



ANGILA VETTER / ELKE ZINSMEISTER

## Die Bibel für alle – der Österreichische Bibelübersetzer auf dem Weg ins Web

Ein Werkstattbericht

Seit 2016 wird in einem Langzeitprojekt der Akademienunion mit Arbeitsstellen an der Universität Augsburg unter Leitung der Bayerischen Akademie der Wissenschaften und an der Berlin-Brandenburgischen Akademie der Wissenschaften in Berlin das Gesamtwerk des sogenannten ‚Österreichischen Bibelübersetzers‘ ediert.

Von der historischen Person sind bislang nur vage Umrisse bekannt. Aus der in den überlieferten Vorreden beschriebenen Arbeitssituation wissen wir, was sein Anliegen war: große Teile der Bibel aus dem Lateinischen ins Deutsche zu übersetzen, zu kommentieren und auszulegen; und aus dem Verbreitungsgebiet seiner Werke können wir grob sagen, wo und wann er dies vermutlich unternahm: in der Zeit um 1320/40 im heutigen Österreich, etwa zwischen den Grenzen der Diözesen Salzburg und Passau. So ergab sich der Notname ‚Österreichischer Bibelübersetzer‘.<sup>1</sup> Er ist ohne Zweifel einer der bedeutendsten deutschen Bibelübersetzer der Zeit vor Luther. Aber selbstverständlich richtet er sich nicht wie Luther nach dem hebräischen bzw. griechischen Urtext, sondern wie nahezu alle mittelalterlichen Übersetzer in Europa nach der Vulgata. Übersetzung, Auslegung und Kommentar werden bei ihm wie in einer mittelalterlichen Predigt eng verbunden, wobei er eine beeindruckende Menge von Quellen einbezieht. Offenbar selbst sehr gut ausgebildet, sagt er von sich, dass er nicht zum Predigen ordiniert sei, er will vielmehr als Laie für

---

<sup>1</sup> Dies nach Freimut LÖSER /Christine STÖLLINGER-LÖSER: Verteidigung der Laienbibel. Zwei programmatische Vorreden des österreichischen Bibelübersetzers der ersten Hälfte des 14. Jahrhunderts, in: Überlieferungsgeschichtliche Editionen und Studien zur deutschen Literatur des Mittelalters (Texte und Textgeschichte 31), hrsg. v. Konrad Kunze /Johannes G. Mayer /Bernhard Schnell, Tübingen 1989, S. 245–313.

Laien die Heilige Schrift in die Volkssprache bringen. Sein Vorhaben bedurfte demnach mehr als ohnehin für eine Übersetzung in die Volkssprache notwendig der Rechtfertigung. Strategien der Verteidigung gegenüber Anklagen des Klerus finden sich im Œuvre des Bibelübersetzers vor allem in den deutschen Vorreden zu seinen einzelnen Werken, in denen er selbstbewusst sein laikales Anliegen vertritt.<sup>2</sup> Gleichzeitig greift er selbst immer wieder vermeintliche Falsch- und Irrgläubige an, grenzt sich also gegen angebliche Fehldeutungen ab. Doch ein genauerer Aufschluss über den ‚Sitz im Leben‘ seiner Werke kann nur aus den Texten selbst gewonnen werden. Dass sich seine Arbeit ausnehmender Beliebtheit erfreute, belegt die umfangreiche Rezeption seiner Werke. Insgesamt sind uns rund 120 Handschriften und Fragmente bekannt.

Das sogenannte ‚Alttestamentliche Werk‘, eine kommentierte und auslegende Übertragung ausgewählter Bücher des Alten Testaments, ist in sechs Textzeugen überliefert; es wurde bereits in einem DFG-geförderten Vorgängerprojekt zur Edition vorbereitet und wird im Rahmen unseres Vorhabens zum Abschluss gebracht. Das ‚Evangelienwerk‘, eine kommentierte und auslegende Übersetzung der vom Bibelübersetzer harmonisierten Evangelien, angereichert durch Apokryphen und Legenden, zählt 30 Textzeugen und liegt in zwei Fassungen vor. Es wird derzeit in Augsburg und Berlin ediert. Daran anschließen wird sich ab 2022 in einem zweiten sechsjährigen Modul die Edition des ‚Psalmenkommentars‘, der uns bislang in 72 Textzeugen bekannt ist und in drei Fassungen vorliegt.

Derzeit arbeiten die Teams in Augsburg und Berlin an der Erschließung der Überlieferungslage und Textgeschichte des ‚Evangelienwerks‘. Das Vorgehen richtet sich dabei an der geplanten digitalen Edition aus, da sich die Daten für die Printedition aus denjenigen, die für die digitale Ausgabe erarbeitet werden, generieren lassen. In der digitalen Edition sollen neben dem handschriftlichen Material in Form von Digitalisaten Forschungsdaten wie Transkriptionen und Kollationen präsentiert werden; die Editionstexte der beiden Fassungen werden durch verschiedene

---

<sup>2</sup> Dazu ausführlich LÖSER/STÖLLINGER-LÖSER: Verteidigung der Laienbibel.

Register erschlossen und mit anderen digitalen Quellen (z. B. digitalen Editionen der Vulgata oder modernen Bibelübersetzungen) und Hilfsmitteln (z. B. Wörterbuchnetz, Wikipedia) verknüpft; eine synoptische Darstellung der beiden Fassungen ermöglicht zudem deren Vergleich.<sup>3</sup>

Die beiden Fassungen des ‚Evangelienwerks‘ unterscheiden sich nicht nur auf inhaltlicher Ebene, sondern auch in der Art und dem Alter ihrer Überlieferungszeugen. Die Erstfassung \*Gö – bezeichnet nach ihrer Leithandschrift Göttweig, Stiftsbibl., Cod. 222 (rot) / 198 (schwarz) (früher C 2) – ist nur in der einen Vollhandschrift auf uns gekommen, zu der sich acht Fragmente und zehn Exzerpthandschriften gesellen. Wir haben es hier also mit einer vorwiegend sehr kleinteiligen Überlieferung zu tun, und über weite Strecken des Textes ist Handschrift Gö ohne Parallelüberlieferung. Der Codex wird in das sechste Jahrzehnt des 15. Jahrhunderts datiert und ist damit die jüngste der erhaltenen Vollhandschriften, die teilweise modernisierende Tendenzen im Sprachstand aufweist. Durch die beschriebene Überlieferungssituation kommt trotzdem nur sie als Leithandschrift der älteren Fassung in Frage. Die Bearbeitung dagegen ist in vier Vollhandschriften erhalten, zwei Fragmenten und drei Exzerpthandschriften. Als Leithandschrift fungiert der älteste Textzeuge, die prächtige, mit über 429 erhaltenen Federzeichnungen illustrierte Pergamenthandschrift S (Schaffhausen, Stadtbibl., Cod. Gen. 8) von um 1340, die damit einen sehr autorzeitnahen Sprachstand bezeugt. Wo sie wegen Blattverlust ausfällt, springen die Handschriften K2 (Klosterneuburg, Stiftsbibl., Cod. 51) von 1415 oder K1 (Klosterneuburg, Stiftsbibl., Cod. 4) von etwa 1410 ein. Nach diesen Handschriften wird die Bearbeitung als Fassung \*SK bezeichnet.<sup>4</sup>

---

<sup>3</sup> „Editionen ‚on demand‘“, nennt Thomas BEIN: Leerstellen edieren? Überlegungen zur Einbindung von Performanz in Editionen mittelalterlicher Literatur, in: *editio* 32 (2018), S. 82–92, hier: S. 90, solche Plattformen; sein Beispiel ist die von Sonja GLAUCH / Manuel BRAUN / Florian KRAGL erstellte digitale Präsentation der ‚Lyrik des deutschen Mittelalters‘: <http://ldm-digital.de>. Verschiedene Textzustände lassen sich dort ebenso wie unterschiedliche Niveaus editorischer Eingriffe der eigenen Fragestellung gemäß anzeigen, arrangieren und ausdrücken.

<sup>4</sup> Als vierte Vollhandschrift kommt Nü (Nürnberg, Germanisches Nationalmuseum, Löffelholz-Archiv D 654) aus der ersten Hälfte des 15. Jahrhunderts hinzu. Mit N (Neiße,

Die inhaltlichen Fassungsunterschiede definieren sich zum einen durch die unterschiedliche Gestaltung der Kapitelübergänge und zum anderen durch Änderungen in der Anordnung des Textes sowohl bei der Reihenfolge der Kapitel als auch in einzelnen Kapiteln bei der Reihenfolge der Absätze.<sup>5</sup> Diese umfangreichen Textumstellungen lassen eine Fassungssynopse im Druck impraktikabel erscheinen, sinnvoll darstellen lässt sie sich allein in der digitalen Edition. Aus pragmatischen Gründen (siehe unten) wurde die Bearbeitung der Fassungen unter den beiden Arbeitsstellen aufgeteilt, die Fassung \*Gö wird von der Berliner Arbeitsstelle, die Fassung \*SK von der Augsburger Arbeitsstelle betreut.

Das Datenmodell für die angestrebte Hybridedition soll gewährleisten, dass die verschiedenen Ebenen der Edition bzw. der ‚eine Text‘ in seinen unterschiedlichen Erscheinungsformen (Handschriftenimage, Transkription, kritischer Text inklusive Metadaten sowie Fassungssynopse) über dynamische Verknüpfungen aufrufbar und transparent gehalten werden. Die Datenerfassung erfolgt in unserem Vorhaben daher auch aus zwei Perspektiven: der ‚werkbasierten‘ und der ‚codexbasierten‘ Perspektive.<sup>6</sup> Unterschiedlich wird dabei das jeweilige Verhältnis von Textualität und Materialität gewichtet, dennoch stehen die über diese Perspektiven erfassten Texte gleichberechtigt nebeneinander, ergänzen und

---

Gymnasialbibl., Cod. A VIII 9) lag bis zum Ende des Zweiten Weltkriegs eine weitere Vollhandschrift vor, wohl eine direkte Abschrift von S und wie diese eine kunstreich ausgestattete Pergamenthandschrift. Sie gilt heute als verschollen, abgesehen von fünf Einzelblättern, die in den USA zum Verkauf angeboten worden waren und heute in der Staatsbibliothek zu Berlin bzw. in amerikanischen Bibliotheken liegen. Zudem existieren eine Reihe von Schwarz-Weiß-Fotos der bebilderten Seiten sowie eine Beschreibung der Vollhandschrift.

<sup>5</sup> Vgl. Gisela KORNRUMPF: Das ‚Klosterneuburger Evangelienwerk‘ des österreichischen Anonymus. Datierung, neue Überlieferung, Originalfassung, in: Deutsche Bibelübersetzungen des Mittelalters. Beiträge eines Kolloquiums im Deutschen Bibel-Archiv, hrsg. v. Heimo Reinitzer unt. Mitarb. v. Nikolaus Henkel (Vestigia Bibliae 9/10 [1987/1988]), Bern 1991, S. 115–131, hier: S. 123–124.

<sup>6</sup> Dies bereits bei Jakob ŠIMEK: Welscher Gast digital. TEI Handbuch (Version 0.6), hier § 3, zur Plattform ‚Welscher Gast digital‘: <https://digi.ub.uni-heidelberg.de/wgd/>; PDF des Handbuchs unter: [https://digi.ub.uni-heidelberg.de/wgd/TEI-Handbuch\\_0-6.pdf](https://digi.ub.uni-heidelberg.de/wgd/TEI-Handbuch_0-6.pdf).

erweitern sich.<sup>7</sup> Die ‚werkbasierte Perspektive‘ strukturiert die Daten primär nach hierarchischen Einheiten des abstrakt gedachten Werks, also nach Büchern, Kapiteln, Absätzen, Sätzen, Wörtern etc. Die ‚codexbasierte Perspektive‘ legt hingegen jeweils eine physische Handschrift zugrunde. In den Transkriptionen werden aus der codexbasierten Perspektive so neben der Varianz und den Lesarten der einzelnen Textzeugen auch die kodikologischen und paläographischen Parameter sowie etwaige Illustrationen erfasst.

Die Datenerfassung erfolgt in XML und richtet sich hierbei nach dem etablierten Standard der TEI-Richtlinien. XML als Basistechnologie der TEI ist unbestritten ein Format, das eine hohe Interoperabilität mit diversen Applikationen besitzt. Probleme für die Interoperabilität bereitet aber die Ebene der Codierung der Daten nach der TEI. Die zahlreichen möglichen Auszeichnungen für ein und dasselbe Phänomen sind sicherlich notwendig, um den jeweils spezifischen Anforderungen der von den verschiedenen geisteswissenschaftlichen Disziplinen bearbeiteten Texte gerecht zu werden, die ‚menschliche Interpretationsebene‘, die damit dem Datensatz beigegeben wird, mag in der menschlichen Kommunikation noch zu deuten sein, die maschinelle Interoperabilität wird damit aber erschwert.<sup>8</sup> Um die Interoperabilität unserer Daten zu erhöhen, richten wir uns daher nach dem *Deutschen Textarchiv Basisformat* (DTABf), das ein Tagset aus der TEI darstellt und explizit auf sogenannte ‚tei.extensions‘ verzichtet.<sup>9</sup> Das DTABf hat zum Ziel „eine umfassende

---

<sup>7</sup> Dies ganz im Sinne des von Patrick Sahle geforderten ‚erweiterten Textbegriffs‘ und eines ‚pluralistischen Textmodells‘, welche die angestrebte Darstellung der Vielschichtigkeit des ‚einen Textes‘ erlauben, vgl. hierzu Patrick SAHLE: *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*, Teil 3: Textbegriffe und Recodierung (Schriften des Instituts für Dokumentologie und Editorik 9), Norderstedt 2013, S. 244f.

<sup>8</sup> Vom „TEI-Paradox“ spricht deshalb auch Desmond SCHMIDT: *Towards an Interoperable Digital Scholarly Edition*, in: *Journal of the Text Encoding Initiative* 7 (2014), par. 12f.

<sup>9</sup> Vgl. dazu ‚Ziel und Fokus des DTA-Basisformats‘ der Einleitung unter: <http://www.deutschestextarchiv.de/doku/basisformat/ziel.html>. Dank unserer Kooperation mit dem DTABf konnten bereits bestimmte, für uns notwendige Elemente ergänzt werden (etwa ‚#in‘ zur Beschreibung von Initialen).

Textaufbereitung [zu] ermöglichen und dabei gleichzeitig Variationsspielräume bei der Annotation so [einzuschränken], dass die Kohärenz der DTA-Texte untereinander gewährleistet wird“.<sup>10</sup>

Dort, wo das DTABf nicht für die Anforderungen unserer Textzeugen ausreicht, greifen wir wieder auf die TEI zurück. Sofern möglich, bemühen wir uns um die Übernahme von Codierungen, die bereits durch andere mediävistische Editionsprojekte in den Diskurs eingebracht wurden.<sup>11</sup> Auf *tei.extensions* versuchen auch wir vollständig zu verzichten. Das verfolgte Konzept bei der Datenerfassung ist demnach reproduktionsorientiert: Die erhobenen Daten sollen möglichst zugleich menschen- als auch maschinenlesbar sein, um sowohl den Austausch und die Zusammenarbeit mit anderen (mediävistischen) Editionsprojekten wie auch die Interoperabilität mit verschiedenen Applikationen zu gewährleisten, wobei stets versucht wird, soweit wie möglich ‚standardnah‘ zu arbeiten.<sup>12</sup> Letzteres schließt die Nutzung bereits vorhandener Software mit ein.<sup>13</sup>

Als digitale Arbeitsumgebung wird von uns *ediarum.MEDIAEVUM* genutzt, das auf *ediarum* basiert, einer aus mehreren Softwarekomponenten entwickelten Arbeitsumgebung des Digital Humanities-Referats TELOTA – IT/ DH der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW).<sup>14</sup> In Zusammenarbeit mit Nadine Arndt vom *Otto von*

---

<sup>10</sup> Zitiert aus dem Abschnitt ‚Ziel und Fokus des DTA-Basisformats‘ der Einleitung unter: <http://www.deutschestextarchiv.de/doku/basisformat/ziel.html>.

<sup>11</sup> Insbesondere zu nennen ist hier die von ŠIMEK hervorragend aufbereitete Dokumentation ‚Welscher Gast digital. TEI Handbuch‘, vgl. Anm. 6.

<sup>12</sup> Siehe dazu auch Martin HOLMES: Whatever happened to Interchange?, in: *Digital Scholarship in the Humanities* 32.1 (2017), insb. par. 3f.

<sup>13</sup> Um zu verhindern, doch wieder zu einer Einzellösung zu kommen und sich damit vom Standard zu entfernen, ist die Evaluation der eingesetzten Software notwendig: Wie hoch ist der Zeitfaktor für die Einarbeitung? Wie sind Interoperabilität und Performance? Welche Abhängigkeiten entstehen, respektive, sind zu bedenken? Ist die Software auch unter Berücksichtigung des zu erwartenden technischen Fortschritts noch tauglich? Auch die lange Laufzeit des Projekts muss berücksichtigt werden; heute schon befassen wir uns mit Fragen der Langzeitdatensicherung und der Pflege der digitalen Editionsplattform nach Abschluss des Projekts, worauf wir bislang aber keine Antwort haben.

<sup>14</sup> Vgl. <http://www.bbaw.de/bbaw-digital/telota>.

*Passau-Projekt*<sup>15</sup> und TELOTA wurde diese Arbeitsumgebung so modifiziert, dass sie den spezifischen Anforderungen gerecht wird, die die Edition umfangreicher und breit überlieferter mittelalterlicher Prosatexte stellen.<sup>16</sup>

Die Arbeitsumgebung erleichtert das Erstellen des kritischen Textes dahingehend, dass die BearbeiterInnen in *oXygen XML Author* nicht in einer Codeansicht, sondern in einer benutzerfreundlichen ‚Autorenansicht‘ arbeiten, die über CSS gestaltet wird.<sup>17</sup> Den BearbeiterInnen stehen dabei mehrere Ansichten zur Auswahl, so dass per Mausclick die für den Arbeitsschritt geeignetste ausgewählt werden kann. Außerdem können die BearbeiterInnen über eine eigene Werkzeuggestreife per Knopfdruck Auszeichnungen vornehmen. So können beispielsweise Einträge in die vier Apparate mit der entsprechenden TEI-Auszeichnung versehen werden. Quellenverweise lassen sich über eine komfortable Auswahlliste mit dem jeweiligen Eintrag innerhalb des bibliographischen Registers in der zentralen XML-Datenbank eXist-db verknüpfen. Darüber hinaus kann, sofern vorhanden, eine Verknüpfung zum Eintrag im Wörterbuchnetz vorgenommen werden. Der gesamte Text kann dadurch einfach und schnell mit TEI-konformer XML ausgezeichnet werden.

Die XML-Datenbank eXist-db dient in der digitalen Arbeitsumgebung als zentrales Repositorium für die XML-Dokumente. Die Datenbank ist auf einem Server installiert und online global mit entsprechenden Zugangsdaten zugänglich. Dadurch können alle ProjektmitarbeiterInnen auf ein und denselben Datenbestand zugreifen und zusammenarbeiten. Zusätzlich besteht die Möglichkeit, auf der Basis von eXist-db, XQuery und XSLT eine Website zu erstellen. Vor der tatsächlichen Publikation der digitalen Edition im Web lässt sich hierdurch der aktuelle Datenbestand leicht durchblättern bzw. durchsuchen. Ebenfalls möglich ist die Ausgabe der aktuellen TEI-XML-Dokumente als PDF über den TEI-Publisher.

---

<sup>15</sup> <http://www.bbaw.de/forschung/otto-von-passau>.

<sup>16</sup> <https://git.rz.uni-augsburg.de/dslma-gitlab/ediarum.mediaevum>.

<sup>17</sup> Vgl. hierzu und im Folgenden die Beschreibung von *ediarum* unter: <http://www.ediarum.org> mit weiterer Literatur.

Theoretisch böte *ediarum.MEDIAEVUM* also eine Arbeitsumgebung, in der von der Transkription der Textzeugen über die Herstellung des kritischen Textes bis zur Publikation der Edition in Print und Web alles möglich ist. Allerdings stand *ediarum.MEDIAEVUM* in den ersten Jahren des Projekts noch nicht zur Verfügung. Die Gründe für diese Verzögerung liegen darin, dass unsere IT-Stelle erst mit großer Verspätung besetzt werden konnte und die Anpassung von *ediarum* zu *ediarum.MEDIAEVUM* deutlich komplexer und damit zeitintensiver ist als erwartet. Zudem gibt es Arbeitsschritte, die mit *ediarum.MEDIAEVUM* nicht (Kollationen) oder mit anderen Tools auf andere Weise besser (Transkription, PDF-Erstellung) bewerkstelligt werden können. Da es also keine ‚End-to-End‘-Lösung für die Erfassung unserer Primärdaten, ihre Aufbereitung und Publikation auf der von uns angestrebten digitalen Editionsplattform gibt<sup>18</sup> und wir mit dem Transkribieren und Edieren nicht warten konnten, bis sie fertiggestellt ist, sind wir auf eine Kombination verschiedener Softwarekomponenten angewiesen. Ein Nachteil des Einsatzes mehrerer Tools und Plattformen ist sicherlich, dass die Kenntnisse zum Arbeiten mit ihnen und die Fähigkeit zu ihrer Verknüpfung vorliegen müssen.<sup>19</sup> Diese Verknüpfung der verschiedenen Tools ist zwar aufwendig, aber immerhin ohne größeren Datenverlust möglich. Ein Vorteil ist, dass wir so das bereits vorhandene technische Know-how (TUSTEP in Berlin, XML in Augsburg) bestmöglich einsetzen können. Es ergibt sich folgender Workflow:

---

<sup>18</sup> Vgl. Zu dieser Problematik: Bridget ALMAS/Emad KHAZRAEE/Matthew Thomas MILLER/Joshua WESTGARD: Manuscript Study in Digital Spaces: The State of the Field and New Ways Forward, *Digital Humanities Quarterly* 12/2 (2018), par. 24.

<sup>19</sup> Siehe dazu ALMAS u. a.: Manuscript Study, par. 5.

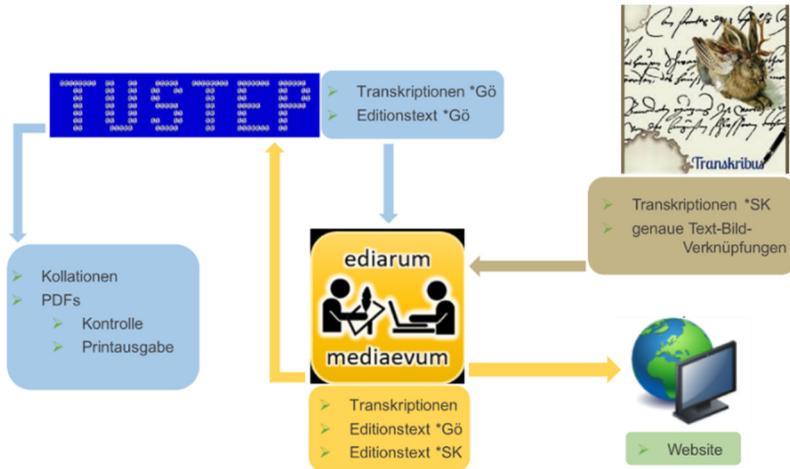


Abb. 1: Workflow Edition ‚Evangelienwerk‘

Historisch bedingt kommt in der Berliner Arbeitsstelle das *Tübinger System von Textverarbeitungsprogrammen* (TUSTEP) für die Erstellung der Transkriptionstexte wie des Editionstextes zum Einsatz.<sup>20</sup> TUSTEP hat sich in der Arbeitsstelle seit langem für die Aufbereitung von Textdaten für die Printedition in der Reihe ‚Deutsche Texte des Mittelalters‘ bewährt. Für eine Reihe wichtiger Arbeitsschritte (z.B. u/v-Ausgleich) liegen dementsprechend erprobte Routinen vor. TUSTEP bietet die Möglichkeit, die Materialität des handschriftlichen Textzeugens in der Transkription zu codieren, die Textzeugen zu kollationieren, den kritischen Text herzustellen sowie Transkriptionen und Edition in einem ansprechenden Layout auszudrucken.

Damit ermöglicht TUSTEP also einen in sich geschlossenen Workflow, mit dem Vorteil, dass NutzerInnen lediglich dieses eine Tool bedienen können müssen. Das in TUSTEP individuell gestaltbare Markup wurde für die Auszeichnung der Transkriptionen und des Editionstextes dabei so gewählt, dass es sich möglichst reibungslos in das für *ediarum.MEDIAEVUM* benötigte TEI/XML konvertieren lässt. Angepasst

<sup>20</sup> Mehr Informationen zu TUSTEP unter: <https://www.tustep.uni-tuebingen.de>.

werden müssen dort noch eine Reihe von Attributen, deren Eingabe in TUSTEP zu aufwendig wäre, ein Teil der kodikologischen Daten, für die in *ediarum.MEDIAEVUM* spezielle Eingabemöglichkeiten vorhanden sind, die eine einheitliche Aufnahme unterstützen, sowie die Verlinkungen zu den Registern und zu externen Ressourcen.

Selbst nach der Einspeisung aller Transkriptionen und des Editions-textes in *ediarum.MEDIAEVUM* wird TUSTEP nicht obsolet sein: Es bietet die einzig adäquate Möglichkeit, die Transkriptionen der verschiedenen Textzeugen zu kollationieren sowie den kritischen Text und seine vier Apparate in einer visuell ansprechenden Form zu setzen. Letzteres ist sowohl für Prüfgänge im Arbeitsablauf als auch für die Herstellung der Druckvorlage für die Printausgabe nötig.

In der Augsburger Arbeitsgruppe werden die Transkriptionen der Textzeugen der Fassung \*SK mit Hilfe der gleichnamigen Softwarekomponente der Plattform *Transkribus* von den studentischen und wissenschaftlichen Hilfskräften erstellt. *Transkribus* wurde zur Erkennung und Transkription historischer Dokumente konzipiert und stellt eine Reihe von Werkzeugen für die automatisierte Erfassung von Dokumenten zur Verfügung, darunter eine computergestützte Handschriftenerkennung (HTR), Bilderkennung (Layout Analysis) und Strukturerkennung (Document Understanding).<sup>21</sup>

Mit der lokal zu installierenden Expertensoftware werden in der leicht zu bedienenden Nutzeroberfläche die Transkriptionen der Handschriften zunächst von den Hilfskräften eingegeben und mit dem in der codexbasierten Perspektive angegebenen Markup versehen. Etwa 100 transkribierte Seiten bilden die Grundlage für den erstellten Algorithmus, der nachfolgend das automatische Auslesen der Transkriptionen, ähnlich der OCR für gedruckte Texte, ermöglicht. Die maschinell ausgelesene Transkription wird von einer Hilfskraft kontrolliert und ggf. verbessert. Zugleich wird die Annotation vorgenommen. Die hierbei erzielten Ergebnisse sind gut, die Fehlerrate bei einer automatisch erstellten

---

<sup>21</sup> Die Plattform *Transkribus* ist derzeit für alle BenutzerInnen frei zugänglich und wird von der Universität Innsbruck bzw. der Gruppe ‚Digitalisierung und Elektronische Archivierung‘ (DEA) betrieben, <https://transkribus.eu/Transkribus/>.

Transkription liegt mittlerweile bei durchschnittlich 4%, was den notwendigen Korrekturvorgang durch die BearbeiterInnen bei gleichzeitig erfolgreicher Kodierung der Transkription deutlich weniger zeitaufwendig macht als das händische Transkribieren mit nachfolgend einzufügender Annotation in XML. Zudem lässt sich bei der maschinell ausgegebenen Transkription aufgrund des DeepLearnings über die immer umfangreicher zur Verfügung stehende Datenmenge bislang eine stetige Verbesserung feststellen. Aus *Transkribus* lassen sich zahlreiche Datenformate ausgeben, u.a. auch XML-Dateien. Die aus *Transkribus* transformierte Transkriptionsdatei, die ‚tT‘, wird, aufgesplittet in Pakete von 25 Seiten der Transkription, in der eXistDB abgespeichert. Diese Aufteilung ist notwendig, da durch die genauen Text-Bild-Verknüpfungen XML-Dateien mit einer Größe von ca. 500 MB entstehen, die die Performance von *oXygen* zu stark einschränken würden. Die hier erstellten Transkriptionen bieten die Grundlage für die Texterstellung der Fassung \*SK. Ein für uns wichtiger Faktor ist zudem die zeilengenaue Verknüpfung von Digitalisat und Transkription, deren Koordinaten in den Dateien später für die Webpräsentation genutzt werden.

Die mit TUSTEP erstellten Transkriptionen der Überlieferungszeugen der Fassung \*Gö werden in der digitalen Edition dagegen nur spaltenweise mit den jeweiligen Digitalisaten verknüpft, was den NutzerInnen aber eine ausreichende Orientierung zwischen Digitalisat und Transkription ermöglichen wird. Der Mehraufwand durch die (erneute) Eingabe der bereits in TUSTEP erstellten Transkriptionen in *Transkribus* stünde in keinem Verhältnis zum Nutzen und würde zudem den vorgegebenen Zeitplan sprengen.<sup>22</sup>

Der in *ediarum.MEDIAEVUM* hergestellte kritische Text basiert für \*SK auf der transformierten Transkriptionsdatei der Leithandschrift, der ‚tT‘. Als ‚Aussagen über den handschriftlichen Text‘ sind die

---

<sup>22</sup> Als *Transkribus* als Tool für \*SK ins Gespräch kam, war für die Erstfassung die einzige Vollhandschrift Gö bereits vollständig, die Mitüberlieferung teilweise transkribiert. Die kleinteilige Exzerpt- und Fragmentüberlieferung böte zudem nicht genug Übungsmaterial für einen Algorithmus, so dass die automatische Handschriftenerkennung hier nicht zum Einsatz kommen könnte.

Annotationen der den handschriftlichen Text strukturierenden Mittel wie Initialen, Majuskeln oder Paragraphen- und Verweiszeichen zwar wichtige Anhaltspunkte für die Herstellung des kritischen Textes, werden aber in Codierung hier nicht benötigt. Über ein XSLT-Script werden diese Angaben in <comments> ausgelagert, Abkürzungen aufgelöst und geringfügige Normalisierungen auf Zeichenbasis vorgenommen; zurück bleibt die ‚transformierte Editionstranskription‘, die ‚tE‘, die den BearbeiterInnen das Arbeiten in werkbasierter Perspektive erleichtert, ohne dass dabei die technische Verknüpfung zu Transkriptions- und Bilddatei verloren ginge. Die tE stellt lediglich eine Zwischenstufe in eben genannter Funktion dar; aus ihr wird der betreffende Kapiteltext herauskopiert und als eigenständige Kapiteldatei abgespeichert. Das Aufsplitten des ODD in verschiedene Datenpakete sichert, indem darüber Verschachtelungen von verschiedenen Strukturen innerhalb eines Dokuments vermieden werden, die technische Funktionalität der Dateien ebenso wie es die leichtere Bearbeitung über verschiedene Perspektiven auf den Text ermöglicht.<sup>23</sup> Analog wurde für \*Gö die Transkription der Leithandschrift Gö in TUSTEP durch eine Austauschroutine in eine der tE entsprechende Datei verwandelt, in der die weitere Bearbeitung als Editionstext erfolgt. Die technischen Verknüpfungen mit dem Digitalisat und der Transkription werden hier nachträglich in *ediarum.MEDIAEVUM* eingefügt und erfolgen seiten- bzw. spaltenweise.

Dem kritischen Text des ‚Evangelienwerks‘ werden vier Apparate beigegeben: In Apparat 1 werden Sofortkorrekturen der Leithandschrift dokumentiert sowie alle Eingriffe gegen die Leithandschrift erläutert, ggf. mit stützenden Lesarten. Daneben werden Layoutphänomene der Leithandschrift und sinntragende Varianz der Parallelüberlieferung der eigenen Fassung vermerkt. Die genaue Verknüpfung des kritischen Textes mit den Transkriptionen der weiteren Textzeugen wird über die Auswahl der betreffenden Textstelle aus dem ‚Lineregister‘ aufrechterhalten. In dieses werden im Fall der in *Transkribus* erfassten Handschriften und Fragmente die aus den Transkriptionen isolierten ‚lineIDs‘ gespeichert,

---

<sup>23</sup> Zum Vorteil dieses Vorgehens siehe SCHMIDT: Towards an Interoperable Digital Scholarly Edition, par. 21f.

die die zeilengenaue Verknüpfung zwischen Digitalisat und Transkription gewährleisten. Im Fall der in TUSTEP erstellten Transkriptionen werden die Gliederungselemente <pb>, <cb>, <lb>, das heißt das Markup für Seiten- und Spaltenwechsel sowie Zeilenumbruch, genutzt.

Die Worterklärungen und Verständnishilfen in Apparat 2 werden zweifach vermerkt, indem die Lemma-Einträge mit den Online-Ausgaben der Wörterbücher sowie mit ihren gedruckten Ausgaben verknüpft werden.

Übersetzte Bibelstellen und die komplexe Verwendung weiterer Quellen durch den Österreichischen Bibelübersetzer werden in Apparat 3 nachgewiesen; Apparat 4 enthält Selbstzitate und Querverweise. Die über ein Register gesteuerte Eingabe der von den BearbeiterInnen identifizierten Quellen sichert dabei sowohl die Gültigkeit der Daten als auch die Möglichkeit unterschiedlicher Ausgabeformen, wie eine Bibliographie für den Druck oder die Durchsuchbarkeit des Textes in der digitalen Ausgabe. Die Auszeichnung von Personen und Orten erfolgt unter Verwendung der Gemeinsamen Normdatei.

Diese Form der Codierung verknüpft so nicht allein die verschiedenen Bestandteile der Edition miteinander, sondern öffnet die Edition nach ihrer Veröffentlichung auch für das Semantic Web. Die digitale Edition verfügt damit über einen beträchtlichen Mehrwert. Vergessen werden darf aber nicht, dass dieser Mehrwert nur durch einen deutlich höheren Aufwand als bei einer reinen Printedition zu erreichen ist.

Unsere Edition des Gesamtwerks des Österreichischen Bibelübersetzers wird Werk und Wirkung jenes Unbekannten nicht allein der Forschungsgemeinschaft, sondern auch interessierten Laien zur Verfügung stellen und so nach mehr als 600 Jahren das tun, was der Bibelübersetzer stets im Sinn hatte: die Bibel für alle zugänglich zu machen.<sup>24</sup> Derzeit

---

<sup>24</sup> Möglichst viele Zugänge und Fragestellungen sollen unsere Daten ermöglichen; klare Einschränkungen lassen sich aber schon jetzt benennen: Völlig zu Recht wird gefordert, das sprachhistorische Interesse an edierten (mittelalterlichen) Texten zu berücksichtigen, siehe hierzu etwa Klaus-Peter WEGERA: Vnd machet sie mit gesehenden augen blind. Zum Problem von Editionen als Datenquelle für sprachhistorische Untersuchungen, in: Zeitschrift für deutsche Philologie 134 (2015), S. 77–95; sowie die Beiträge in: Vom Nutzen der Editionen. Zur Bedeutung moderner Editorik für die Erforschung von Literatur- und

befindet sich die Online-Visualisierung der erarbeiteten Inhalte in der Planung; während für die Print-Publikation der beiden Fassungen Texte Erscheinungsform und Layout klar sind, stellen uns die veränderten Lese- und Wahrnehmungsmöglichkeiten im neuen Medium vor ganz eigene Herausforderungen.

---

Kulturgeschichte (Beihefte zu editio 39), hrsg. v. Thomas Bein, Berlin 2015. Die hierfür notwendigen Angaben zu Wortgrenzen sowie Wortannotationen und -disambiguierungen innerhalb des Textes sind im zeitlichen Rahmen unseres Vorhabens jedoch nicht leistbar.

## Bibliographische Hinweise

Alle Internetlinks wurden am 10.06.2020 überprüft.

### Forschungsliteratur

- ALMAS, Bridget/KHAZRAEE, Emad/MILLER, Matthew Thomas/WESTGARD, Joshua: Manuscript Study in Digital Spaces. The State of the Field and New Ways Forward, in: *Digital Humanities Quarterly* 12.2 (2018); <http://www.digitalhumanities.org/dhq/vol/12/2/index.html>.
- BEIN, Thomas (Hrsg.): *Vom Nutzen der Editionen. Zur Bedeutung moderner Editorik für die Erforschung von Literatur- und Kulturgeschichte* (Beihefte zu editio 39), Berlin 2015.
- BEIN, Thomas: Leerstellen edieren? Überlegungen zur Einbindung von Performanz in Editionen mittelalterlicher Literatur, in: *editio* 32 (2018), S. 82–92.
- HOLMES, Martin: Whatever happened to Interchange?, in: *Digital Scholarship in the Humanities* 32.1 (2017); doi.org/10.1093/lc/fqw048.
- KORN RUMPF, Gisela: Das ‚Klosterneuburger Evangelienwerk‘ des österreichischen Anonymus. Datierung, neue Überlieferung, Originalfassung, in: *Deutsche Bibelübersetzungen des Mittelalters. Beiträge eines Kolloquiums im Deutschen Bibel-Archiv*, hrsg. v. Heimo Reintzer unt. Mitarb. v. Nikolaus Henkel (Vestigia Bibliae 9/10 [1987/1988]), Bern 1991, S. 115–131.
- LÖSER, Freimut/STÖLLINGER-LÖSER Christine: Verteidigung der Laienbibel. Zwei programmatische Vorreden des österreichischen Bibelübersetzers der ersten Hälfte des 14. Jahrhunderts, in: *Überlieferungsgeschichtliche Editionen und Studien zur deutschen Literatur des Mittelalters (Texte und Textgeschichte 31)*, hrsg. v. Konrad Kunze/Johannes G. Mayer/Bernhard Schnell, Tübingen 1989, S. 245–313.
- SAHLE, Patrick: *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*, Teil 3: Textbegriffe und Recodierung (Schriften des Instituts für Dokumentologie und Editorik 9), Norderstedt 2013.
- SCHMIDT, Desmond: Towards an Interoperable Digital Scholarly Edition, in: *Journal of the Text Encoding Initiative* 7 (2014); doi.org/10.4000/jtei.979.
- WEGERA, Klaus-Peter: Vnd machet sie mit gesehenden augen blind. Zum Problem von Editionen als Datenquelle für sprachhistorische Untersuchungen, in: *Zeitschrift für deutsche Philologie* 134 (2015), S. 77–95.

### Internetlinks

- Deutsches Textarchiv-Basisformat;  
<http://www.deutschestextarchiv.de>.
- Digitale Arbeitsumgebung ediarum.MEDIAEVUM:  
<https://git.rz.uni-augsburg.de/dslma-gitlab/ediarum.mediaevum>.
- ediarum (TELOTA, Berlin-Brandenburgische Akademie der Wissenschaften);  
<http://www.ediarum.org>.
- ‚Lyrik des deutschen Mittelalters. Digitale Edition‘;  
<http://ldm-digital.de>.

,Otto von Passau: Die 24 Alten';

<http://www.bbaw.de/forschung/otto-von-passau>.

ŠIMEK, Jakub: Welscher Gast digital. TEI Handbuch (Version 0.6);

[https://digi.ub.uni-heidelberg.de/wgd/TEI-Handbuch\\_0-6.pdf](https://digi.ub.uni-heidelberg.de/wgd/TEI-Handbuch_0-6.pdf).

TELOTA;

[http://www.bbaw.de/telota/ueber\\_uns](http://www.bbaw.de/telota/ueber_uns).

Transkribus;

<https://transkribus.eu/Transkribus>.

TUSTEP;

<https://www.tustep.uni-tuebingen.de>.

,Welscher Gast digital';

<https://digi.ub.uni-heidelberg.de/wgd/>.