Datenmodellierung

Woche 7: Textmodellierung

Institut für Digital Humanities, Historisch-Kulturwissenschaftliche Informationsverarbeitung | Prof. Dr. Øyvind Eide

Regeln

- Alle **Elemente** müssen geöffnet und geschlossen werden

```
<p>Some text here.</p>
```

- **Attribute** werden nur einmal definiert

```
<p n="1">Some text here.</p>
```

- Overlapping ist nicht erlaubt

```
*<l><placeName>Vened</l>ig</placeName>
```

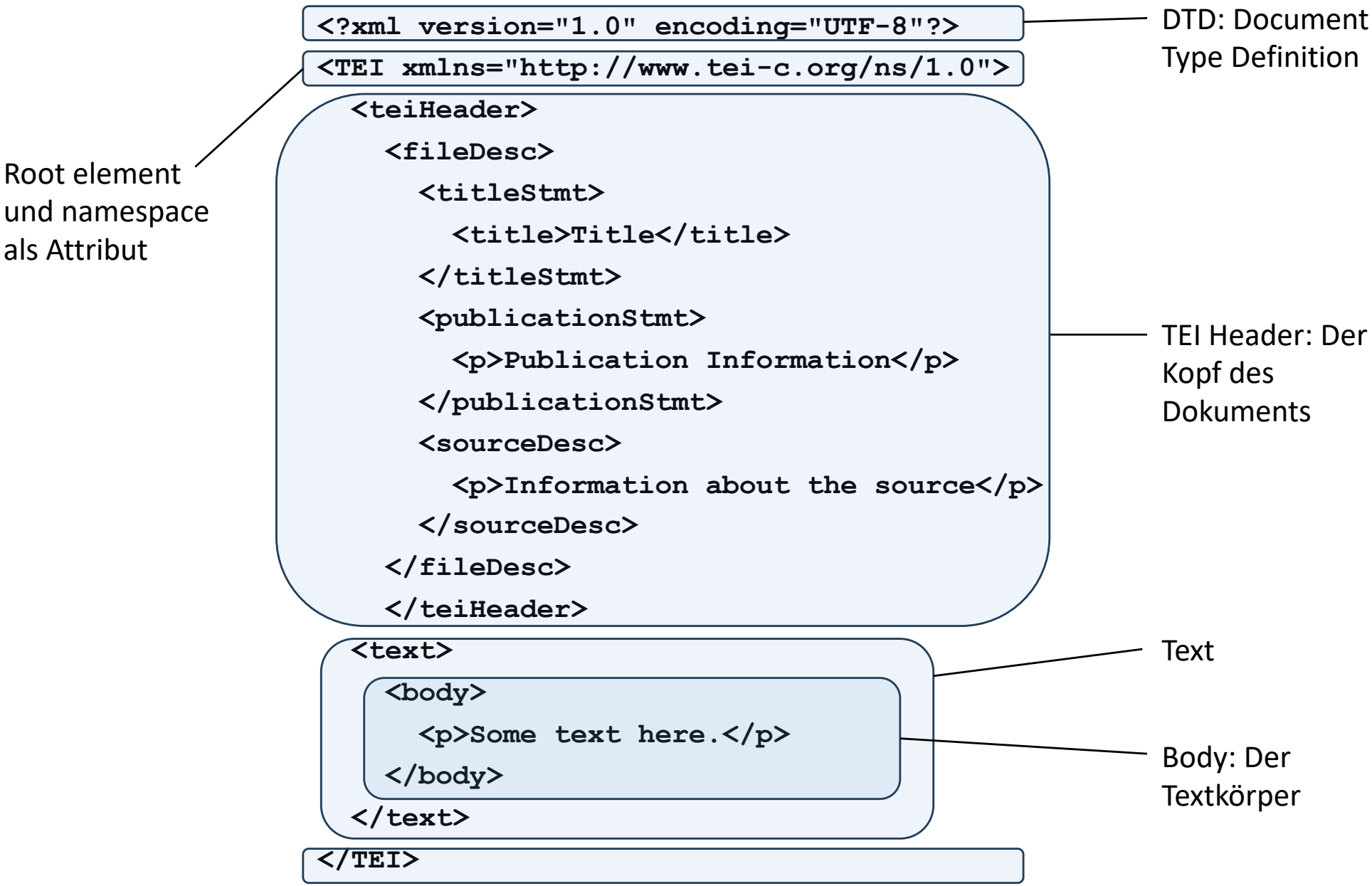
- Self-contained elements lösen viele Probleme:

```
<lb/><placeName>Vened <lb/>ig</placeName> ist schön
```

- Die Struktur ist hierarchisch!

Universität zu Köln

Ein TEI-Dokument



Beispiel

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Exercise</title>
      <respStmt>
        <resp>Digital edition by</resp>
          <name>Elisa Cugliana</name>
        </respStmt>
      </titleStmt>
      <publicationStmt>
        <p>Not intended for publishing</p>
      </publicationStmt>
      <sourceDesc>
        <bibl>
          <author>Rilke, Rainer M.</author>
          <title>Gedichte</title>
          <pubPlace>Frankfurt am Main</pubPlace>
          <publisher>Fischer Taschen Bibliothek</publisher>
          <date>2016</date>
        </bibl>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  ...
```

Beispiel

```
...
  <text>
    <body>
      <div type="poem">
        <head>Spätherbst in <placeName>Venedig</placeName></head>
        <lg type="stanza" n="1">
          <l n="1">Nun treibt die Stadt schon nicht mehr wie ein
                    Köder,</l>
          <l n="2">der alle aufgetauchten Tage fängt.</l>
          <l n="3">Die gläsernen Paläste klingen spröder</l>
          <l n="4">an deinen Blick. Und aus dem Gärten hängt</l>
        </lg>
        <lg type="stanza" n="2">
          <l n="5">der Sommer wie ein Haufen Marionetten</l>
          <l n="6">kopfüber, müde, umgebracht.</l>
          <l n="7">Aber vom Grund aus alten Waldskeletten</l>
          <l n="8">steigt Willen auf: als sollte über Nacht</l>
        </lg>
        <!-- The text goes on -->
      </div>
    </body>
  </text>
</TEI>
```

Philologisches Intermezzo: Mittelalterliche Textualität

- Textträger: Stein, Holz, Metall, Papyrus, Pergament, Papier
  - ➢ Einfluss auf die Texte selbst

ᛖᚲᚺᛚᛖᚹᚨᚷᚨᛊᛏᛁᛉ:ᚺᛟᛚᛏᛁᛃᚨᛉ:ᚺᛟᚱᚾᚨ:ᛏᚨᚹᛁᛞᛟ

ek hlewagastiz : holtijaz : horna : tawido

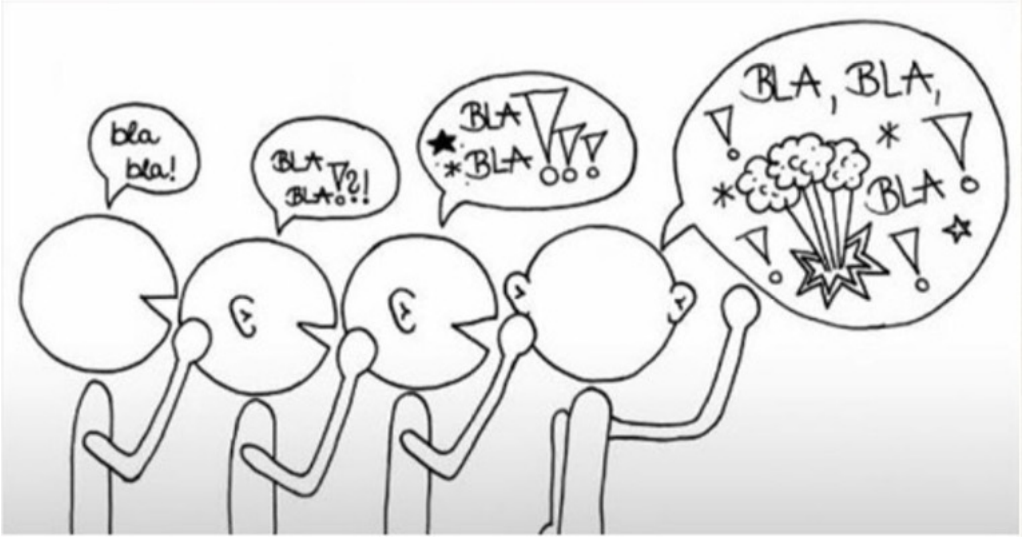
„Ich, Hlewagastiz, Holtijaz, machte das Horn.“

https://samlinger.natmus.dk/do/asset/10765

https://de.m.wikipedia.org/wiki/Goldhörner_von_Gallehus

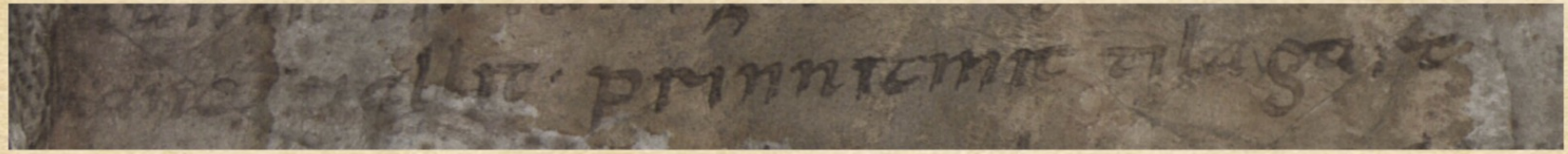
Überlieferung von mittelalterlichen Texten

- Originale Werke: Oft verloren gegangen
- Texte oft mündlich vorgetragen
- Überlieferung: Texte werden manuell abgeschrieben
- Mechanische Fehler + bewusste Änderungen beim Abschreiben
  - auch um die Texte für eine neue Leserschaft anzupassen)
  - ➢ Varianten (Form und Inhalt)
- Textbausteine werden benutzt um andere Texte zu verfassen
- Autor:innen bleiben meistens anonym



http://www.peterhartauer.de/wp-content/uploads/2019/02/Chinese%20Whispers_Peter%20Hartauer_final%20.pdf

Handschriften lesen



mano uallit · prinnitmit tilagart

mano vallit · prinnit mittilagart

Mano fallit, prinnit mittilagart

Der Mond fällt, der Mittelgart brennt

Faksimile

Diplomatisch

Semi-diplomatisch

Interpretative

Übersetzung

Universität zu Köln

Texte normalisieren

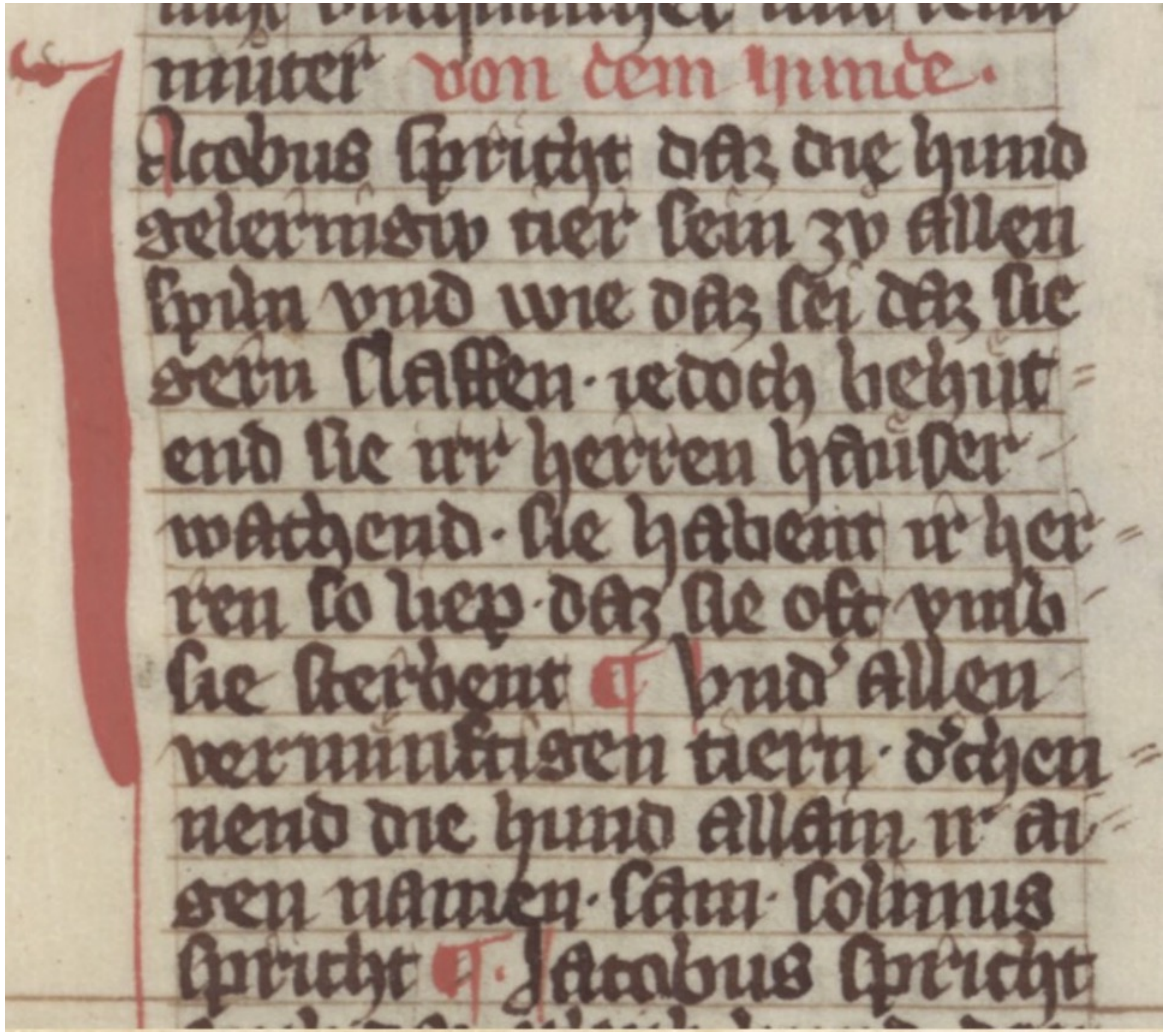
- Im Mittelalter gab es keinen orthographischen Standard
- Man konnte die Wörter so schreiben, wie man es für richtig hielt

  A Linguistic Atlas of Late Mediaeval English records sixteen spellings of the first-person personal pronoun and more than five hundred of the preposition ‚through'. (Cruse 2017)
- Man konnte auch die Worte in scriptio continua (ohne Leerzeichen) schreiben, um Pergament zu sparen
- In wissenschaftlichen Editionen werden die Texte oft normalisiert: die Orthographie wird nach einem gewissen Standard „angepasst"
- Das verbessert Lesbarkeit, Durchsuchbarkeit und Analyse der Texte

Unser Material: Das Buch der Natur
von Konrad von Megenberg (1309-1374)

- Erste Naturgeschichte in deutscher Sprache
  - 106 Textzeugen + 6 Drucke des 15. Jh.s
- Enzyklopädie in acht Büchern
- Hauptquelle ist der "Liber de natura rerum" (1240) des Dominikaners Thomas von Cantimpré
- Ziel des Buches ist die Gotteserkenntnis
- Bei der Bezeichnung der Objekte verwendete Konrad thüringische, ostfränkische und bayerische Wörter

Beispiel der Inhalte

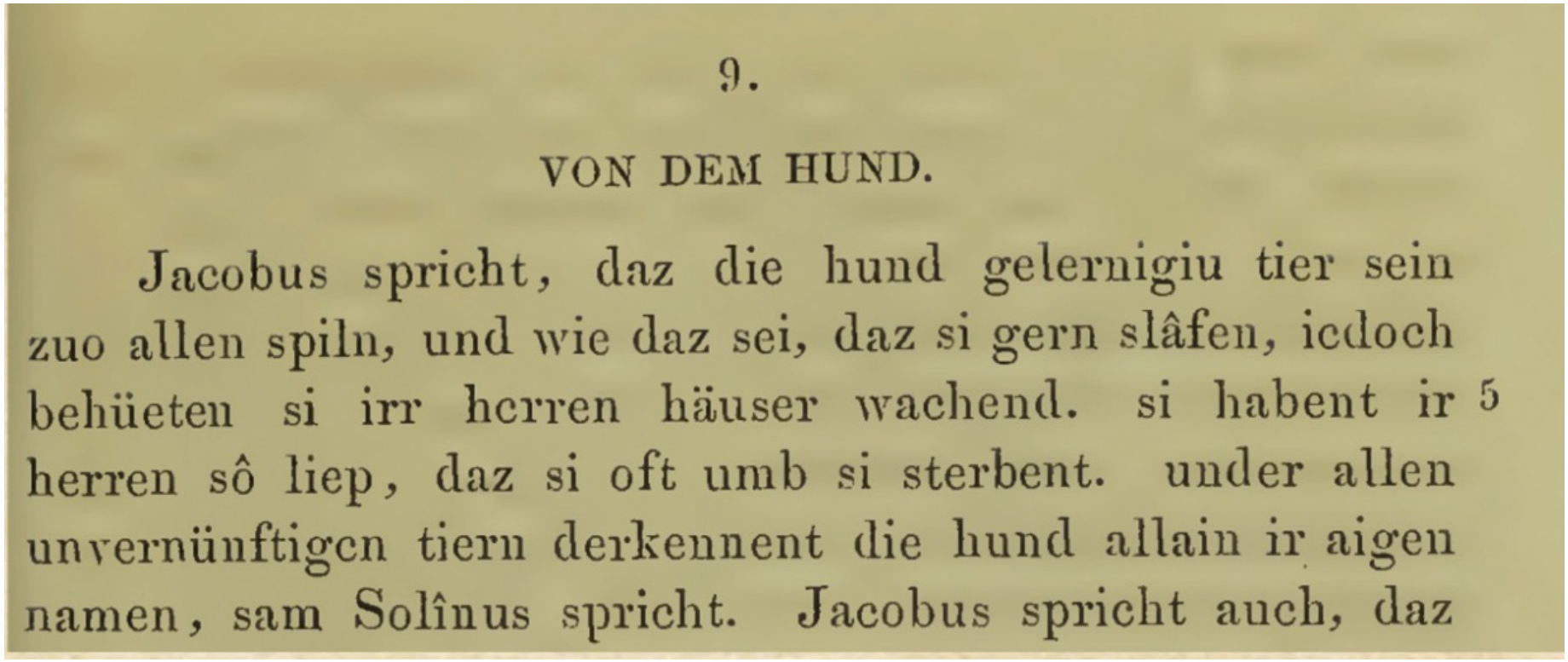


München, BSB Cgm 38

https://www.bavarikon.de/object/bav:BSB-HSS-00000BSB00043227?view=meta

Universität zu Köln

Beispiel der Inhalte



9.

VON DEM HUND.

Jacobus spricht, daz die hund gelernigiu tier sein zuo allen spiln, und wie daz sei, daz si gern slâfen, icdoch behüeten si irr herren häuser wachend. si habent ir herren sô liep, daz si oft umb si sterbent. under allen unvernünftigen tiern derkennent die hund allain ir aigen namen, sam Solînus spricht. Jacobus spricht auch, daz

Pfeiffer (1846)

https://dn790007.ca.archive.org/0/items/b24880681/b24880681.pdf

Beispiel der Inhalte

Universitätsbibliothek Heidelberg, Cod. Pal. germ. 311

https://digi.ub.uni-heidelberg.de/diglit/cpg311/0164/image,info,thumbs

Universität zu Köln

Unser Tool: Oxygen XML Editor

- Hier Demo Version downloaden:

Mac:
https://www.oxygenxml.com/xml_editor/download_oxygenxml_editor.html?os=macOS

Windows:
https://www.oxygenxml.com/xml_editor/download_oxygenxml_editor.html?os=Windows

Linux:
https://www.oxygenxml.com/xml_editor/download_oxygenxml_editor.html?os=Linux

- Musterdatei buchdernatur.xml (Ilias) in Oxygen öffnen

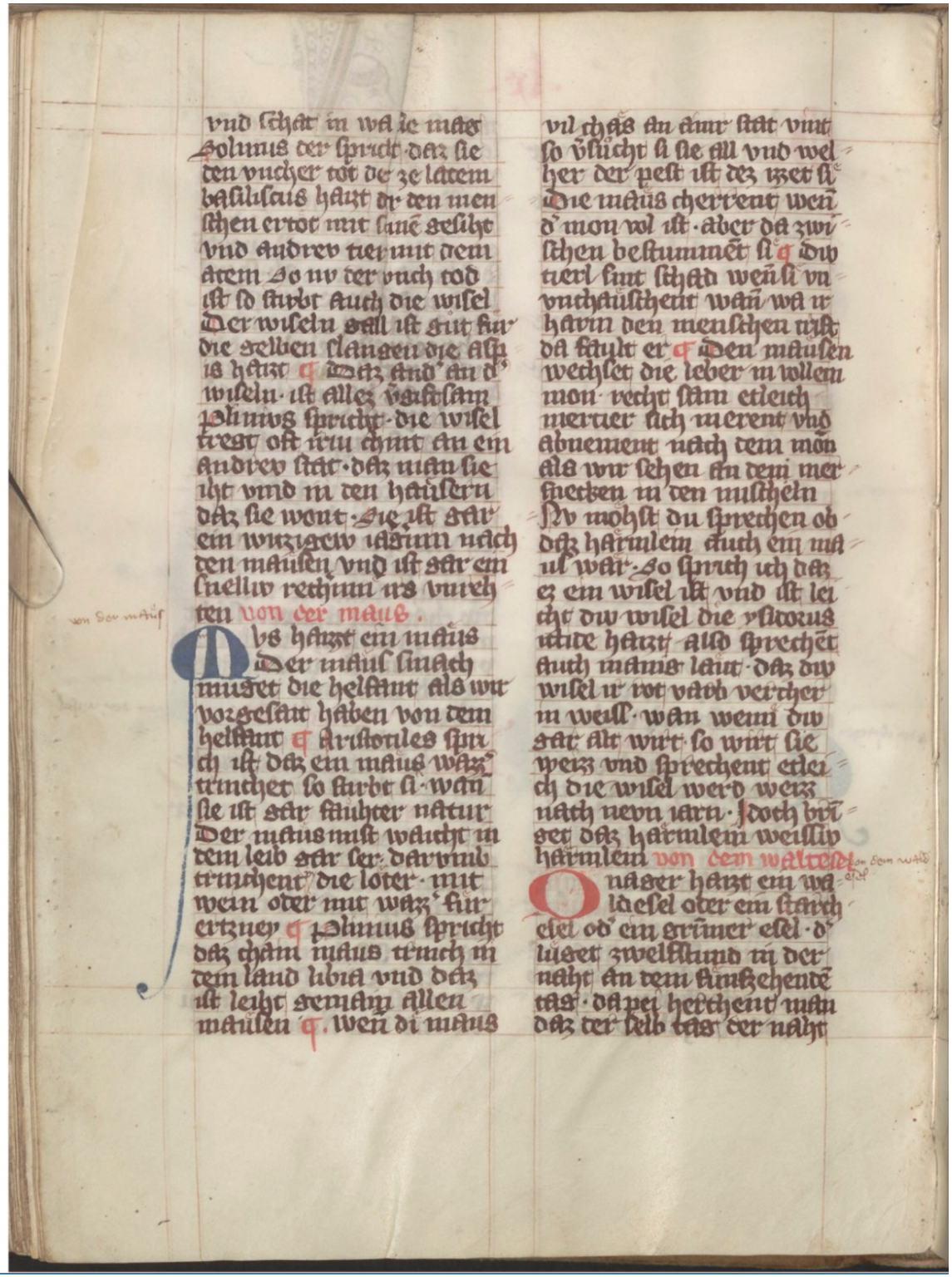
https://www.ilias.uni-koeln.de/ilias/goto.php/crs/6743345

Von der Maus

- Welche Aspekte fallen Ihnen auf?
- Was würden Sie annotieren, wenn Sie könnten?
- Zu zweit:

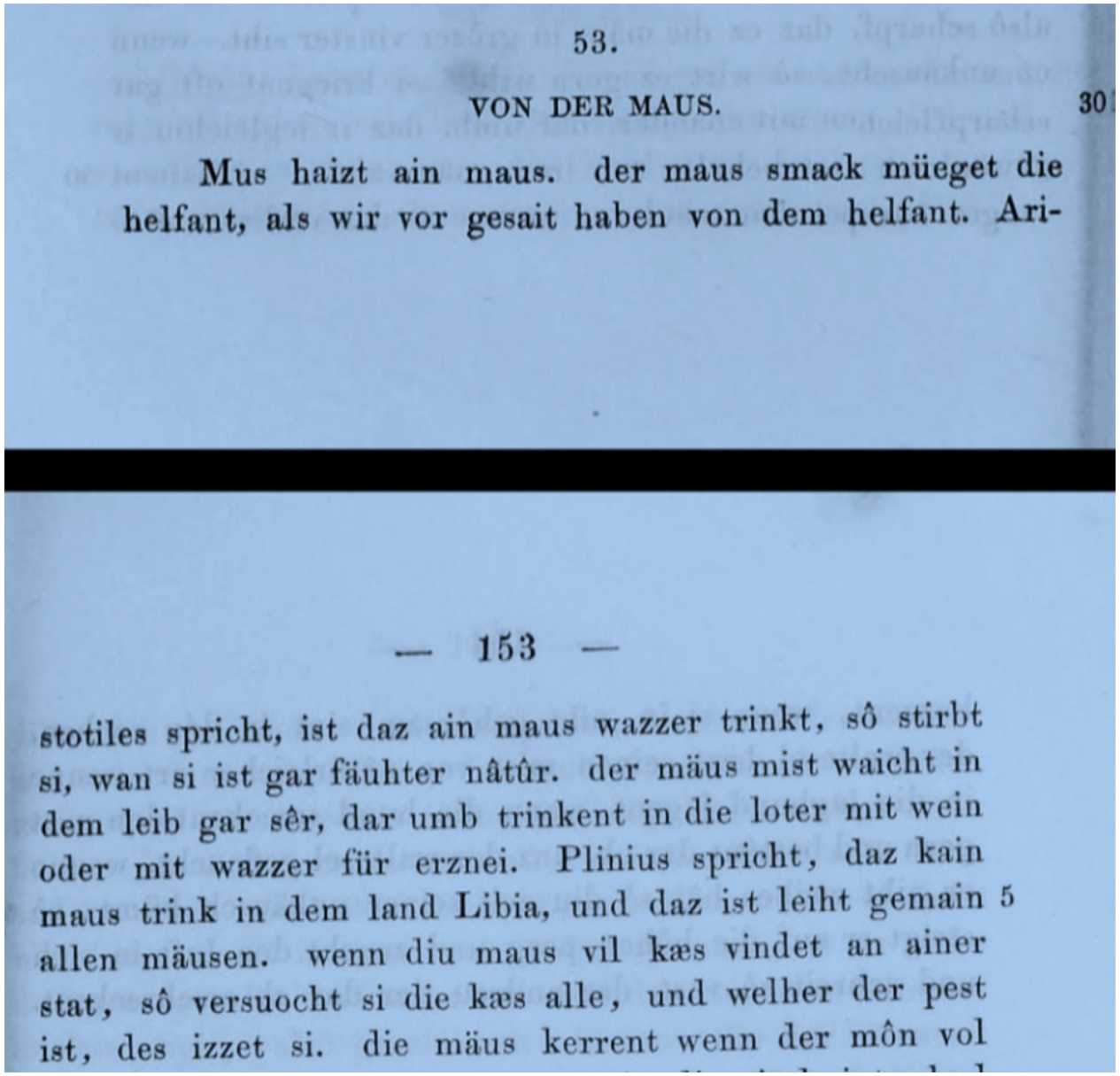
TEI tags suchen!

https://tei-c.org/release/doc/tei-p5-doc/en/html/



München, BSB Cgm 38

Von der Maus: Materialien



53.

VON DER MAUS.

Mus haizt ain maus. der maus smack müeget die helfant, als wir vor gesait haben von dem helfant. Aristotiles spricht, ist daz ain maus wazzer trinkt, sô stirbt si, wan si ist gar fäuhter nâtûr. der mäus mist waicht in dem leib gar sêr, dar umb trinkent in die loter mit wein oder mit wazzer für erznei. Plinius spricht, daz kain maus trink in dem land Libia, und daz ist leiht gemain allen mäusen. wenn diu maus vil kæs vindet an ainer stat, sô versuocht si die kæs alle, und welher der pest ist, des izzet si. die mäus kerrent wenn der môn vol

— 153 —

Mus heisst eine Maus. Der Geruch der Maus ist dem Elephanten zuwider, wie ich oben in dem Abschnitt vom Elephanten gesagt habe. Aristoteles behauptet, das eine Maus sterben muss, wenn sie Wasser trinkt, da sie sehr feuchter Natur ist. Mäusekoth wirkt innerlich sehr erweichend auf den Darminhalt, desshalb trinken liderliche Leute ihn mit Wein oder Wasser als Arznei. Plinius bemerkt, in Lybieu trinke keine Maus, was vielleicht auf alle Mäuse passt. Findet die Maus irgendwo viele Käse, so versucht sie die Käse sämmtlich und frisst dann von dem besten. Bei Vollmond geben die Mäuse einen zischenden Ton von sich, in der Zwischenzeit sind sie stumm.

Pfeiffer (1846)

Übersetzung

Studienleistung

- Codieren Sie das Kapitel „Von der Maus” in TEI XML vom Anfang bis „des izzet sie”.
- Als Text nehmen sie die Edition von Pfeiffer (links), aber berücksichtigen Sie die Zeilenumbrüche sowie die Farben, die in der Handschrift BSB Cgm 38 (die vorhergehende Folie) benutzt werden. In der <msDesc> des <teiHeader> ersetzen Sie die Handschrift aus Heidelberg mit der Handschrift aus München.
- Im Falle von Abbreviaturen, benutzen Sie die Struktur mit

```
<choice>
  <abbr></abbr>
  <expan></expan>
</choice>
```

- Wenn Sie möchten, können Sie auch eine Übersetzung des Textes ins moderne Deutsch hinzufügen (anhand eines **innerhalb** des
- Mit Oxygen, transformieren Sie die XML-Datei in eine HTML-Datei
- Bitte reichen Sie ***beide*** Dateien ein
  - XML/TEI *und* HTML

Universität zu Köln