



ASTRID BÖHM / HELMUT W. KLUG

Quellenorientierte Aufbereitung historischer Texte im Rahmen digitaler Editionen

Das Problem der Transkription in mediävistischen Editionsprojekten¹

I. Einleitung

Zieht man die Ausstattung und die Arbeitsweise an einem modernen geisteswissenschaftlichen Arbeitsplatz in Betracht, so sind Editionen heute zwangsläufig digital. Alle weiteren Produkte dieser Arbeit wie diverse Analysen, Visualisierungen oder gar Druckerzeugnisse sind daher Derivate einer digitalen Datenbasis. Das digitale Medium erlaubt es, gesteigerte Ansprüche an die Aufbereitung einer historischen Quelle zu stellen, sodass damit Daten erzeugt werden, die auch einer Nachnutzung zugeführt werden können. Die dem zugrundeliegende Idee geht auf die Grazer Dynamische Editions-methode² zurück, in der besonderer Wert auf eine ultradiplomatistische Basistransliteration gelegt wird. Im Rahmen des bundesgeförderten Hochschulraumstrukturmittelpjekts ‚KONDE - Kompetenznetzwerk Digitale Edition‘ wurde diese Forderung als eine zentrale Zielsetzung formuliert:

[...] Entwicklung von Prototypen zur Texteingabe bei der Erfassung handschriftlicher (historischer) Dokumente in Abstimmung mit, aber auch in Erweiterung der geltenden Standards auf einer möglichst fein codierten Mikrotransliteration

¹ Der vorliegende Text wurde im Team abgefasst. Der Schwerpunkt der Vorarbeiten in Bezug auf die praktische Umsetzung beim Transkribieren und die Überlegungen zu den Transkriptionsrichtlinien wurde von Astrid Böhm geleistet; das Workflow-Konzept, die Zeichenmodellierung und die technische Umsetzung wurde von Helmut W. Klug bereitgestellt.

² Andrea HOFMEISTER-WINTER: Das Konzept einer ‚Dynamischen Edition‘ dargestellt an der Erstausgabe des „Brixner Dommesnerbuches“ von Veit Feichter (Mitte 16. Jh.). Theorie und praktische Umsetzung (Göppinger Arbeiten zur Germanistik 706), Göppinger 2003.

(mehrschichtig-dynamische Editionstechnik) und damit sowohl paläografische, text- als auch sprachanalytische Untersuchungen.³

Der vorliegende Aufsatz soll jenen Workflow präsentieren, der unter dieser Zielsetzung entstanden ist. Das Ziel der Transkriptionsleistung muss eine möglichst umfangreiche Basistranskription in TEI-XML sein, die als Forschungsdaten *open access* publiziert wird. Diese Forschungsdaten sollen nicht nur dem Editionsprojekt selbst in Form der detaillierten Quellenerschließung und Analyse, sondern auch beliebigen anderen Nachnutzern für die weitere Verarbeitung in den unterschiedlichsten Forschungsbereichen zur Verfügung stehen. Die theoretischen Überlegungen wurden in zwei weiteren Projekten, in denen aktive Transkriptionsarbeit geleistet wurde, intensiv getestet und weiterentwickelt: Im vom Land Steiermark geförderten Projekt *Etablierung eines Mittelalterlabors an der Universität Graz: Digitale, fachwissenschaftliche und wissenschaftskommunikatorische Aufbereitung der Handschrift Graz, UB, Ms. 1609* war eine Zielsetzung die Transkription der historischen Quelle einer rund 500 Blatt umfassenden, latein- und deutschsprachigen zusammengesetzten Handschrift.⁴ Im vom österreichischen Wissenschaftsfonds geförderten internationalen Kooperationsprojekt *Cooking Recipes of the Middle Ages: Corpus, Analysis, Visualisation (CoReMa)* müssen knapp 60 Kochrezepttextsammlungen, die ca. 880 Blatt umfassen, transkribiert werden.⁵

Die Feinheit der Transkription muss dabei von Textzeugen zu Textzeugen und abhängig von der eigentlichen Forschungsfrage⁶ individuell

³ ‚Projekt‘, in: KONDE – Kompetenznetzwerk Digitale Edition. HRSM-Projekt, hrsg. v. Zentrum für Informationsmodellierung – Austrian Centre for Digital Humanities, <http://www.digitale-edition.at>.

⁴ Mittelalterlabor. Transkription der Handschrift Graz, UB, Ms. 1609, hrsg. v. Helmut W. Klug unt. Mitarb. v. Astrid Böhm/Elisabeth Raunig, April 2019, <hdl.handle.net/11471/521.60> (GAMS 521.60).

⁵ Das Projekt *CoReMA* wird vom fwf unter der Nr. I 3614 gefördert und in Kooperation mit der Universität Tours (Prof. Bruno Lauriou, ANR-17-CE27-0019-01) durchgeführt. Die Projektwebsite wird ab Herbst 2019 sukzessive veröffentlicht: <http://gams.uni-graz.at/context:corema>.

⁶ Im *Mittelalterlabor* war das die intensive Quellenerschließung und die Zurverfügungstellung der Forschungsdaten. In *CoReMA* würde der normalisierte Lesetext für die

festgelegt werden. In den oben genannten Projekten steht aufgrund der umfangreichen Quellenlage die Erarbeitung eines gut argumentier- und umsetzbaren Minimalansatzes im Mittelpunkt – der aber für die Nachnutzung der Editionsdaten dennoch maximale Möglichkeiten bieten soll. Dieser Transkriptionsprozess ist begleitet von einer ‚dynamischen Modellierung‘ – diese Bezeichnung ist in Anlehnung an die eben genannte Editionsmethode gewählt – der Name veranschaulicht aber auch den Zugang zum Transkriptionsprozess als Ganzem: Da wir in der Regel mit vorerst unbekanntem Quellen der spätmittelalterlichen Fachliteratur arbeiten, verfeinert sich das zugrunde gelegte Modell mit jedem Transkriptions- und Korrekturdurchlauf. Das betrifft sowohl die Mikro- wie auch die Makrostruktur der Texte. Die Grundlage für eine erste Modellierung der vorhandenen Informationen einer Quelle ist eine stichprobenartige Auswertung,⁷ nach der ein Modell entworfen wird, das die Quelle so detailliert wie möglich abbildet: In einem ersten Schritt wird der Aufbau der Quelle – Blatt, Seiten und andere Gliederungselemente – abgebildet. Dieses Grundmodell wird mit einem mehr oder weniger fein annotierten Text und Anmerkungen zu Problemstellen gefüllt. Im Zuge der iterativen Arbeit mit dem Text werden die Anmerkungen sukzessive aufgelöst und das Modell ausgebaut bzw. um editorische Eingriffe erweitert.

II. Editionsmethode

Für die Transkription der Texte wurde das mehrstufige Editionsmodell der ‚Dynamischen Edition‘ von Andrea Hofmeister adaptiert. Ausgehend von einer gemeinsamen feingranularen Transkription des Quellentextes zielt das Konzept der Dynamischen Edition auf eine möglichst ‚breite‘ und vielfältige Nutzung der gewonnenen Erkenntnisse ab. Paläografische Interessen sollen dabei genauso berücksichtigt werden wie literaturwissenschaftliche Anliegen und sprachwissenschaftliche Fragestellungen. Ältere bzw. gedruckte Editionen unterliegen häufig gewissen medienbedingten Einschränkungen: Während PaläografInnen unterschiedliche

semantische Annotation ausreichen, dennoch werden Basistranskriptionen aller deutschsprachigen Quellen angefertigt!

⁷ Vgl. die ähnlichen Ausführungen zur 10%-Studie bei der Schriftbefundung unten.

Ausführungen von Glyphen Aufschluss über Entstehungszeit und -ort geben können, sind beispielsweise LiteraturwissenschaftlerInnen eher an einem ‚gut lesbaren‘ Text interessiert, mit dem fachspezifische Fragestellungen bearbeitet werden können; ein normalisierter Text ist jedoch unbrauchbar, wenn historiologische Auswertungen vorgenommen werden sollen. Die Liste der Einschränkungen kann beliebig fortgesetzt werden.⁸ Das Problem bzw. die unterschiedliche Herangehensweise an Texte ist also nichts Neues und wird in den Editionswissenschaften seit langem diskutiert.⁹ Welche paläografischen Informationen sind für wen wichtig? Welche können ohne größeren Verlust vereinheitlicht/vernachlässigt werden? Edierende stehen zu Beginn jedes Editionsprojektes vor der Aufgabe diese Fragen für die individuelle Arbeit zu lösen. Sahle hält dazu fest, dass es der Forschung bislang nicht gelungen ist, eine Definition „relevante[r] Informationen im Bereich der Orthografie zu geben oder eine eindeutige Grenze zwischen grafischen (ev. auch: kalligrafischen) Zufälligkeiten und semiotischen Bedeutungsträgern zu ziehen.“¹⁰ Dieses Problem lässt sich auch mit den vorliegenden Ausführungen nicht lösen, aber die Projekte übergreifende Zusammenfassung bietet zumindest eine Möglichkeit, Teile der vielfältigen Anliegen, die an die Edition gestellt werden, zu bedienen und individuelle Lösungen zu präsentieren, die durchaus auf eine generische Anwendung abzielen. Die Mehrstufigkeit der Dynamischen Edition¹¹ ist die ideelle Grundlage des beschriebenen Workflows.

⁸ Vgl. Karin KRANICH-HOFBAUER: Editionswissenschaften als interdisziplinäre Grundwissenschaft. Über Berührungspunkte und deren Überwindung, in: Jahrbuch der Oswald von Wolkenstein Gesellschaft 12 (2000), S. 49–64, hier: S. 53f.

⁹ Vgl. als Überblick Patrick SAHLE: Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, Teil 1: Das typografische Erbe, Nordstedt 2013, S. 143–167.

¹⁰ SAHLE: Digitale Editionsformen, S. 72.

¹¹ Vgl. HOFMEISTER-WINTER: Dynamische Edition, S. 110–117.

Den Kern des Modells bildet eine ‚deskriptive Transliteration‘¹² oder auch ‚Basistransliteration‘¹³. Diese erste Editionsstufe¹⁴ bietet eine möglichst genaue Dokumentation des handschriftlichen Befundes mit möglichst geringer Interpretation, hier findet sich das Maximum an paläografischen Informationen. Diese sind in einer TEI-XML-Wiedergabe der Quelle abgebildet. In weiteren Transformationsstufen werden die gewonnenen Informationen systematisch durch grafetische Analyse reduziert, um zu einer philologisch nutzbaren ‚Lesefassung‘ zu gelangen, die unterschiedliche Eingriffe bezüglich der Vereinheitlichung der Grafie aufweist. Der Grad der Normalisierung hängt von der Intention des Editors / der Editorin ab, bzw. welchem NutzerInnenkreis und welchen Fragestellungen der Text dienen soll. Durch die aufeinander aufbauenden Editionsstufen kommt es trotz der Reduktion zu keinem Informationsverlust, sie ermöglichen vielmehr die Beantwortung unterschiedlicher Fragestellungen.

Im Transkriptionsprojekt *Mittelalterlabor* wurde das mehrstufige Konzept der Dynamischen Edition verfeinert und in seiner Form im Prinzip auch für das Projekt *CoReMA* beibehalten: Das Grundanliegen ist dabei, möglichst umfangreiche paläografische und linguistisch relevante Informationen aus der Handschrift zu dokumentieren, um damit einen Arbeitsschritt zu leisten, auf dem auch andere Forschende aufbauen können. Die Umsetzung wurde dabei den veränderten technischen und wissenschaftsrelevanten Bedingungen angepasst.¹⁵

Die Quelldaten werden in TEI-XML modelliert zur Verfügung gestellt. Die Komplexität der Annotation macht eine direkte Transkription der Texte in ein XML-Dokument überaus mühsam und darüber hinaus

¹² Wolfgang FEIGS: Deskriptive Edition auf Allograph-, Wort- und Satzniveau, demonstriert an handschriftlich überlieferten, deutschsprachigen Briefen von H. Steffens, Teil 1: Methode, Bern 1979.

¹³ Vgl. HOFMEISTER-WINTER: Dynamische Edition, S. 101.

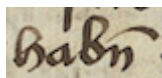
¹⁴ Zur detaillierten Beschreibung der einzelnen Stufen vgl. HOFMEISTER-WINTER: Dynamische Edition, S. 100–109.

¹⁵ Das ursprüngliche Modell berücksichtigt die Möglichkeiten der digitalen Edition noch nicht.

höchst fehleranfällig. Eine grundlegende Fragestellung war es daher, einen möglichst intuitiven, aber dennoch effektiven Weg zu finden, auch umfangreiche Textkorpora zu transkribieren.

Die Basistransliteration erfolgt mittels eines proprietären Markups, das aus alphanumerischen Codes gebildet wird. Generell soll versucht werden, den Text ausschließlich mithilfe von Zeichen zu codieren/transkribieren, die in der ASCII-Zeichentabelle vorhanden sind. Die konkrete Codierung ist u. a. einem pragmatischen Ansatz geschuldet, denn gearbeitet wird ausschließlich mit den auf einer Standardtastatur verfügbaren Zeichen. Gerade bei der Transkription von großen Textmengen ist jeder zusätzliche Tastenanschlag bzw. Mausklick zu vermeiden, um den Transkriptions-Prozess so effizient wie möglich zu gestalten. Die in diesem Fall gewählte Codierung ist eine Mischung aus konkreter (ikonischer) und abstrakter (symbolischer) Codierung. Für die Verwendung konkreter Zeichen spricht die einfache Assoziation, die sich aus der Ähnlichkeit zwischen gewähltem Code und den real existierenden Zeichen des Textes ergibt.

Als Beispiel sei hier auf eine der am häufigsten verwendeten Abbrüviaturen in deutschsprachigen Texten des Spätmittelalters verwiesen, den sogenannten ‚Nasalstrich‘. Der Nasalstrich – oder um seine Funktion im historischen Kontext besser zu beschreiben, da dieses Abbrüviaturzeichen nicht nur bei gekürzten Nasalen zu finden ist, – der ‚unbestimmte Kürzungsstrich‘, steht in der Regel über Vokalen und kann die Kürzungen unterschiedlicher Buchstaben anzeigen. Meist handelt es sich um ein ‚fehlendes‘ *m*, *n*, *e* oder *d*, aber auch Kombinationen wie *en*, *em*, *de* usw. Für die Codierung des konkreten Abbrüviaturzeichens wird in unserem Workflow das Minussymbol (-) verwendet, die gekürzten Buchstaben werden dann mittels einer Ziffer bestimmt. Die Codierung des Wortes *haben*, in dem das *e* mittels Kürzungsstrich gekennzeichnet ist, wird dann wie in Abbildung 1 dargestellt.



hab-1n

Abb. 1: Synoptische Gegenüberstellung des Wortes *haben* als Ausschnitt aus dem digitalen Faksimile und als transkribierter Text mit proprietärem Markup

Bedingt dadurch, dass ein und dasselbe Kürzungszeichen für unterschiedliche Kürzungen stehen kann, benötigt es die abstrakte numerische Codierung, um auf den oder die exakten Buchstaben zu verweisen. Aufgrund der leichteren Memorierbarkeit der Codierung wurde soweit wie möglich mit dem Gleichheitsprinzip gearbeitet. Der proprietäre Code hat sich als eine effektive Arbeitsmethode bewährt, die es ermöglicht, auch umfangreiche Texte, die von Hand transkribiert werden, relativ ‚schnell‘ zu erfassen. Als Einschränkung gilt allerdings, dass die Quelle ein überschaubar eingeschränktes Zeichensonderinventar aufweisen sollte: Deutschsprachige spätmittelalterliche Fachtexte können damit jedenfalls gut bearbeitet werden. Das Problem ist in diesem Fall nicht die generelle Konzeption des Workflows oder die technische Umsetzung, sondern der involvierte menschliche Faktor, denn für z. B. lateinische Texte, die eine überdurchschnittlich große Anzahl an unterschiedlichen Abbrueviaturzeichen aufweisen, ist das vorgestellte System des proprietären Markups mit alphanumerischen Codes nicht geeignet. Ähnliche Ansätze finden sich auch in anderen Projekten, wie etwa bei *Lyrik des deutschen Mittelalters*.¹⁶ Die EditorInnen haben aufgrund des Ausgangsmaterials (deutschsprachige Texte des 12. bis 14. Jhs.) und dem Anliegen mehrere ‚Textebenen‘ zu erfassen mit der gleichen Problematik zu kämpfen.¹⁷

Jegliches Zeicheninventar, das im ersten Schritt über das proprietäre Markup erfasst wurde, wird mithilfe der TEI innerhalb der <encoding-Desc> im Detail beschrieben und über das <g>-Element in den Fließtext des TEI-XML eingebunden. Mittels des TEI-XML Textmodells wird die historische Quelle möglichst deskriptiv dargestellt:¹⁸ Seiten-, Spalten- und

¹⁶ *Lyrik des deutschen Mittelalters*, hrsg. v. Manuel Braun/Sonja Glauch/Florian Kragl, <http://www.ldm-digital.de>.

¹⁷ Näheres dazu findet sich bei Florian KRAGL: Pseudo-Markup. Eine Eselsbrücke zwischen manueller und maschineller Textverarbeitung, in: *Magazin für Editionswissenschaften* 1 (2015), S. 7–14.

¹⁸ Vgl. dazu die detaillierten Ausführungen zu den verwendeten TEI-Elementen, Attributen und Attributwerten in den Editionsrichtlinien zur Edition der Grazer Handschrift Ms. 1609: Editionsrichtlinien, in: *Mittelalterlabor. Transkription der Handschrift Graz, UB, Ms. 1609*, hrsg. v. Helmut W. Klug unt. Mitarb. v. Astrid Böhm/Karin Kranich/Elisabeth Raignig, April 2019, hdl.handle.net/11471/521.60 (GAMS. 521.60.).

Zeilenlayout werden als makrostrukturelle Elemente der Quelle erhalten. Als Elemente der Mikrostruktur werden Überschriften, Initialen, Rubrizierungen, Strichelungen, Unterstreichungen mit den entsprechenden TEI-Elementen und Attributwerten ausgezeichnet. Marginalen werden, wenn sie Teil des Haupttextes sind, ihrem Charakter nach annotiert (z. B. als Einfügung) und an der passenden Stelle in den Text integriert oder mit dem Element <note> kommentiert. Kustoden werden mit dem Element <note> gekennzeichnet, und mit einem beschreibenden Kommentar versehen. Im Zuge des CoReMA-Projekts werden auch semantische Einheiten (z. B. Rezepte) mit dem Element <seg> im Basistext annotiert. Auf Wort- und Grafenebene werden Revisionen, Abkürzungen sowie editorische Rekonstruktionen mit den entsprechenden TEI-Elementen modelliert. Der Quellentext wird zeichengetreu wiedergegeben, nicht lesbare oder unklare Passagen werden mit dem TEI-Element <unclear> gekennzeichnet.

Im Zuge der deskriptiven Basistranskription wird ein für die Schrift der Quelle typisches Zeichensonderinventar angelegt und in der <charDecl>, welche die TEI zur Beschreibung von Einzelzeichen anbietet, gesammelt. Diese dient neben der Systematisierung in erster Linie dazu, eine möglichst handschriftennahe/diplomatische Ansicht erzeugen zu können, die zur Kollationierung mit dem digitalen Faksimile im Zuge des internen Workflows dient. Darüber hinaus werden die Zeichen in der Character Declaration auf Basis eines defacto Community-Standards – der Medieval Unicode Font Initiative (MUFI)¹⁹ – beschrieben. Um ein vollständiges Zeichensonderinventar erstellen zu können, ist es mitunter auch erforderlich, Zeichen zu beschreiben, zu denen es bisher keine Entsprechung in der MUFI gibt. In derartigen Fällen wurden im ersten Schritt optisch ähnliche Unicodezeichen gewählt, um eine diplomatische Darstellung zu gewährleisten: Die aus der MUFI Character Recommendation in die Character Declaration einer historischen Schrift (hier: spätmittelalterliche Kursive) übernommenen Zeichen müssen auch nicht

¹⁹ MUFI character recommendation. Characters in the official Unicode Standard and in the Private Use Area for Medieval texts written in the Latin alphabet, Version 4.0, hrsg. v. Odd Einar Haugen, 22.12.2015, hdl.handle.net/1956/10699.

zwingend der in der MUFI assoziierten historischen Schrift entsprechen – im Vordergrund steht bei der Zeichenrepräsentation die optische Übereinstimmung. Eine distinkte Unterscheidbarkeit ist ohnehin durch die Annotation der Zeichen gegeben.

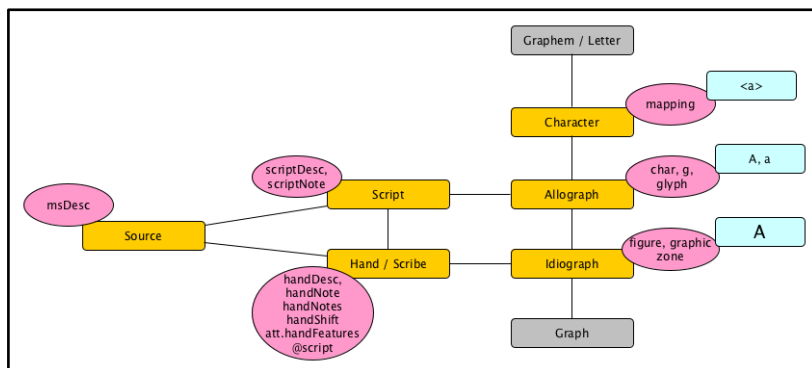


Abb. 2: Schematische Darstellung der Beschreibungsebenen von Zeichen in Relation zur Quellenbeschreibung

Die Sonderinventarliste gliedert sich in alphabetische Zeichen, Diakritika, Kürzungszeichen und Interpunktionszeichen und folgt bei der Hierarchisierung der Einträge weitestgehend dem Beschreibungsmodell von Handschrift, wie es im Rahmen des DigiPal-Projekts²⁰ entwickelt wurde.²¹ Abbildung 2 zeigt die einzelnen Beschreibungsebenen in einem hierarchischen Modell vom Graphem über dessen individuelle Ausformungen bis hin zum Graf als Teilzeichen. Die detaillierte Beschreibung der Handschrift wurde dabei in Relation zur Beschreibung der Quelle gestellt

²⁰ Peter A. STOKES: Describing Handwriting, Part IV: Recapitulation and Formal Model, in: Blog, DigiPal, 14.10.2011, <http://www.digipal.eu/blog/describing-handwriting-part-iv-recapitulation-and-formal-model>. Vgl. zur Diskussion um die Konzeption auch die anderen Teile dieser Post-Serie.

²¹ Es hat bei diesem Thema den Anschein, dass jede Disziplin ihr eigenes Vokabular und ihre eigene Hierarchie bildet, teilweise werden Begriffe in derselben Disziplin sogar unterschiedlich verwendet (vgl. dazu SAHLE: Editionsformen 1, S. 158f.): Dieser Prozess mag intellektuell herausfordernd sein, lösungsorientiert ist er nicht. Für hier gewählte Systematik stehen rein praktische Überlegungen im Zentrum.

und um mögliche Elemente der TEI, die dafür genutzt werden können, angereichert.

In der praktischen Umsetzung mit TEI-XML kann das Element `<charDecl>` als Kindelemente `<char>` haben, was dem Konzept des ‚Character‘ der Grafik entspricht; mithilfe des Elements `<glyph>` wird der ‚Allograf‘ näher beschrieben: So wird als ‚Character‘ das „LATIN SMALL LETTER I“ als `<i>` geführt, als Allograf das „LATIN SMALL LETTER DOTLESS I“ `<i>`. Die in der Grafik dargestellte Hierarchie kann aufgrund der Einschränkungen der TEI (`<char>`- und `<glyph>`-Elemente können nur auf der gleichen Ebene und nicht verschachtelt geschrieben werden) nur linear und nicht hierarchisch realisiert werden.

```
<char xml:id="i">
  <charName>LATIN SMALL LETTER I</charName>
  <mapping/>
</char>
<glyph      corresp="#i"          xml:id="inodot"          ana="allograph"
resp="https://bora.uib.no/handle/1956/10699" source="p.48">
  <glyphName>LATIN SMALL LETTER DOTLESS I</glyphName>
  <mapping type="normalized">i</mapping>
  <mapping type="transcription">i2</mapping>
  <mapping type="unicode_codepoint" subtype="LatExtA">0131</mapping>
  <mapping type="encoding" subtype="html_entity">&#305;</mapping>
  <mapping type="encoding" subtype="unicode_symbol">i</mapping>
</glyph>
```

Abb. 3: Code Snippet aus der Character Declaration: Beschrieben werden hier `<i>` mit und ohne Punkt-Superskript.

In den Einträgen zu den einzelnen individuellen Allografenbeschreibungen werden mithilfe von XML-Attributen und den entsprechenden Attributwerten folgende Informationen abgebildet (vgl. Abbildung 3): Der Wert des `@corresp`-Attributs weist auf eine mögliche hierarchische Verknüpfung hin. Jeder Eintrag erhält einen eindeutigen Identifikator, der auf den Entity Name der MUFI Character Recommendation zurückgeht. Jeder Eintrag wird systematisch beschrieben als Allograf, Abbrueviatur, Superskript oder Satzzeichen; Abbrueviaturen wiederum werden nochmals in allgemeines Kürzungszeichen, Brevigraf²² oder Kontraktion differenziert. Die Benennung und Beschreibung der Zeichen beruht in der Regel

²² Details zur Terminologie siehe unten.

auf der MUFI, deshalb wird mit entsprechenden Attributen (@resp, @source) auch auf diese Quelle mit Seitenreferenz verwiesen. Ist ein Zeichen oder eine Zeichenkombination dort nicht vorhanden, wird dessen/deren Name analog zu den Unicode-Namen aufgebaut. Das Element <mapping> gibt die Darstellung eines Zeichens in unterschiedlichen Kontexten an: Aufgenommen sind die jeweiligen Werte des Unicode Codepoint und für die Darstellung die HTML-Entity und das Uniconesymbol. Zusätzlich sind für die automatisierte Verarbeitung noch das proprietäre Markup und die normalisierte Zeichendarstellung hinterlegt. Zeichen und Zeichenkombinationen, die im ASCII-Zeichensatz vorhanden sind, werden nicht im Detail beschrieben.

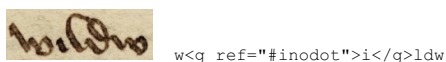


Abb. 4: Synoptische Gegenüberstellung der Phrase *wildw* (,willst du') als Ausschnitt aus dem digitalen Faksimile und als TEI-XML modellierter Text, indem das *i* ohne Punkt mit dem <g>-Element und dem @ref-Attribut annotiert ist.

Der damit teilweise bis auf Teilzeichenebene modellierte und annotierte Text (vgl. Abbildung 4) steht dann zur Weiterbe- und -verarbeitung bereit. Wenn eine Character Declaration einmal für alle auszuzeichnenden Zeichen angelegt worden ist, kann diese als nachnutzbare Ressource im Rahmen vergleichbarer historischer Schriften leicht für neue Projekte adaptiert werden: Zeichen können gelöscht oder hinzugefügt werden oder das proprietäre Markup kann auf die Vorlieben anderer Transkribierender angepasst werden.

III. Transkriptionsrichtlinien

Bei der Festlegung der Transkriptionsregeln galt es in einem ersten Schritt zu entscheiden, welche Informationen aus der Handschrift mit welcher Beschreibungstiefe übernommen werden sollen und können.²³ Die Basistransliteration soll prinzipiell so nahe wie möglich am historischen Text bleiben, das bedeutet einerseits, die Makrostruktur bezüglich der Seiten- und Zeilenumbrüche beizubehalten, andererseits werden

²³ Dieser Prozess konnte nur durch intensive Diskussion mit FachkollegInnen zum Ziel geführt werden, unser besonderer Dank gilt hier Dr. Karin Kranich für den wertvollen Input.

auch die Besonderheiten der historischen Schrift bezüglich einzelner Grafe berücksichtigt.

Um etwaige Schreiberspezifika zu dokumentieren, wurde von jedem Überlieferungsträger eine 10%-Studie sowie eine systematisierte Handbeschreibung angelegt, die eine paläografische Analyse beinhaltet. Dabei werden 10% des vorliegenden Textes vor Beginn der eigentlichen Transkription intensiv begutachtet. Basierend auf den Erkenntnissen dieser Analyse wird dann eine Schrift- bzw. Handbeschreibung erstellt, in der bestimmte Leitgrafe beschrieben werden, in der Besonderheiten, wie unübliche Bogenverbindungen oder Schaftlängen dokumentiert sind, in der genauere Aussagen zu Revisionsvorgängen getroffen werden (z. B. welche Techniken bei Tilgungen zum Einsatz kamen: Streichung, Bleichung, Rasur, Überklebung, Expungierung, ...) usw. Die Handbeschreibung dient dem genaueren Schriftbefund, der in der Transkription als solcher nicht dokumentiert ist.

Transkribiert werden außerdem alle Phänomene, die lautliche Unterschiede zur Folge haben (z. B. diakritische Zeichen) oder andere Bedeutungsunterschiede (beispielsweise Superskripte zur Unterscheidung von einzelnen Grafen n/u) ausmachen. Ebenfalls annotiert werden Abbreviaturen und Revisionen. Gleiches gilt für makrostrukturelle Gliederungselemente, d. h. Rubrizierungen, Initialen, Auszeichnungsschriften. Bei der Codierung auf Grafebene haben wir uns entschieden, die unterschiedlichen Grafe einstöckiges-a und zweistöckiges-a, rundes-s und Schaft-s, gerades-r und rundes/Rotunda-r sowie rundes-z und geschwänztes-z zu annotieren.²⁴ Wohingegen die verschiedenen Ausformungen von beispielsweise b, d, l, h, k mit oder ohne Schleife nicht

²⁴ Gegen die Normalisierung von vermeintlich rein grafischen Varianten oder einer Normalisierung nach rein phonologischen Prinzipien spricht sich Kranich-Hofbauer aus. Editorisch bearbeitete Texte sollen die Ausgangstexte im ersten Schritt „deutlich und ‚unkorrigiert‘“ wiedergeben. „Graphische Varianten sind relevant, um anhand von Frequenzanalysen, der Beobachtung von Prä- und Postgraphien auf Gebrauchsnormen schließen zu können“ (Karin KRANICH-HOFBAUER: s/f - ein Fall für die Normalisierung? Ein Beitrag zur Edition spätmittelalterlicher Gebrauchstexte, in: *De consolatione philologiae, Studies in Honor of Evelyn S. Firchow*, hrsg. v. Anna Grotans/Heinrich Beck/Anton Schwob, Band 1, Göttingen 2000, S. 217–230, hier: S. 229).

annotiert werden. Die Auswahl der annotierten Grafvarianten erfolgte unter den Gesichtspunkten:

- eindeutig erkennbare unterschiedliche Ausformungen von Grafvarianten,
- vermeintliche Stellungenabhängigkeit von beispielsweise rundem- bzw. Schaft-s in der historischen Schrift,
- Hinweise, die bestimmte Buchstabenformen zur Schrift liefern und damit beispielsweise eine zeitliche Einordnung erleichtern²⁵.

Verzichtet wurde auf nähere Differenzierung einzelner Majuskeln. Hier ist die Schreibervarianz zu groß, als dass es machbar und sinnvoll erschien, diese teils sehr individuell ausgeführten Buchstaben gesondert zu kennzeichnen.

Superskripte bzw. diakritische Zeichen werden durchgehend, aber unabhängig von ihrer konkreten Realisierung annotiert. D.h. Tremata und Umlautzeichen werde zwar als solche ausgezeichnet, aber ob es sich dabei um punkt-, strich- oder tildenförmige Ausführung handelt, wird nicht berücksichtigt. Ebenfalls annotiert werden Vokale, die in Form von Superskripten auftauchen sowie der i-Punkt. Gerade Untersuchungen auf Teilzeichenebene können wichtige Erkenntnisse zur Frage der Schreiberhände liefern.²⁶

²⁵ Zu Grundüberlegungen bezüglich der Datierung von Handschriften anhand paläografischer Befunde siehe auch Karin SCHNEIDER: Paläographie und Handschriftenkunde für Germanisten. Eine Einführung, Tübingen 2014, S. 13 und 17f.

²⁶ Bezüglich der i-Punkt Analyse siehe auch Wernfried HOFMEISTER/Andrea HOFMEISTER-WINTER/Georg THALLINGER: Forschung am Rande des paläographischen Zweifels. Die EDV-basierte Erfassung individueller Schriftzüge im Projekt DAmalS*, in: Kodikologie und Paläographie im digitalen Zeitalter, (Schriften des Instituts für Dokumentologie und Editorik 2), hrsg. v. Malte Rehbein/Patrick Sahle/Torsten Schaßan, Nordstedt 2009, S. 261–292. Vgl. auch Wernfried HOFMEISTER/Astrid BÖHM/Helmut W. KLUG: Die deutschsprachigen Marginaltexte der Grazer Handschrift UB, Ms. 781 als interdisziplinärer Prüfstein explorativer Revisionsforschung und Editionstechnik, in: editio 30 (2016), S. 14–33., hier bes. S. 18.

Alle Abkürzungen werden systematisiert und annotiert. Unterschieden wird dabei zwischen Brevigrafen²⁷ und Kürzungszeichen, wobei hier die Grenze nicht immer eindeutig gezogen werden kann. Außerdem werden Kontraktionen wie etwa die deutschsprachige *daz*-Kürzung ebenfalls annotiert. Die Kürzungen liegen damit systematisiert und im Detail beschrieben vor.²⁸

Gliederungselemente wie Interpunktionszeichen, Terminatoren, Unterstreichungen, Rubrizierungen, Initialen, Strichelungen und Wort-trennzeichen werden transkribiert und annotiert. Diese Elemente stellen wichtige Informationen bezüglich Textgliederung, Textentstehung, Schreibprozess und Schreibpraxis dar und bieten somit Hilfestellung beim Textverständnis.

²⁷ Diese eigentlich sehr nützliche Kategorie wird in der deutschsprachigen Paläografie kaum benutzt. Vgl. zum Brevigraph: Alpo HONKAPOHJA: Manuscript abbreviations in Latin and English. History, typologies and how to tackle them in encoding, in: *Studies in Variation, Contacts and Change in English* 14 (2013), ch. 2.4, <http://www.helsinki.fi/varieng/series/volumes/14/honkapohja/>. Ein ähnliches Phänomen beschreibt vergleichbar aber schon Paul Arnold GRUN mit dem Begriff *Sigle*. Vgl. Paul Arnold GRUN: Schlüssel zu alten und neuen Abkürzungen. Wörterbuch lateinischer und deutscher Abkürzungen des späten Mittelalters und der Neuzeit mit historischer und systematischer Einführung für Archivbenutzer, Studierende, Heimat- und Familienforscher. Nachbildungen der Originale (Grundriß der Genealogie 6), Limburg a. d. Lahn 1966, S. 7.

²⁸ Um Vollständigkeit zu erreichen, müsste dieses System zumindest noch um die *Nomina Sacra* oder Zeichen von Kurzschriften wie den ‚tironischen Noten‘, die aber in unseren Quellen nicht vorkommen, erweitert werden. Auch hier zeigt der (historische) Überblick, dass die betroffenen Disziplinen, die sich mit Kürzungen befassen, und allen voran die beiden Hauptdisziplinen – Paläografie und Linguistik – mit unterschiedlichem Vokabular über dieselben Phänomene sprechen. Vgl. dazu z. B. einführend allgemein SCHNEIDER: Paläographie, S. 86–91; aus linguistischer Perspektive Franz SIMMLER: Prinzipien der Edition von Texten der Frühen Neuzeit aus sprachwissenschaftlicher Sicht, in: *Probleme der Edition von Texten der Frühen Neuzeit. Beiträge zur Arbeitstagung der Kommission für die Edition von Texten der Frühen Neuzeit*, hrsg. v. Lothar Mundt/Hans-Gert Roloff/Ulrich Seelbach, Berlin 1992, S. 53–61; vor allem in Bezug auf den editorischen Umgang mit Abkürzungen SAHLE: *Editionsformen* 1, S. 76–78; sowie neu ordnend Jürgen RÖMER: *Geschichte der Kürzungen. Abkürzungen in deutschsprachigen Texten des Mittelalters und der Frühen Neuzeit*, Göttingen 1999. Die in unseren Projekten vorgelegte Systematisierung versucht eine vereinfachende, disziplinenferne Einteilung auf Basis des Schriftbefunds.

Genauso wichtig für die Erforschung der Textgenese und Überlieferungssituation sind die im Text befindlichen Revisionsspuren: Im Sinne des ‚fürsorglichen Skriptoriums‘²⁹ geben diese Phänomene Einblick in die Textproduktion. Tilgungen, Einfügungen und marginale Ergänzungen werden transkribiert und kommentierend annotiert. Eine genauere Einordnung, wie sie von Clausen und Klug angeregt wird³⁰, wurde dabei bei der Transkription der Ms. 1609 exemplarisch getestet; auf eine detaillierte, systematische Annotation der Revisionsspuren in den Kochrezepttextsammlungen wurde vorerst aber verzichtet.

Die Entscheidung, welche Informationen die Transkription enthalten soll, unterlag nicht zuletzt dem Rahmen der Möglichkeiten, der u. a. durch Projektziele, zur Verfügung stehender Zeit und Arbeitskraft definiert ist. Es wird mit dieser differenzierten Modellierung aber eine möglichst breite Basis geschaffen, die es anderen WissenschaftlerInnen ermöglicht, die Daten für eigene Forschungsfragen zu erweitern. Wenn keine feingranularen Unterschiede dokumentiert wurden, wie z. B. bei den Superskripten oder im CoReMA-Projekt bei den Revisionsspuren, so beinhaltet die XML-Datei ausreichend Annotationen, dass diese Informationen bei Bedarf mit wenig Aufwand nachgetragen werden können. Eine Edition kann nie alle Bedürfnisse aller möglichen InteressentInnen abdecken! Umso wichtiger ist es, das Potenzial digital transkribierter / edierter Texte aufzuzeigen. Mit dieser Transkriptionsmethode wird eine Grundlage für verschiedene Forschungsansätze gelegt, auf denen weitere Forschung aufbauen kann.

IV. Transkriptionsworkflow

Die einzelnen Arbeitsschritte hin zur Basistranskription einer dynamischen Edition umfassen für uns:

²⁹ Vgl. Jürgen WOLF: Das „fürsorgliche“ Skriptorium. Überlegungen zur literarhistorischen Relevanz von Produktionsbedingungen, in: *Das Mittelalter 7* (2002), S. 92–109.

³⁰ Hans CLAUSEN/Helmut W. KLUG: Schreiberische Sorgfalt. Der Einsatz digitaler Verfahren für die textgenetische Analyse mittelalterlicher Handschriften, in: *Textgenese in der digitalen Edition* (Beihefte zu *editio* 45), hrsg. v. Anke Bosse/Walter Fanta, Berlin 2019/2020, S. 139–151.

- Schriftanalyse/Schriftbefundung,
- die Aufbereitung der Quelle für die Transkription in Transkribus,
- den ersten Transkriptionsdurchgang mit Transkribus,
- mehrere Korrekturdurchgänge mithilfe einer diplomatischen Darstellung und Transkribus,
- der Schritt aus der Transkriptionssoftware hinaus und die endgültige XSLT-Transformation der proprietären Codes in TEI-XML-Daten,
- das Einarbeiten editorischer Eingriffe in die XML-Datei.

Vor Beginn der eigentlichen Transkription wird mittels der bereits beschriebenen 10%-Studie eine Schrift- bzw. Handbeschreibung angefertigt. Durch die Analyse der Schrift werden auch Phänomene ersichtlich, die evtl. nicht durch die allgemeinen Transkriptionsrichtlinien abgedeckt sind. Sollten sich diese als so wichtig erweisen, dass sie in die Basistransliteration aufgenommen werden müssen, kann die Codierung dementsprechend angepasst werden.

Für die Roh-Transkription verwenden wir jene Arbeitsumgebung, die über Transkribus³¹ zur Verfügung gestellt wird. Die eigentlich für die automatische Handschriftenerkennung und -transkription von historischen Dokumenten vorgesehene Plattform bietet mehrere Vorteile für das Projekt: Besonders hervorzuheben ist zum einen die gute Bild-Text-Verknüpfung im Texteditor, die eine zeilengenaue Synopse aufweist, zum anderen hat Transkribus eine seitenbasierte Struktur, sodass durch die ad personam-Zuteilung von Seitenbereichen kollaboratives Arbeiten möglich wird. Die Eingabemaske erinnert darüber hinaus stark an bekannte Textverarbeitungsprogramme, die Möglichkeiten der Annotation sind vielfältig, wenngleich mit der Komplexität des Markups die Fehleranfälligkeit steigt. Die eingegebenen Daten lassen sich u. a. als XML exportieren und können dann in externen Programmen weiterverarbeitet werden.

³¹ <https://transkribus.eu/Transkribus/>. Im CoReMA-Projekt kommt Version 1.4.2 der Software zum Einsatz.

Der konkrete Workflow besteht aus mehreren Schritten: Nach der Anlage eines Projektes innerhalb von Transkribus müssen die Faksimiles auf den Server der Betreiber hochgeladen werden, danach erfolgt die automatische – mittlerweile sehr gut funktionierende – Segmentierung, bei der Textregionen und Textzeilen ausgezeichnet und mit dem Textbereich im Editor verknüpft werden. Der Output muss auf eventuelle Fehler oder nicht erwünschte Zuordnungen kontrolliert und korrigiert werden. Danach kann die eigentliche Transkription des Textes beginnen. Vorsicht ist aber bei der Annotation der Daten geboten: Transkribus bietet eine große Anzahl an vorgegebenen Elementen, mit denen Texte ausgezeichnet werden können, sowie die Möglichkeit, eigene Elemente zu kreieren (vgl. Abbildung 5).

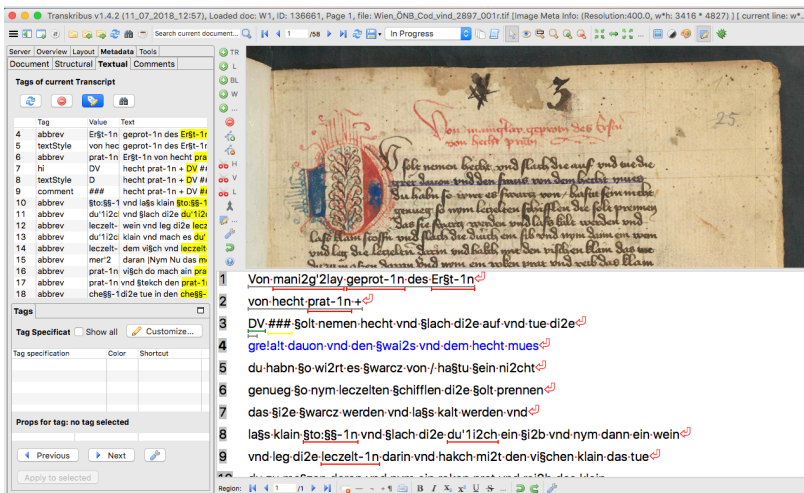


Abb. 5: Das Interface von Transkribus im Transkriptionsprofil

Allerdings führt z. B. das Mehrfachtagging von Zeichen oder Zeichenketten beim XML-Output zu Problemen mit der Verschachtelung einzelner Elemente, im schlimmsten Fall also zu einem nicht validen XML-Dokument. Um ein nachträgliches Ausbessern der Fehler zu vermeiden, sollte unbedingt vor der Transkription ausprobiert werden, welche Elemente sich problemlos miteinander kombinieren lassen und welche

nicht. Das schlussendlich aus Transkribus gewonnene XML-Dokument, das noch das proprietäre Markup enthält, durchläuft verschiedene XSLT-Transformationen und wird in Zuge dessen in das projektspezifische XML-Modell umgewandelt, auf dem dann weitere Bearbeitungsschritte aufbauen.

Zur Überprüfung der Transkription ist aber weder Transkribus noch das exportierte und transformierte XML-Dokument geeignet: In Transkribus behindert die proprietäre Kodierung eine Kollationierung, im XML ist die Annotationsdichte derart hoch, dass der Lesefluss nachhaltig gestört ist. Für den Schritt der Kollationierung von Transkription und Original wird mithilfe von XSL-Transformation eine Kombination aus diplomatischer Zeichenfassung und der normalisierten Darstellung angefertigt (vgl. Abbildung 6).

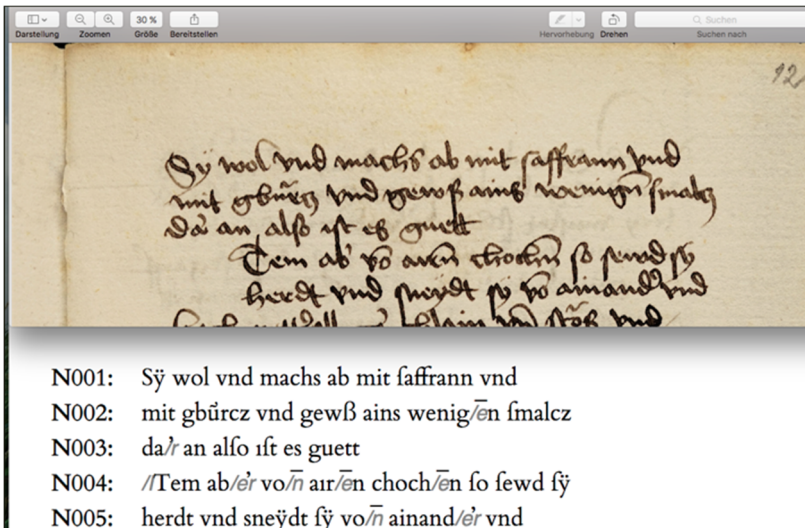


Abb. 6: Synoptische Darstellung von Quellendigitalisat und Kollationierungstext mit diplomatischer Zeichendarstellung und aufgelösten Kürzungen in grauer Schriftfarbe von fol. 12r der HS Graz, UB, Ms. 1609

In dieser Darstellung lässt sich der Text gut und zügig kontrollieren. Die ‚diplomatische‘ Version erleichtert die Überprüfung der Transkription. Für diese Ansicht werden Unicode-Zeichen verwendet, die das

proprietäre Markup in Zeichensymbole auflösen, die eine möglichst zeichengenaue Darstellung der Handschrift gewährleisten sollen. Parallel dazu zeigen die normalisierten Elemente, ob beispielsweise Abbrüviaturen richtig aufgelöst wurden. In Abbildung 6 ist die gleichzeitige Anzeige des Kürzungszeichens sowie die Auflösung der Abbrüviatur ersichtlich (vgl. z. B. *wenigen* in Zeile 2). Diese Textversion ist ausschließlich für den internen Workflow konzipiert. Die Ausgangsdaten für diese Ansicht sind die XML-Datei mit dem entsprechend annotierten Text und die Zeichenbeschreibung der `<charDecl>`, in der alle weiteren Darstellungsoptionen zum speziellen Zeicheninventar zentral gesammelt werden.

Auf diesen Daten bauen alle weiteren Textderivate, die im Zuge einer Edition erstellt werden sollen, auf. Für das CoReMA-Projekt sind das u. a. eine Text-Bild-Synopse, in der ein moderat normalisierter Text als editorisch erschlossene Textvariante den Abbildungen der Quelle gegenübergestellt wird und ein sehr stark normalisierter Text, der als Grundlage für die semantische Annotation herangezogen wird.

V. Zusammenfassung und Ausblick

Mit dem Transkriptions-Workflow, der im Mittelalterlabor-Projekt intensiv getestet und im CoReMA-Projekt vollständig ausgearbeitet wurde, liegt nun ein Arbeitsplan vor, der Projektteams mit unterschiedlichen Aufgaben produktiv unterstützt. Die von uns adaptierte ‚Dynamische Editionsmethode‘ hat sich in der Datenproduktion bewährt und kann vor allem durch die zentral veränderbare Character Declaration sowie die hierarchisch aufeinander aufbauenden XSL-Transformationen, welche den Transkriptionsprozess unterstützen, sehr produktiv eingesetzt werden. Durch die Arbeit mit den unterschiedlichen Handschriften des CoReMA-Projekts, die jede für sich ein ‚kleines Editionsprojekt‘ darstellt, hat sich bestätigt, dass die Datenerfassung und die damit verbundenen Arbeitsschritte in Bezug auf individuelle Anpassungen generisch und flexibel gestaltet sind. Insgesamt wurde hier ein effektiver Workflow erarbeitet, mit dem es auch möglich ist, große Textkorpora unter Einhaltung eines hohen Detailniveaus mit moderatem Zeitaufwand zu bewältigen. Damit ist es möglich, dass der zugrundeliegende Workflow auch für andere

Transkriptions-/Editionsvorhaben mit nur geringfügigen Anpassungen übernommen werden kann.

Als Ausblick kann in diesem Zusammenhang auf die Bereitstellung der notwendigen Dateien (exemplarische Character Declaration, XSLT-Dateien, TEI-XML-Testdatensätze) und einer detaillierten Anleitung im Rahmen des KONDE-GitHub-Repository³² verwiesen werden, das die praktische Ergänzung zu den hier vorliegenden theoretischen Ausführungen darstellt. Bevor das passieren kann, muss aber der Schnelldigkeit der digitalen Welt Rechnung getragen und die XSL-Transformationen an die Änderungen der neuesten Transkribus Versionen angepasst werden: Das verspricht u.a. die Möglichkeit, bereits beim Export aus Transkribus die Transformationen eines durch Tagging erzeugten individuellen Markups und vielleicht sogar des proprietären Markups in TEI-XML durchführen zu können.

³² <https://github.com/KONDE-AT>. Unser Projektteam steht auch gerne für Fragen und Hilfestellungen zur Verfügung, wenn die Aussicht besteht, dadurch mehr tieferenschlossene Basistranskriptionen zur Verfügung zu haben!

Bibliographische Hinweise

Alle Internetlinks wurden am 10.07.2019 überprüft.

Forschungsliteratur

- CLAUSEN, Hans / KLUG, Helmut W.: Schreiberische Sorgfalt. Der Einsatz digitaler Verfahren für die textgenetische Analyse mittelalterlicher Handschriften, in: Textgenese in der digitalen Edition (Beihefte zu editio 45), hrsg. v. Anke Bosse/Walter Fanta, Berlin 2019/2020, S. 139–151).
- FEIGS, Wolfgang: Deskriptive Edition auf Allograph-, Wort- und Satzniveau, demonstriert an handschriftlich überlieferten, deutschsprachigen Briefen von H. Steffens, Teil 1: Methode, Bern 1979.
- GRUN, Paul Arnold: Schlüssel zu alten und neuen Abkürzungen. Wörterbuch lateinischer und deutscher Abkürzungen des späten Mittelalters und der Neuzeit mit historischer und systematischer Einführung für Archivbenutzer, Studierende, Heimat- und Familienforscher. Nachbildungen der Originale (Grundriß der Genealogie 6), Limburg a.d. Lahn 1966.
- HOFMEISTER-WINTER, Andrea: Das Konzept einer ‚Dynamischen Edition‘ dargestellt an der Erstausgabe des „Brixner Dommessnerbuches“ von Veit Feichter (Mitte 16. Jh.). Theorie und praktische Umsetzung (Göppinger Arbeiten zur Germanistik 706), Göppingen 2003.
- HOFMEISTER, Wernfried / HOFMEISTER-WINTER, Andrea / THALLINGER, Georg: Forschung am Rande des paläographischen Zweifels. Die EDV-basierte Erfassung individueller Schriftzüge im Projekt DAMaS*, in: Kodikologie und Paläographie im digitalen Zeitalter, (Schriften des Instituts für Dokumentologie und Editorik 2), hrsg. v. Malte Rehbein/Patrick Sahle/Torsten Schaßan, Nordstedt 2009, S. 261–292.
- HOFMEISTER, Wernfried / BÖHM, Astrid / KLUG, Helmut W.: Die deutschsprachigen Marginaltexte der Grazer Handschrift UB, Ms. 781 als interdisziplinärer Prüfstein explorativer Revisionsforschung und Editionstechnik, in: editio 30 (2016), S. 14–33.
- HONKAPOHJA, Alpo: Manuscript abbreviations in Latin and English. History, typologies and how to tackle them in encoding, in: Studies in Variation, Contacts and Change in English 14 (2013), ch. 2.4, <http://www.helsinki.fi/varieng/series/volumes/14/honkapohja/>.
- KRAGL, Florian: Pseudo-Markup. Eine Eselsbrücke zwischen manueller und maschineller Textverarbeitung, in: Magazin für Editionswissenschaften 1 (2015), S. 7–14.
- KRANICH-HOFBAUER, Karin: Editionswissenschaften als interdisziplinäre Grundwissenschaft. Über Berührungängste und deren Überwindung, in: Jahrbuch der Oswald von Wolkenstein Gesellschaft 12 (2000), S. 49–64.
- KRANICH-HOFBAUER, Karin: s / f - ein Fall für die Normalisierung? Ein Beitrag zur Edition spätmittelalterlicher Gebrauchstexte, in: De consolatione philologiae, Studies in Honor of Evelyn S. Firchow, hrsg. v. Anna Grotans/Heinrich Beck/Anton Schwob, Band 1, Göppingen 2000, S. 217–230.
- RÖMER, Jürgen: Geschichte der Kürzungen. Abbraviaturen in deutschsprachigen Texten des Mittelalters und der Frühen Neuzeit, Göppingen 1999.

- SAHLE, Patrick: Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, Teil 1: Das typografische Erbe, Norderstedt 2013, S. 143–167.
- SCHNEIDER, Karin: Paläographie und Handschriftenkunde für Germanisten. Eine Einführung, Tübingen 2014.
- SIMMLER, Franz: Prinzipien der Edition von Texten der Frühen Neuzeit aus sprachwissenschaftlicher Sicht, in: Probleme der Edition von Texten der Frühen Neuzeit. Beiträge zur Arbeitstagung der Kommission für die Edition von Texten der Frühen Neuzeit, hrsg. v. Lothar Mundt/Hans-Gert Roloff/Ulrich Seelbach, Berlin 1992, S. 53–61.
- WOLF, Jürgen: Das „fürsorgliche“ Skriptorium. Überlegungen zur literarhistorischen Relevanz von Produktionsbedingungen, in: Das Mittelalter 7 (2002), S. 92–109.

Internetlinks

- CoReMA – Cooking Recipes of the Middle Ages. Corpus – Analysis – Visualisation, hrsg. v. Helmut W. Klug unt. Mitarb. v. Astrid Böhm/Christian Steiner, Feb. 2020; hdl.handle.net/11471/562.10 (GAMS.562.10).
- KONDE –Kompetenznetzwerk Digitale Edition. HRSM-Projekt, hrsg. v. Zentrum für Informationsmodellierung - Austrian Centre for Digital Humanities; <http://www.digitale-edition.at>.
- Lyrik des deutschen Mittelalters, hrsg. v. Manuel Braun/Sonja Glauch/Florian Kragl; <http://www.ldm-digital.de>.
- Mittelalterlabor. Transkription der Handschrift Graz, UB, Ms. 1609, hrsg. v. Helmut W. Klug unt. Mitarb. v. Astrid Böhm/Elisabeth Raunig, April 2019; hdl.handle.net/11471/521.60 (GAMS 521.60).
- MUFI character recommendation. Characters in the official Unicode Standard and in the Private Use Area for Medieval texts written in the Latin alphabet, Version 4.0, hrsg. v. Odd Einar Haugen, 22.12.2015; hdl.handle.net/1956/10699.
- STOKES, Peter A.: Describing Handwriting, Part IV: Recapitulation and Formal Model, in: Blog, DigiPal, 14.10.2011; <http://www.digipal.eu/blog/describing-handwriting-part-iv-recapitulation-and-formal-model>.