
            </publicationStmt>
            <sourceDesc>
                <No source: this is an original work.</p>
            </sourceDesc>
        </fileDesc>
    </teiHeader>
    <text>
        <body>
            <p> This is about the shortest TEI document imaginable.</p>
        </body>
    </text>
</TEI>
```

Header

➜ includes metadata

mandatory elements:

1. <fileDesc> (file description): bibliographic description of the electronic text
2. <encodingDesc> (encoding description): description of the relation of the electronic text to its source
3. <profileDesc> (profile description): description of the context in which the electronic text was created + classification of information

\<fileDesc>

➔ the file description is probably the most important part of the header

➔ it includes:

- ◆ \<titleStmt> (title statement)

- ◆ \<publicationStmt> (publication statement)

- ◆ \<sourceDesc) (source description)

- ◆ \<encodingDesc> (encoding description)

- ◆ \<profileDesc> (profile description)

<titelStmt>

➔ first mandatory subsection of the <fileDesc>

➔ must at least contain:

◆ <title>: title for the electronic text

◆ <author>, <editor>, <sponsor>, <funder>, <principal>: the parties that have been involved in the realisation of the electronic text (not necessarily all)

◆ <respStmt> (responsibility statement): to encode responsibilities such as <name> of a party + their <resp> in a list

\<pubStmt\>

➔   publisher, status, publication place, address, availability, date

our publication statement:

```
<publicationStmt>
<publisher> Centre for Translation Studies Düsseldorf  </publisher>
<status>not yet published </status>
<pubPlace> Düsseldorf </pubPlace>
<address> <name type="institution" > Centre for Translation Studies Düsseldorf  </name>
<name type="institution" > Heinrich-Heine-Universität Düsseldorf  </name>
<street> Universitätsstraße 1  </street>
<postCode> 40225 </postCode>
<name type="city"> Düsseldorf </name>
<name type="country"> Germany </name>
</address>
<availability>  not yet available  </availability>
<date> TBD </date>
</publicationStmt>
```

\<sourceDesc\>

➔ the BIBLIOGRAPHIC source, not the actual one

➔ in our case, title, storyteller + translator

```
<sourceDesc>
<bibl> <title></title> <storyteller></storyteller> Translated by <editor
role="translator"> Tasun Tidorchibe </editor> </bibl>
</sourceDesc>
```

\<encodingDesc\>

➔ second major element of the \<fileDesc\>

➔ may contain:

◆ \<editorialDecl\> (editorial practice declaration)

◆ \<projectDesc\> (project description)

◆ \<tagsDecl\> (tagging declaration) → may contain additional renditions (like font + colour)

◆ \<refsDecl\> (reference system declaration)

◆ \<classDecl\> (classification declaration)

\<profileDesc\>

➜ third major element of the \<fileDesc\>

➜ used to document non-bibliographic information

◆ e.g. information about the creation of the text

◆ languages used in the text

◆ classification of the text

→ information about the actual source

\<profileDesc\> - example

```
<profileDesc>
        <creation> Originally told in
        <place xml:id="Sobii, Saboba">
                <placeName> Sobii, Saboba</placeName>
                <location>
                        <geo>9.71, 0.32</geo>
                </location>
        </place>
        transcribed folktale from Likpakpaln as told by <name type="storyteller">Binibaaliin </name> on <date
when="2022">22nd March 2022</date> in <name type="city">Sobii, Saboba</name>, <name type="country">Ghana</name>
        </creation>
        <langUsage>
                <language ident="lik">Likpakpaln</language>
                <language ident="en">English</language>
        </langUsage>
        <textClass>placeholder tbd</textClass>
</profileDesc>
```

<body>: basic elements

- <div>: text divisions

- <head>: title + subtitle

- <p>: paragraph

- <q>: spoken words / quotes

- <l>: lines in poems / songs

\<div>

- Short for division

- Texts can be divided into chapters, sections etc.

- When a new section starts, use the division tag

- This tag will be closed when the division ends

  - \<div type="introduction">

Foreignisation vs. Domestication

Foreignised Text

Left with no other option, he decided to go alone, carrying with him all the instruments needed to perform the <u>njeen</u> and <u>kinachuŋ</u>: he picked up the <u>ligangaln</u>; he tied <u>tibaln</u> around his ankles; he strapped the <u>kilon</u> to one shoulder; he slung an axe over the other; he picked up his <u>uduun</u> in his hand; he tied all the other <u>tinachinchen</u> around his waist and hit the road running.

Domesticated Text

Left with no other option, he decided to go alone, carrying with him all the instruments needed to perform the <u>procession</u> and <u>traditional funeral dance</u>: he picked up the <u>drum</u>; he tied <u>bangles</u> around his ankles; he strapped an <u>armpit drum</u> to one shoulder; he slung an axe over the other; he picked up his <u>gong</u> in his hand; he tied all the other <u>musical instruments</u> around his waist and hit the road.