

Faithful Transcriptions

TEI-Handbuch

Staatsbibliothek zu Berlin und Universitätsbibliothek Leipzig, 2021

Einleitung

Die Transkriptionen werden nach den P5 Guidelines der Text Encoding Initiative (TEI; <https://tei-c.org/>) erstellt und basieren auf dem Modul Representation of Primary Sources (<https://www.tei-c.org/release/doc/tei-p5-doc/en/html/PH.html>).

Nachfolgend werden die Struktur und die Elemente der TEI-Dokumente beschrieben, die im Rahmen des Transkribathons entstehen. Ein TEI-Dokument enthält dabei mehrere Seiten einer Handschrift. Da die Transkriptionen kollaborativ hergestellt werden (mehrere Personen steuern Transkriptionen aus einer Handschrift bei) und mit der Annotationsfunktion des Mirador-Viewers erstellt werden, haben Sie als Teilnehmende nur mit einem Ausschnitt des gesamten TEI-Dokuments zu tun. Die übergreifenden Teile wie der TEI-Header werden vom Organisationsteam erstellt, andere Elemente, wie z.B. die Bildkoordinaten, werden beim Export aus Mirador automatisch erstellt. Im untenstehenden Schema sind diejenigen Elemente, mit denen Sie bei der Transkription befasst sein werden, **blau hinterlegt**.

TEI-Header

<teiHeader>

Der TEI-Header enthält das Element **<fileDesc>**. Dieses enthält die Elemente **<titleStmt>**, **<publicationStmt>** und **<sourceDesc>**.

<titleStmt> enthält das Element **<title>**, darin wird „Transkription aus“ und der Titel/die Inhaltsangabe der Handschrift genannt.

<publicationStmt> enthält das Element **<authority>** mit Angabe der Veranstaltung: „Angefertigt im Rahmen des Workshops "Faithful Transcriptions" der Staatsbibliothek zu Berlin und der Universitätsbibliothek Leipzig“, das Element **<date>** mit Angabe des Jahres sowie das Element **<availability>** mit dem Kind-Element **<licence>** und der Angabe der Rechteinformation: „Veröffentlicht unter **<ref target="https://creativecommons.org/publicdomain/zero/1.0/">CC0 1.0 Universal Licence**“.

<sourceDesc> enthält das Element **<msDesc>** mit dem Kind-Element **<msIdentifier>**. Dieses wiederum hat die Kind-Elemente **<settlement>** mit Angabe des Aufbewahrungsortes, **<repository>** mit Angabe der bestandshaltenden Institution und **<idno>** mit Angabe der Signatur der Handschrift.

<TEI xmlns="http://www.tei-c.org/ns/1.0">

<teiHeader>

<sourceDoc>

<surface n="51v" xml:id="Mgf516_051v_image" type="digital" facs="https://content.staatsbibliothek-berlin.de/dc/ulx="0" uly="0" lrx="3190" lry="4785">

<zone xml:id="Mgf516_051v_page" type="page" ulx="199"

<line xml:id="Mgf516_051v_101"><w xml:id="Mgf516_051v_101_w01">oiffo</w>

<w xml:id="Mgf516_051v_101_w02">oiffo</w>

<w xml:id="Mgf516_051v_101_w03">oiffo</w>

<w xml:id="Mgf516_051v_101_w04">oiffo</w>

<w xml:id="Mgf516_051v_101_w05">oiffo</w>

<w xml:id="Mgf516_051v_101_w06">oiffo</w>

<w xml:id="Mgf516_051v_101_w07">oiffo</w>

<w xml:id="Mgf516_051v_101_w08">oiffo</w>

<w xml:id="Mgf516_051v_101_w09">oiffo</w>

<w xml:id="Mgf516_051v_101_w10">oiffo</w>

Inhalt

Einleitung	1
TEI-Header	1
TEI-SourceDoc	2
Transkriptionsregeln	3
Anhang	4
Beispiel	4
TEI-Schema	5
Übersicht Sonderzeichen/Abkürzungszeichen in Unicode ...	7
TEI-Elemente und Attribute im Faithful-Transcriptions-	
Schema: Übersicht	8



Staatsbibliothek
zu Berlin
Preußischer Kulturbesitz

UNIVERSITÄTS
BIBLIOTHEK
LEIPZIG



TEI-SourceDoc

<sourceDoc>

Innerhalb des <sourceDoc> gibt es ein Kind-Element <surface> für jede Digitalisat-Seite. Die TEI-Dokumente der Einzelseiten werden hier vom Organisationsteam in die Handschriften-Gesamtdatei eingefügt.

<surface>

<surface> hat folgende Attribute: @facs: Url des Digitalisat-Images; @n: Canvas-Id aus dem IIIF-Manifest; @ulx/uly/lrx/lry: Bildkoordinaten der gesamten Digitalisat-Seite.

<surface> enthält ein Kind-Element <zone> für die Handschriftenseite. Das Element <surface> wird beim TEI-Export aus Mirador automatisch vergeben.

<zone type="page">

Dieses Element wird manuell in einer Annotation angelegt, das die ganze Handschriftenseite (ohne die schwarzen Ränder des Digitalisats) umfasst. Das Element hat folgende Attribute: @type="page", @xml:id: Signatur_Blattzahl aufgenullt, z.B. "Mgf516_006v", @ulx/uly/lrx/lry: Bildkoordinaten der Handschriftenseite. Die xml:id muss hier manuell eingegeben werden. Die Bildkoordinaten werden beim Export automatisch vergeben.

<zone type="page"> kann folgende Kind-Elemente enthalten: <zone type="fw" | type="figure"> für Paratexte und Schmuckelemente, <line> für Textzeilen.

<zone type="fw">

Dieses Element wird manuell in einer Annotation für Paratexte angelegt. Es hat folgende Attribute: @type="fw", @xml:id: Signatur_Blattzahl aufgenullt_laufende Nummer, @ulx/uly/lrx/lry: Bildkoordinaten der Annotation. Die Attribute xml:id und Bildkoordinaten werden beim Export automatisch vergeben.

Das Element enthält die Transkription des Paratextes (s.u., Transkriptionsregeln).

<zone type="figure">

Dieses Element wird manuell in einer Annotation für Buchschmuckelemente (Illustrationen, Zierinitialen, Randleisten) angelegt. Es hat folgende Attribute: @type="figure", @xml:id: Signatur_Blattzahl aufgenullt_laufende Nummer, @ulx/uly/lrx/lry: Bildkoordinaten der Annotation. Die Attribute xml:id und die Bildkoordinaten werden beim Export automatisch vergeben.

Es enthält das Element <note> mit einer verbalen Beschreibung des Buchschmuckelements. Das Element <note> kann optional auch einen Verweis auf eine ikonographische Klassifikation (z.B. Iconclass) enthalten mit dem Element <ref> und den Attributen @type="Name der Klassifikation" @target="URL".

Eigennamen (z.B. der dargestellten biblischen Figuren) können mit einem <w>-Tag und den Attributen @type="name" subtype="person | place | org | work" lemmaRef="Link zur GND-ID" codiert werden. Die GND-ID ist der Identifier des Eigennamens in der Gemeinsamen Normdatei, die unter folgender Adresse durchsucht werden kann: <https://ognd.bsz-bw.de>.

<line>

Textzeilen werden manuell in einer Annotation pro Zeile angelegt und mit dem Element <line> codiert. Dieses hat folgende Attribute: @n: Zeilennummer in Leserichtung, @xml:id: Signatur_Blattzahl aufgenullt_laufende Nummer, @ulx/uly/lrx/lry: Bildkoordinaten der Annotation. Die Attribute xml:id und die Bildkoordinaten werden beim Export automatisch vergeben.

Das <line>-Element enthält die Transkription der entsprechenden Textzeile. Jedes transkribierte Wort steht dabei in einem <w>-Tag, der beim Export aus Mirador automatisch erstellt wird und daher nicht manuell eingefügt werden muss. Für die Transkription gelten die folgenden Regeln:

Transkriptionsregeln:

- Grundsätzlich wird buchstabengetreu transkribiert, d.h. Groß- und Kleinschreibung sowie i/j-, u/v- und s/f-Schreibung folgen der Handschrift. Sonderzeichen und Abkürzungszeichen werden, sofern nicht über die Tastatur vorhanden oder durch Kopieren eingefügt, in Unicode (hex) codiert, z.B. F;al für fal. Eine Liste der häufigsten Sonderzeichen als Kopiervorlage sowie ihrer Unicode-Codierung findet sich im Anhang. Beim Kopieren von kombinierten Zeichen (beispielsweise a mit Nasalstrich ã) muss das Spatium zwischen Buchstabe (a) und Sonderzeichen (ã) in Mirador gelöscht werden, damit die Kombination funktioniert.
- Die Interpunktion und die Setzung von Worttrennungszeichen folgen der Handschrift.
- Hervorhebungen (z.B. Initialen, Rubrizierung) werden mit dem Element **<hi>** codiert. **<hi>** hat das Attribut @rend, das spezifiziert, ob es sich um ein rubriziertes Wort handelt (rend="color: red") oder um ein Zierelement (z.B. Majuskelstrichelung) an einem Buchstaben (rend="decoration-color: blue") oder um eine Initiale (rend="ital-color: red"). Bei Initialen, die mehr als eine Zeile umfassen, wird außerdem die Größe mit dem Attribut @n="Anzahl der umfassten Zeilen" angegeben. Handelt es sich um eine aufwändig gestaltete (z.B. historisierte) Initiale, die auch als eigene <zone type="figure"> codiert wurde, kann außerdem auf die xml:id dieser <zone> verwiesen werden mit dem Attribut: @corresp="#xml:id der Zone". Das Attribut @corresp sollte erst nachträglich im TEI-Export-Dokument eingefügt werden, da die xml:id der Zone erst beim Export automatisch generiert wird.
- Abkürzungen werden sowohl handschriftengetreu wiedergegeben als auch aufgelöst. Dies wird mit dem Element **<choice>** codiert.
 - <choice> enthält die Kind-Elemente <abbr> und <expan>. **<abbr>** enthält die Transkription des abgekürzten Wortes, wobei der von der Abkürzung betroffene Buchstabe/das Abkürzungszeichen mit dem Element **<am>** gekennzeichnet wird. **<expan>** enthält die Transkription des aufgelösten Wortes, wobei der aufgelöste Teil mit dem Element **<ex>** gekennzeichnet wird. Beispiel:
<choice><abbr>nam<am>ē</am></abbr><expan>nam<ex>en</ex></expan></choice>
- Handschriftliche Tilgungen werden mit dem Element **** codiert. Dies kann mit dem Attribut @rend spezifiziert werden: @rend="strikethrough | overwritten | erasure | adapted"
- Handschriftliche Ergänzungen werden mit dem Element **<add>** codiert. Die Platzierung der Ergänzung kann mit dem Attribut @place spezifiziert werden: @place="above | below | inline | superimposed | rmargin | lmargin | betweenlines"
- Eigennamen können mit einer Referenz zur Gemeinsamen Normdatei (GND) versehen werden (s. <https://ognd.bsz-bw.de>). Dafür wird der Eigenname mit dem Element **<w>** codiert. Dies hat die Attribute @type="name" subtype="person | place | org | work" lemmaRef="Link zur GND-ID"

Anhang

Beispiel

Im folgenden Beispiel ist ein Screenshot einer annotierten Handschriftenseite im Mirador-Viewer zu sehen. Darauf folgt ein vollständiges TEI-Dokument für diese annotierte Seite. Die Anteile der TEI-Dokuments, die im Rahmen des Transkribathons von den Teilnehmenden beige-steuert werden, sind dabei **blau hinterlegt**. Die restlichen Bestandteile des TEI-Dokuments werden entweder vom Organisationsteam vergeben (z.B. der `teiHeader`) bzw. automatisch beim TEI-Export aus Mirador erstellt (z.B. die Bildkoordinaten und die `<w>`-Tags).

Handschriftenportal

Alles ▾ ms. germ. fol. 516 SUCHEN

Niederrheinische Historienbibel

Annotations

7 Annotationen werden angezeigt

OBJEKT: [6V [18]]

<zone type="page" xml:id="Mgf516_006v"></zone>

<zone type="fw"><hi rend="color: red">Gene</hi></zone>


<zone type="figure"><note>Opferung <w type="name" subtype="person" lemmaRef="http://d-nb.info/gnd/118555898">Isaaks.</w></note></zone>

<line n="05"><choice><abbr>nam<am>ē</am></abbr><expan>nam<ex>en</ex></expan></choice> des ewigen godes</line>

<line n="22">me der berge den ich dych</line>

<line n="23">dan wysen <hi rend="decoration-color: red">H</hi>erumb ftoyn</line>

<line n="24">de <hi rend="decoration-color: red">A</hi>braham up des nachtes</line>



18 von 716 · 6v [18]

TEI-Schema

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Transkription aus der Niederrheinischen Historienbibel</title>
      </titleStmt>
      <publicationStmt>
        <authority>Angefertigt im Rahmen des Workshops "Faithful Transcriptions" der
          Staatsbibliothek zu Berlin und der Universitätsbibliothek Leipzig</authority><date>2021</date>
        <availability>
          <licence>Veröffentlicht unter <ref
            target="https://creativecommons.org/publicdomain/zero/1.0/" />CC0 1.0 Universal
          Licence</licence>
        </availability>
      </publicationStmt>
      <sourceDesc>
        <msDesc>
          <msIdentifier>
            <settlement>Berlin</settlement>
            <repository>Staatsbibliothek zu Berlin - Preußischer Kulturbesitz</repository>
            <idno>Ms. germ. fol. 516</idno>
          </msIdentifier>
          </msDesc>
        </sourceDesc>
      </fileDesc>
    </teiHeader>
    <sourceDoc>
      <surface facs="https://content.staatsbibliothek-berlin.de/dc/757246826-0018/full/full/0/default.jpg"
        n="https://content.staatsbibliothek-berlin.de/dc/757246826-0018/canvas" ulx="0" uly="0"
        lrx="3190" lry="4785">
        <zone type="page" xml:id="Mgf516_006v" ulx="217" uly="421" lrx="3087" lry="4330">
          <zone type="fw" ulx="1638" uly="528" lrx="1912" lry="679" xml:id="Mgf516_006v_001">
            <hi rend="color: red">Gene</hi>
          </zone>
          <zone type="figure" ulx="354" uly="1131" lrx="2999" lry="2590" xml:id="Mgf516_006v_002">
            <note>Opferung <w type="name" subtype="person" lemma-ref="http://d-nb.info/gnd/118555898">
```

```
>Isaaks.</w></note>
</zone>
<line n="05" ulx="844" uly="1072" lrx="1740" lry="1150" xml:id="Mgf516_006v_003">
  <w>
    <choice>
      <abbr>nam<am>ē</am></abbr>
      <expan>nam<ex>en</ex></expan>
    </choice>
  </w>
  <w>des</w>
  <w>ewigen</w>
  <w>godes</w>
</line>
<line n="22" ulx="1936" uly="2748" lrx="2847" lry="2831" xml:id="Mgf516_006v_004">
  <w>me</w>
  <w>der </w>
  <w>berge</w>
  <w>den</w>
  <w>ich</w>
  <w>dych</w>
</line>
<line n="23" ulx="1922" uly="2806" lrx="2803" lry="2903" xml:id="Mgf516_006v_005">
  <w>dan</w>
  <w>wysen </w>
  <w><hi rend="decoration-color: red">H</hi>erumb</w>
  <w>ftoyn</w>
</line>
<line n="24" ulx="1922" uly="2890" lrx="2852" lry="2968" xml:id="Mgf516_006v_006">
  <w>de</w>
  <w><hi rend="decoration-color: red">A</hi>braham</w>
  <w>up</w>
  <w>des</w>
  <w>nachtes</w>
</line>
</zone>
</surface>
</sourceDoc>
</TEI>
```

Übersicht Sonderzeichen/Abkürzungszeichen in Unicode

MUFI Code chart: <https://mufi.info/m.php?p=muficodechart>

Zeichen	Anwendung	Code Point in Unicode (einfügen mit odePoint;)
f	Lang-s	017F
z	Geschwänztes z	0292
er	er- / re-Kürzung	035B
o	(Nasal-)strich	0304
us	us-Kürzung	A770
o	Hochgestellte us-Kürzung	1DD2
con	con-Kürzung	A76F
ra	ra-Kürzung	1DD3
ur	ur-Kürzung	1DD1
p	per-Kürzung	A751
p	pro-Kürzung	A753
q	q-Kürzung	A757
q	q-Kürzung	A759
rum	rum-Kürzung	A75D
-que	-que-Kürzung	A76B
¶	Capitulumzeichen	2E3F
o	Tilde	0303

Allgemeine Literaturhinweise zu Abkürzungen:

- Karin Schneider, Paläographie und Handschriftenkunde für Germanisten. Eine Einführung, 3., durchges. Aufl., Berlin 2014, S. 86-91.
(auch als eBook: DOI: 10.1515/9783110338676)
- Bernhard Bischoff, Paläographie des römischen Altertums und des abendländischen Mittelalters, 4., durchges. u. erw. Aufl., Berlin 2009, S. 202-223.
(Englische Übersetzung auch als eBook verfügbar: DOI: 10.1017/CBO9780511809927, dort S. 150-168)
- A. Cappelli, Lexicon abbreviatarum. Dizionario di abbreviature latine ed italiane, 5. Aufl., Milano 1954.
Deutsche Übersetzung auch als Digitalisat verfügbar:
<http://www.ub.uni-koeln.de/cdm/ref/collection/mono20/id/7918>
Datenbankversion:
<https://www.adfontes.uzh.ch/ressourcen/abkuerzungen/cappelli-online>
- Datenbank Abbreviationes Online. Electronic Dictionary of Medieval Latin Abbreviations (Im Datenbank-Informationssystem DBIS können Sie die Verfügbarkeit in Ihrer Bibliothek prüfen: https://dbis.uni-regensburg.de/frontdoor.php?titel_id=2506)

TEI-Elemente und Attribute im Faithful-Transcriptions-Schema: Übersicht

<surface>	facs ="URL_BILDDATEI" ulx ="X_OBEN" uly ="Y_OBEN" [Koordinaten Ecke oben links] lrx ="X_UNTEN" lry ="Y_UNTEN" [Koordinaten Ecke oben rechts]	Digitalisat-Seite, schließt die Textteile ein
<zone>	xml:id ="SIGNATUR_SEITE_LAUFENDE NUMMER" type ="page" → Seite, darunter: type ="figure" → Illustration, Zierinitiale type ="fw" → Paratexte wie Kolumnentitel, Reklamanten etc. ulx ="X_OBEN" uly ="Y_OBEN" [Koordinaten Ecke oben links] lrx ="X_UNTEN" lry ="Y_UNTEN" [Koordinaten Ecke oben rechts]	Layout-Einheit (Seite, Paratext, Schmuckelement)
<note>		Anmerkung, z.B. verbale Beschreibung einer Illustration
<ref>	type ="Name der verwendeten Klassifikation" (z.B. Iconclass) target ="URL"	Verweis auf eine ikonographische Klassifizierung der Illustration, z.B. Iconclass
<line>	xml:id ="SIGNATUR_SEITE_LAUFENDE NUMMER" n ="Zeilennummer in Leserichtung" ulx ="X_OBEN" uly ="Y_OBEN" [Koordinaten Ecke oben links] lrx ="X_UNTEN" lry ="Y_UNTEN" [Koordinaten Ecke oben rechts]	Zeile
<w>	type ="name" subtype ="person place org work" lemmaRef ="Link zur GND-ID"	Eigennamen werden mit @type/subtype gekennzeichnet und mit @lemmaRef zur GND verlinkt
<hi>	<i>Schmuckrubrizierung (z.B. Majuskelstrichelung):</i> rend ="decoration-color: red" <i>Ganze/Mehrere Buchstaben/Wörter in anderer Farbe:</i> rend ="color: red" <i>Initiale:</i> rend ="initial-color: blue" n ="3" corresp ="#xml:id der Zone"	Schmuckelemente, Hervorhebungen, Initialen
<choice>	<choice> <abbr><am></am></abbr> <expan><ex></ex></expan> </choice>	Abkürzungen (<abbr>: ganzes abgekürztes Wort; <am>: eigentliche Abkürzung; <expan>: ganzes aufgelöstes Wort; <ex>: aufgelöste Abkürzung)
	rend ="strikethrough overwritten erasure adapted"	Tilgung
<add>	place ="above below inline superimposed rmargin lmargin betweenlines"	Ergänzung