



TEI
Basics

Lernziele

- Was ist Textkodierung?
- Was ist TEI?
 - Was könnte es mir bringen?
 - Wie kann ich es nutzen?

Voraussetzungen

- Was ist XML?
 - Strukturelemente?
 - Wohlgeformtheit?
- Was ist XML Schema?
 - Validität?

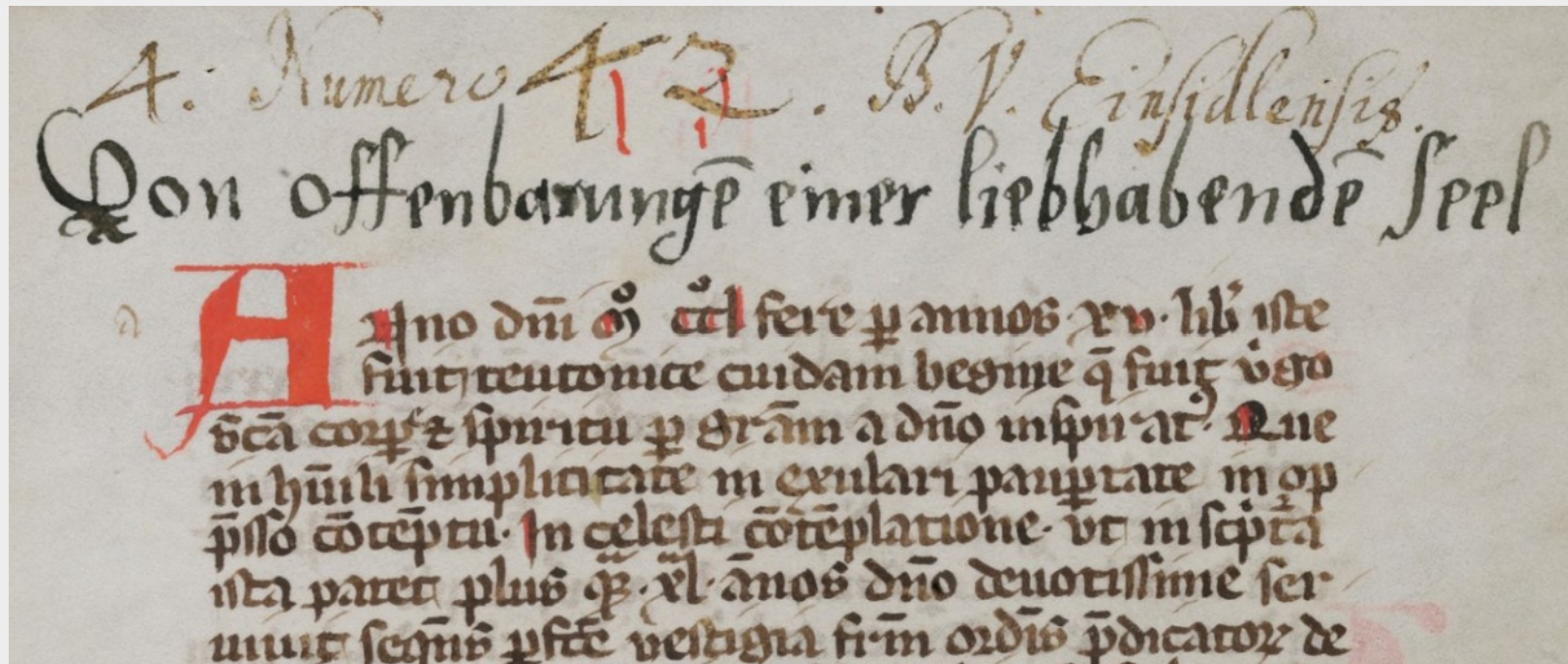


Text
Encoding

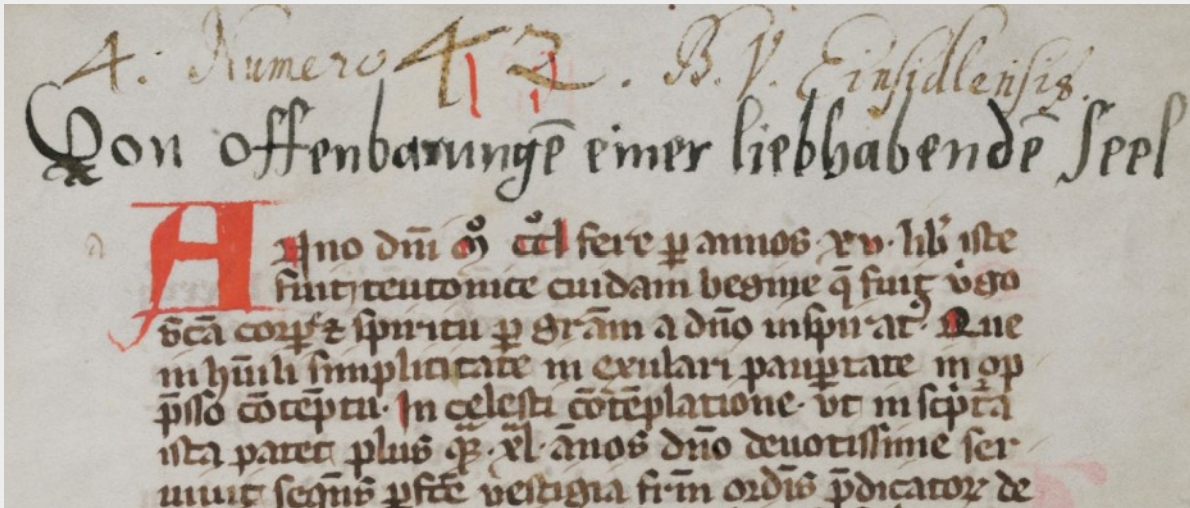
Was ist Text Encoding?

```
<code>  
text text  
text<code/>text  
<code>text text</code>  
text  
</code>
```

Was ist Text Encoding?



Was ist Text Encoding?



Von Offenbarungē[n] einer liebhabende[n] Seel

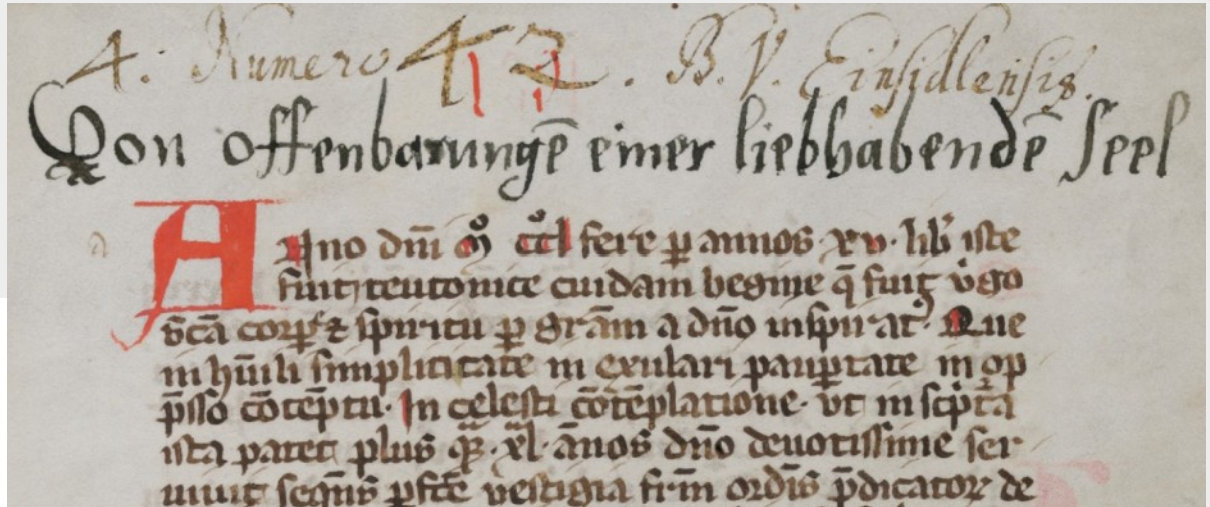
Anno d[o]m[ini] MCCL fere p[er] annos XV lib[er] iste

Was ist Text Encoding?

Von Offenbarunge[n] einer liebhabende[n] Seel

Anno d[o]m[ini] MCCL fere p[er] annos XV lib[er] iste

Was ist Text Encoding?

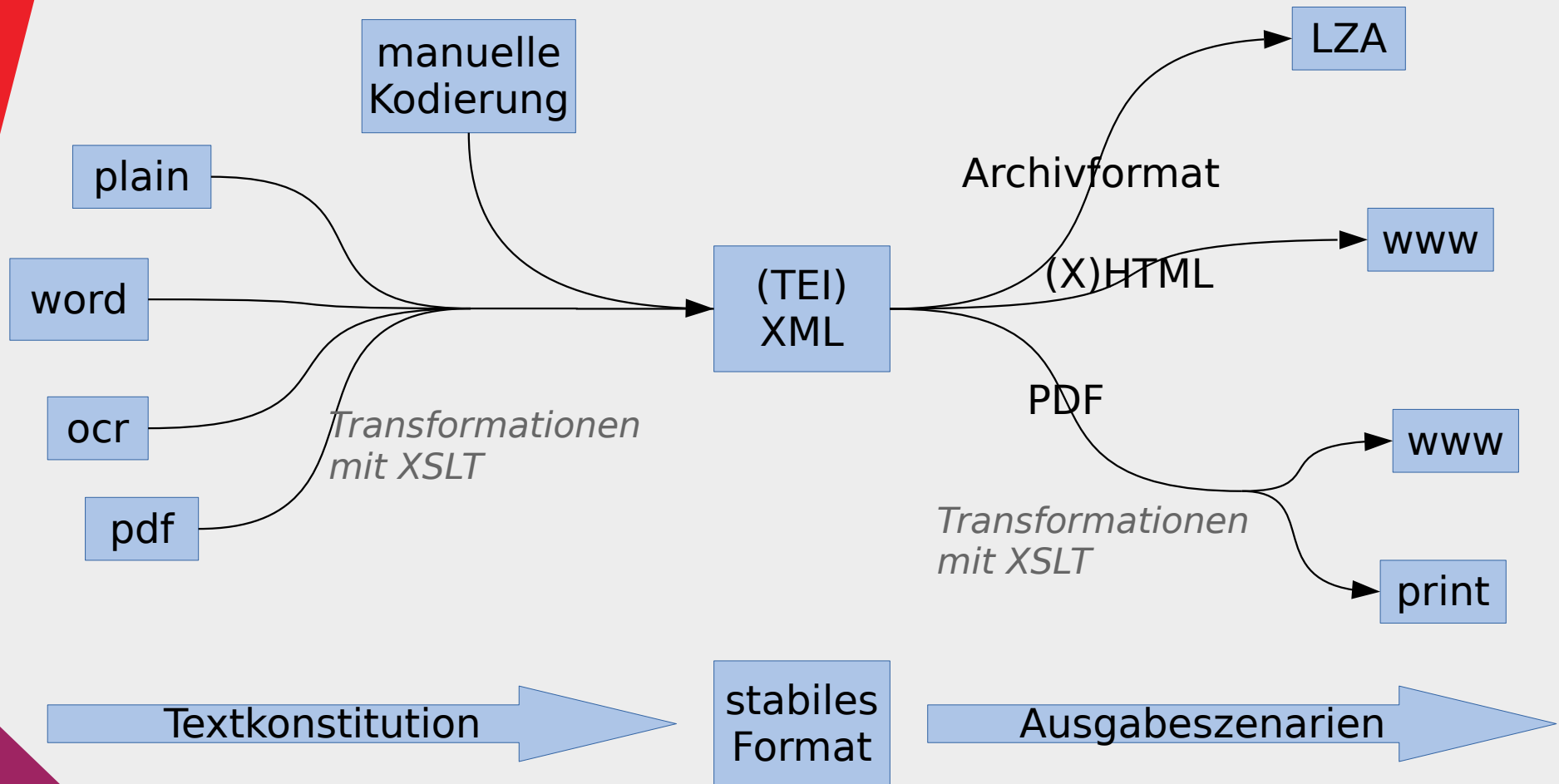


```
<head>Von offenbarung<choice>
  <abbr>e</abbr>
  <expn>en</expn>
</choice> einer liebhabend<choice>
  <abbr>e</abbr>
  <expn>en</expn>
</choice> Seel</head>
<p><figure type="initial"/><date><hi rend="red">A</hi>nno d<choice>
  <abbr>m</abbr>
  <expn>omni</expn>
</choice>
  <num type="roman" value="1250">M<sup>o</sup>CC<sup>o</sup>L</num></date> fere <choice>
  <abbr>p</abbr>
  <expn>per</expn>
</choice> annos <num type="roman" value="15">XV</num>
<choice>
  <abbr>lib'</abbr>
  <expn>liber</expn>
</choice> iste</p>
```

Von Offenbarungē[n] einer liebhabende[n] Seel

Anno d[o]m[ini] MCCL fere p[er] annos XV lib[er] iste

XML-basierter Workflow (Beispiel)



Paradigma des Text Encoding

- WYSIWYM before WYSIWYG
- getrennte Bereiche:
Text - Annotation - Präsentation
- Digitale Edition



XML
Sandbox

Textvorlage

furu ike ya
kawazu tobikomu
mizu no oto
– *Matsuo Bashō*

Der alte Weiher:
Ein Frosch springt hinein.
Oh! Das Geräusch des Wassers.

Mögliche Umsetzung 1

```
<?xml version="1.0" encoding="UTF-8"?>
<haiku>
  <autor>Matsuo Bashō</autor>
  <text xml:lang="JP" xml:id="haiku1">
    <v n="1">furu ike ya</v>
    <v n="2">kawazu tobikomu</v>
    <v n="3">mizu no oto</v>
  </text>
  <text xml:lang="DE" orig="#haiku1">
    <v n="1">Der alte Weiher:</v>
    <v n="2">Ein Frosch springt hinein.</v>
    <v n="3">Oh! Das Geräusch des Wassers.</v>
  </text>
  <!-- Übers. Roland Barthes -->
</haiku>
```

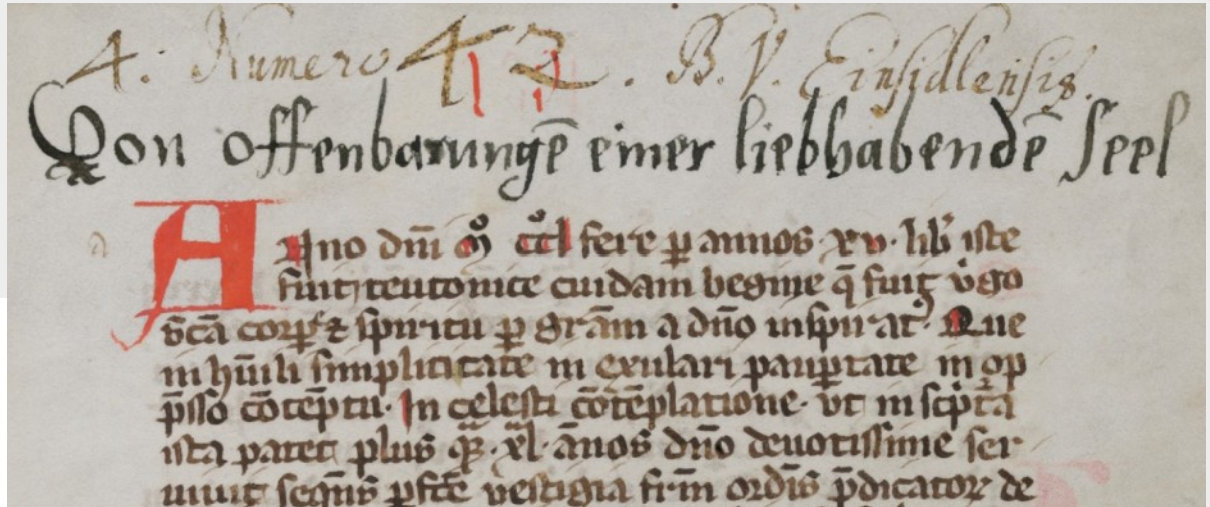
Mögliche Umsetzung 2

```
<?xml version="1.0" encoding="UTF-8"?>
<text>
  <body>
    furu <t ref="#pond">ike<t> ya<eol/>
    kawazu tobikomu<eol/>
    <t ref="#water">mizu<t/> no oto<eol/>
    – <i>Matsuo Bashō</i><eol/>
    <eol/>
    Der alte <t ref="#pond">Weiher<t>:<eol/>
    Ein Frosch springt hinein.<eol/>
    Oh! Das Geräusch des
      <t ref="#water">Wassers</t>.<eol/>
  </body>
  <terms>
    <term xml:id="pond"/>
    <term xml:id="water"/>
  </terms>
</text>
```



Text
Encoding
Initiative

Was will die TEI?

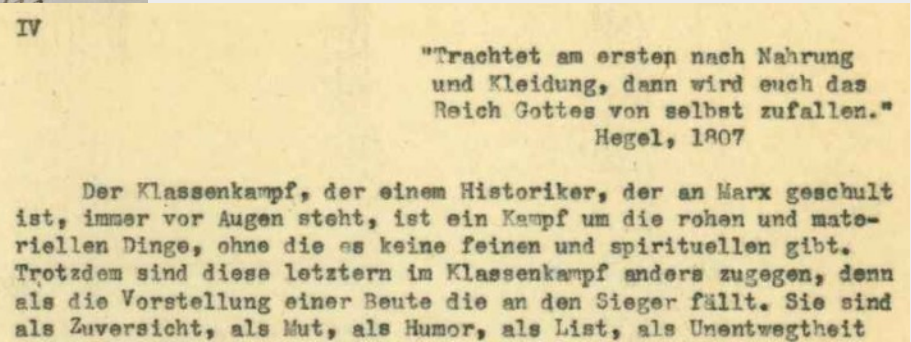
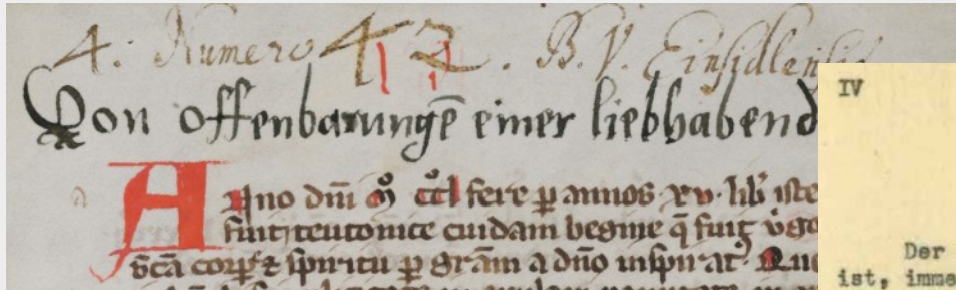


```
<head>Von offenbarung<choice>
  <abbr>e˘</abbr>
  <expn>en</expn>
</choice> einer liebhabend<choice>
  <abbr>e˘</abbr>
  <expn>en</expn>
</choice> Seel</head>
<p><figure type="initial"/><date><hi rend="red">A</hi>nno d<choice>
  <abbr>m˘</abbr>
  <expn>omini</expn>
</choice>
<num type="roman" value="1250">M˚CC˚L˚</num></date> fere <choice>
  <abbr>p</abbr>
  <expn>per</expn>
</choice> annos <num type="roman" value="15">XV</num>
<choice>
  <abbr>lib'</abbr>
  <expn>liber</expn>
</choice> iste</p>
```

Von Offenbarungē[n] einer liebhabende[n] Seel

Anno d[o]m[ini] MCCL fere p[er] annos XV lib[er] iste

Was will die TEI?



der Zeit
 en Herrschen-
 men ihr
 eliotropismus
 , die am Him-
 inbarste
 terialist

Bilder: Mechthild von Magdeburg, Das fließende Licht der Gottheit. Codex Einsidlensis 277, fol. 2r, DOI: 10.5076/e-codices-sbe-0277, CC BY-NC 4.0; Pedanius Dioscorides: Wild Cucumber, URL: https://commons.wikimedia.org/wiki/File:Wild_Cucumber_in_Arabic_Dioscorides.jpg, Public Domain; Walter Benjamin: On the Concept of History, 1940, URL: https://commons.wikimedia.org/wiki/File:04_typoskript.jpg, Public Domain; Corriere della sera, 05.03.1876, URL: https://commons.wikimedia.org/wiki/File:Corriere_5_marzo_1876.jpg, Public Domain.

Text Encoding Initiative

- aktiv seit 1987
- älter als XML (!?)
- international
(= westliche Welt ...)
- seit 2001 als
Konsortium (TEI-C)

Text Encoding Initiative

- Website <https://tei-c.org/>
 - Mitgliedertreffen
 - Zeitschrift „jTEI“
 - Mailingliste „TEI-L“
 - viele Tools
 - Guidelines ♥

Die **TEI** Guidelines

- aktuelles Release: TEI P5
- letzte Version:
3.6.0 (16. Juli 2019)
- GitHub:
github.com/TEIC/TEI
- Zenodo-DOI:
[10.5281/zenodo.3413525](https://doi.org/10.5281/zenodo.3413525)

Die TEI Guidelines

- ca. 2.000 Printseiten
 - modular aufgebaut
 - alles nützt irgendjemandem
 - niemand nutzt alles
- in ständiger Entwicklung
 - durch TEI-C und die SIGs
- „de-facto“-Standard
 - keine Norm, aber oft eine Empfehlung
- selbst anpassbar
 - z.B. an nicht-lateinische Schriftkulturen

Die TEI Guidelines

- unvollkommen
- undogmatisch
- „best practice collection“
- Empfehlung
 - erster Schritt:
schemafreies Modellieren
 - zweiter Schritt:
TEI Guidelines checken,
erst anders machen,
... und dann doch gemäß TEI machen

Haiku in TEI

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Frosch-Haiku</title>
        <author>Matsuo Bashō</author>
      </titleStmt>
      <publicationStmt>
        <p>Winter School 2020</p>
      </publicationStmt>
      <sourceDesc>
        <bibl xml:id="s1">https://de.wikipedia.org/wiki/Haiku</bibl>
        <bibl xml:id="s2">Roland Barthes: Das Reich der Zeichen. Suhrkamp,
          Frankfurt am Main 1981, ISBN 3-518-11077-2.</bibl>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <head><title>Der Frosch</title></head>
      <lg xml:lang="JP" xml:id="t1" source="#s1">
        <l>furu ike ya</l>
        <l><term>kawazu</term> tobikomu</l>
        <l>mizu no oto</l>
      </lg>
      <lg xml:lang="DE" xml:id="t2" corresp="#t1" source="#s2">
        <l>Der alte Weiher:</l>
        <l>Ein <term>Frosch</term> springt hinein.</l>
        <l>Oh! Das Geräusch des Wassers.</l>
      </lg>
    </body>
  </text>
</TEI>
```