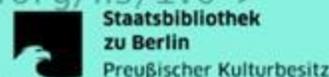


# Faithful Transcriptions

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



UNIVERSITÄTS  
BIBLIOTHEK *ubl* LEIPZIG

<surface

facts="https://content.staatsbibliothek-berlin  
source="https://content.staatsbibliothek-berl  
uly="0" lrx="3190" lry="4785">

<zone type="page" xml:id="Mgf516\_006v" ulx="1  
<line n="01" ulx="842" uly="773" lrx="183

<w>weder</w>

<w>in</w>

<w>dat</w>

<w>lant</w>

<w>der</w>

<w>pa-</w>

</line>

</zone>

</surface>

</sourceDoc>

</TEI>

*gryfft  
bedeket*

*So sal den here des  
betzalen dat vint*





# Transkribathon – Konzept





# Transkribathon-Datenset



Transkribathon „Faithful Transcriptions“



# Nachnutzungsszenarien

- Ground truth-Datensets für HTR/Layoutanalyse
- Grundlage für digitale Editionen
- Grundlage für linguistische Analysen

Staatsbibliothek zu Berlin Preußischer Kulturbesitz SBB-Lab

HOME DATEN DEMOS EVENTS

### Datensets des Transkribathons „Faithful Transcriptions“

Von Mai bis Juli 2021 richtete die Staatsbibliothek zu Berlin in Kooperation mit der Universitätsbibliothek Leipzig das Transkribathon „Faithful Transcriptions“ aus, ein digitales Crowd-Sourcing-Projekt zu theologischen Handschriften des Mittelalters. Über 100 Teilnehmer fertigten dabei TEI/xml-codierte Transkriptionen von 181 Handschriftenseiten an. Als Arbeitsplattform diente der IIF-basierte Handschriftenportal, in Entwicklung befindlichen Handschriftenportals. Die Datensets stehen nach Abschluss der Qualitätskontrolle zur Publikation auf Zenodo zur Verfügung.

#### Übersicht – alle Transkriptionen

<b>Materialien:</b>	Transkriptionen aus 12 mittelalterlichen Handschriften in deutscher, niederländischer, Textura, Textualis, Gotische Kursive, Bastarda
<b>Schriften:</b>	Textura, Textualis, Gotische Kursive, Bastarda
<b>Anzahl Seiten:</b>	181
<b>Anzahl Zeilen:</b>	8952
<b>Formate:</b>	xml-codierte Transkriptionen nach TEI-Schema; html
<b>Lizenzen:</b>	CC0
<b>Links:</b>	Transkriptionen (TEI/xml) auf Zenodo (DOI: 10.5281/zenodo.5582483) Transkriptionen (TEI/xml und html) im SBB-Lab Dokumentation (TEI-Schema) Bericht über den Transkribathon in 027.7 Zeitschrift für Bibliothekskultur

zenodo Search Upload Communities Log in Sign up

October 20, 2021 Dataset Open Access

## Faithful Transcriptions Data Set: TEI/XML-encoded Transcriptions of Medieval Theological Manuscripts

77 views 39 downloads

Indexed in OpenAIRE

Publication date: October 20, 2021  
DOI: 10.5281/zenodo.5582483

Keywords: Digital Humanities, Medieval manuscripts, Transcription, Crowdsourcing, Handwritten Text Recognition

Related identifiers: Documented by 10.21428/1bfadeb6.d3bbd02 (Journal article)

License (for files): Creative Commons Zero v1.0 Universal

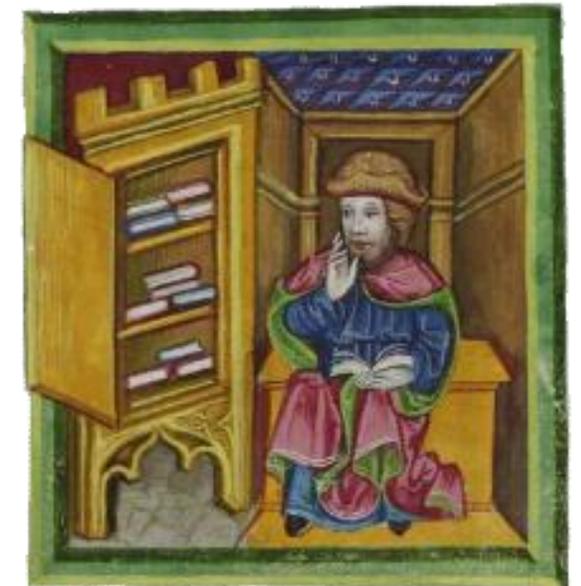
Versions

Preview

Faithful Transcriptions

<https://lab.sbb.berlin/datensets-transkribathon/>

<https://www.doi.org/10.5281/zenodo.5582483>





# Rahmenprogramm

**3. Mai:**  
**Beginn der Transkriptionsphase und Tutorium**

**6. Mai:** Eröffnungsveranstaltung mit Keynote-Vortrag  
*alte kuntschaft und ouch niuwe. Digitale  
Perspektiven auf die Mediävistik*

Prof. Dr. Andrea Rapp / Dr. Luise Borek (Technische Universität Darmstadt)

**11. Mai:** Vortrag  
**Einführung in die Text Encoding Initiative (TEI)**

Torsten Schaßan (Herzog August Bibliothek Wolfenbüttel)

**18. Mai:** Tutorium

**25. Mai:** Tutorium

**30. Mai:** **Ende der Transkriptionsphase**

**10. Juni:** Vortrag  
**heiEDITIONS: Die TEI an der  
Universitätsbibliothek Heidelberg**

Dr. Jakub Šimek (Universitätsbibliothek Heidelberg)

**15. Juni:** Vortrag  
**OCR4all – Möglichkeiten und Grenzen einer  
Erfassung historischer Drucke und Handschriften**

Dr. Christian Reul (Universität Würzburg)

**1. Juli:** Abschlussveranstaltung: Ergebnispräsentation  
und Podiumsdiskussion

***Podiumsdiskussion: Transkribathon im Kontext.  
Aktuelle Perspektiven aus  
Geisteswissenschaften, Informatik und Digital  
Humanities***

Prof. Dr. Racha Kirakosian (Universität Freiburg); Dr. Frederike Neuber (Berlin-  
Brandenburgische Akademie der Wissenschaften); Leander Seige  
(Universitätsbibliothek Leipzig)

Moderation: Dr. Christoph Mackert (Universitätsbibliothek Leipzig)



# Material zur Anleitung

## Informationen zur Transkription

### Transkriptionsschema: Codierung in TEI

Die Transkription wird nach einem TEI-konformen Schema in der Markup-Sprache xml codiert. Weitere Informationen zum Schema und den Transkriptionsregeln finden Sie in den untenstehenden Anleitungen.

### Anleitungen

Hier finden Sie verschiedene Anleitungen für die praktische Transkriptionsarbeit:

- **Anleitung mit Videos** zur Transkription in Mirador
- **TranskriptionsanleitungV3** als PDF
- **TEI\_HandbuchV3** zum Transkribathon mit Erklärung des verwendeten xml-Schemas
- **TEI\_SchemaV3\_Übersicht**: tabellarische Übersicht zu den im xml-Schema verwendeten Elementen und Attributen

Wenn Sie Fragen zum Transkribieren haben, können sie diese gern in unseren Tutorien oder auch jederzeit per Mail an [lab@sbb.spk-berlin.de](mailto:lab@sbb.spk-berlin.de) stellen.

## Anleitungen

Hier finden Sie verschiedene Anleitungen für die praktische Transkriptionsarbeit:

- **Anleitung mit Videos** zur Transkription in Mirador
- **TranskriptionsanleitungV3** als PDF
- **TEI\_HandbuchV3** zum Transkribathon mit Erklärung des verwendeten xml-Schemas
- **TEI\_SchemaV3\_Übersicht**: tabellarische Übersicht zu den im xml-Schema verwendeten Elementen und Attributen



# Material zur Anleitung



Staatsbibliothek  
zu Berlin  
Preußischer Kulturbesitz

UNIVERSITÄTS  
BIBLIOTHEK  
ubl  
LEIPZIG

The screenshot displays the 'Handschriftenportal' interface. At the top, there is a search bar with 'Ms.germ.fol. 1108' and buttons for 'SUCHEN' and 'ERWEITERTE SUCHE'. The main content area shows a manuscript page with a large illuminated initial 'D' and a miniature of a figure holding a staff. The page is annotated with XML tags. On the left, a sidebar shows 'Historienbibel' and 'Annotationen' with a list of 2 annotations. The first annotation is for the page '58v [120]' and contains the XML code: 

```
<zone type="page" xml:id="Mgf1108_058v">
  <line n="18">vnd er sprach zu got warumb haftu</line>
```



# Material zur Anleitung

## Informationen zur Transkription

### Transkriptionsschema: Codierung in TEI

Die Transkription wird nach einem TEI-konformen Schema in der Markup-Sprache xml codiert. Weitere Informationen zum Schema und den Transkriptionsregeln finden Sie in den untenstehenden Anleitungen.

### Anleitungen

Hier finden Sie verschiedene Anleitungen für die praktische Transkriptionsarbeit:

- [Anleitung mit Videos](#) zur Transkription in Mirador
- [TranskriptionsanleitungV3](#) als PDF
- [TEI\\_HandbuchV3](#) zum Transkribathon mit Erklärung des verwendeten xml-Schemas
- [TEI\\_SchemaV3\\_Übersicht](#): tabellarische Übersicht zu den im xml-Schema verwendeten Elementen und Attributen

Wenn Sie Fragen zum Transkribieren haben, können sie diese gern in unseren Tutorien oder auch jederzeit per Mail an [lab@sbb.spk-berlin.de](mailto:lab@sbb.spk-berlin.de) stellen.

## Anleitungen

Hier finden Sie verschiedene Anleitungen für die praktische Transkriptionsarbeit:

- [Anleitung mit Videos](#) zur Transkription in Mirador
- [TranskriptionsanleitungV3](#) als PDF
- [TEI\\_HandbuchV3](#) zum Transkribathon mit Erklärung des verwendeten xml-Schemas
- [TEI\\_SchemaV3\\_Übersicht](#): tabellarische Übersicht zu den im xml-Schema verwendeten Elementen und Attributen





# Technische Umgebung



Staatsbibliothek  
zu Berlin  
Preußischer Kulturbesitz

UNIVERSITÄTS  
BIBLIOTHEK  
ubl LEIPZIG

Testversion des Handschriftenportals

## Handschriftenportal

<https://alpha.handschriftenportal.de>

Basierend auf dem International  
Image Interoperability Framework  
und dem IIF-Viewer Mirador.





# Vielen Dank!



**Staatsbibliothek  
zu Berlin**  
Preußischer Kulturbesitz

UNIVERSITÄTS  
BIBLIOTHEK *ubl* LEIPZIG

Dr. Nicole Eichenberger, Dr. Hedwig Suwelack (Staatsbibliothek zu Berlin),  
Annika Schröer (Universitätsbibliothek Leipzig)

Weitere Informationen zum Transkribathon:

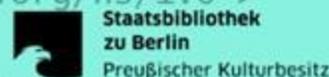
<https://lab.sbb.berlin/faithful-transcriptions/>

Kontakt:

[lab@sbb.spk-berlin.de](mailto:lab@sbb.spk-berlin.de)

# Faithful Transcriptions

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



UNIVERSITÄTS  
BIBLIOTHEK *ubl* LEIPZIG

<surface

facts="https://content.staatsbibliothek-berlin  
source="https://content.staatsbibliothek-berl  
uly="0" lrx="3190" lry="4785">

<zone type="page" xml:id="Mgf516\_006v" ulx="183  
<line n="01" ulx="842" uly="773" lrx="183

<w>weder</w>

<w>in</w>

<w>dat</w>

<w>lant</w>

<w>der</w>

<w>pa-</w>

</line>

</zone>

</surface>

</sourceDoc>

</TEI>

*gryfft  
bedeket*

*So sal den here des  
betzalen dat vint*

