



Separatum aus:

THEMENHEFT 12

*Elisabeth Lienert / Joachim Hamm
Albrecht Hausmann / Gabriel Viehhauser (Hrsg.)*

Digitale Mediävistik

Perspektiven der Digital Humanities für die Altgermanistik

Publiziert im November 2022.

Die BmE Themenhefte erscheinen online im BIS-Verlag der Carl von Ossietzky Universität Oldenburg unter der Creative Commons Lizenz [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/). Die ›Beiträge zur mediävistischen Erzählforschung‹ (BmE) werden herausgegeben von PD Dr. Anja Becker (München) und Prof. Dr. Albrecht Hausmann (Oldenburg). Die inhaltliche und editorische Verantwortung für das einzelne Themenheft liegt bei den jeweiligen Heftherausgebern.

<http://www.erzaehlforschung.de> – Kontakt: herausgeber@erzaehlforschung.de
ISSN 2568-9967

Zitiervorschlag für diesen Beitrag:

Šimek, Jakub: heiEDITIONS – eine Heidelberger Infrastruktur für Editionen (nicht nur) mittelalterlicher Texte, in: Lienert, Elisabeth/Hamm, Joachim/Hausmann, Albrecht/Viehhauser, Gabriel (Hrsg.): Digitale Mediävistik. Perspektiven der Digital Humanities für die Altgermanistik, Oldenburg 2022 (BmE Themenheft 12), S. 27–46 (online).

Jakub Šimek

heiEDITIONS – eine Heidelberger Infrastruktur für Editionen (nicht nur) mittelalterlicher Texte

Abstract. Die seit 2018 an der Universitätsbibliothek Heidelberg unter dem Namen heiEDITIONS entwickelte Infrastruktur für digitale Editionen baut auf Erfahrungen in der Digitalisierung, in der virtuellen Rekonstruktion von Sammlungen und in der Handschriftenerschließung auf und fügt sich als Teil der Heidelberg Research Infrastructure (heiRIS) zu den in Heidelberg strategisch vorangetriebenen Maßnahmen zur Förderung wissenschaftlichen Publikationswesens im Open Access. Besonderer Wert wird darauf gelegt, dass die Datenmodellierung und Visualisierung den vielfältigen Dimensionen der edierten Gegenstände gerecht werden. Gerade den Bedarfen der Altgermanistik kommt dabei eine wichtige Rolle zu.

Die Universitätsbibliothek Heidelberg ist seit 2014 auf dem Gebiet digitaler Editionen aktiv und baut dafür seit 2018 mit [heiEDITIONS](#) eine dedizierte Infrastruktur auf. Dieses Engagement ist eingebettet in die langfristige Entwicklungsstrategie des Hauses und hat damit u. a. die folgenden Hintergründe:

In den meisten digitalen Editionsprojekten kommt den Digitalisaten der Textträger eine wesentliche Rolle zu. Bei der Anfertigung von Digitalisaten – der Digitalisierung historischen Schriftguts – verfügt die Universitätsbibliothek über reiche Erfahrungen, die bis 2001 zurückreichen.^[1] Unser Digitalisierungsmanagementsystem [DWork](#) (Heidelberger Digitalisierungsworkflow) ist seit Langem etabliert und wird auch von anderen Institutionen genutzt. Die Digitalisierung von historischen Textträgern, besonders

auch von den deutschsprachigen Handschriften der Bibliotheca Palatina, gehört an der UB zu den Schwerpunkten ihrer langfristigen Ausrichtung. Digitalisate der Textträger sind für die meisten digitalen Editionsprojekte ein wichtiger Bestandteil der Edition, nicht nur weil sie als offen einsehbare Editionsgrundlage für nachhaltige editorische Transparenz sorgen: Sie werden auch selbst zu einem Gut an sich und sichern die Textüberlieferung angesichts nicht auszuschließender künftiger Verschlechterung der Originale (oder gar eines Restrisikos ihres Verlusts) zusätzlich ab.

Ein anderer Aspekt sind unsere Erfahrungen mit der digitalen Zusammenführung historisch zusammenhängender Textsammlungen. Dabei handelt es sich einerseits um heute dislozierte Bestände ehemaliger und in der Vergangenheit zerstreuter Bibliotheken, die digital rekonstruiert werden können (hier sind die digitale Zusammenführung der Bibliothek des Reichsklosters Lorsch² und die kürzlich abgeschlossene Digitalisierung der heute außerhalb Heidelbergs aufbewahrten Handschriften der ehemaligen Bibliotheca Palatina³ zu nennen). Andererseits geht es um Zusammenstellungen von historischen Textträgern eines Werkes, sodass die Parallelüberlieferung (die sich örtlich durchaus auch auf viele Standorte verteilen kann) in einer ›virtuellen Bibliothek‹ von einer zentralen Stelle aus überblickt werden kann. In solchen Fällen ist es ein ideell (literaturhistorisch) ausgemachtes ›Werk‹, das den Zusammenhang einer digitalen Sammlung stiftet. Als prominente Beispiele sind die virtuellen Bibliotheken der Editionsprojekte ›Welscher Gast digital‹ oder ›Iwein – digital‹ zu nennen.⁴ Manchmal gelingt es dabei ebenfalls, heute örtlich zerstreute Teile eines ursprünglichen Textzeugen wieder zusammenzuführen.⁵ Begleitend bietet es sich auch an, solche Sammlungen um Digitalisate moderner Abschriften und Editionen des jeweiligen Werkes zu ergänzen, was besonders dann von großem Wert ist, wenn der ursprünglich zugrundeliegende historische Textträger nicht mehr greifbar ist.

Des Weiteren sieht sich die Universitätsbibliothek Heidelberg in der Pflicht, ihre eigenen historischen Bestände nicht nur sicher aufzubewahren

und für die Zukunft zu erhalten, sondern sie auch kontinuierlich zu erschließen. Das erfolgt traditionell über die katalogisierende Beschreibung und neuerdings über die bereits erörterte Digitalisierung, darüber hinaus aber auch über die Förderung von Editionen. Dies als Teil des eigenen Auftrags zu betrachten, ist vielleicht nicht selbstverständlich, kann aber als ein konsequenter Folgeschritt nach der Bild-Digitalisierung gesehen werden, der neuerdings durch maschinelle Texterkennungsverfahren zusätzlichen Aufwind bekommt, da mit der Unterstützung von *machine learning* gewonnene Rohtextdaten zur Grundlage von Editionen werden können. Wenn wissenschaftliche Bibliotheken sich editorisch engagieren, kann man sich an das alte Selbstverständnis klösterlicher Skriptorien erinnert fühlen, die als ›Abteilungen‹ von Klosterbibliotheken für die Bewahrung und Weitergabe von Texten zu sorgen hatten; der steigende Bedarf der Wissenschaft an nachhaltigen und den Steuerzahler durch nicht-gewinnorientierte Preisgestaltung entlastenden Anbietern von digitalen Open-Access-Plattformen ist jedoch ganz und gar aktuell.

So wird unser Einsatz für Editionen dadurch begünstigt und fügt sich auch deswegen organisch zu unserer übrigen Hausstrategie, weil einer unserer Entwicklungsschwerpunkte das wissenschaftliche Publikationswesen ist. Die Verlage [heiUP](#) (Heidelberg University Publishing), [heiBOOKS](#) (Heidelberger E-Books), die Publikationsplattformen unserer Fachinformationsdienste [arthistoricum.net](#) (Fachinformationsdienst Kunst, zusammen mit der Sächsischen Landesbibliothek – Staats- und Universitätsbibliothek Dresden), [Propylaeum](#) (Fachinformationsdienst Altertumswissenschaften, zusammen mit der Bayerischen Staatsbibliothek München) und [FID4SA](#) (Fachinformationsdienst Südasiens) sowie die zahlreichen von uns gehosteten Zeitschriften sind nur einige Belege für diesen Fokus unserer Arbeit.

Die genannten Umstände mögen das Engagement der Universitätsbibliothek Heidelberg, mit dem Aufbau ihrer Editionsinfrastruktur [heiEDITIONS](#) digitale Editionen zu ermöglichen, zu fördern und letztlich auch

auf Dauer zu betreiben, plausibel erscheinen lassen. Sie würden aber vielleicht nicht vollumfänglich einen solchen Vorstoß rechtfertigen, gäbe es für digitale Editionen nicht gewichtige wissenschaftsimmanente Gründe. Die bisherige praktische Erfahrung zeigt, dass die Erstellung digitaler Editionen durchaus teurer ist, länger dauert und mit größeren Arbeitsaufwänden verbunden ist als das Edieren über herkömmliche Druckausgaben. Eine Universitätsbibliothek, die wie die Heidelberger den Wert des gedruckten Buches nach wie vor hochhält und angesichts der Verantwortung für ihre historischen Bestände täglich daran erinnert wird, ihre Arbeit nicht an den nächsten paar Jahren, sondern an Jahrhunderten auszurichten, kann auch den Vorwurf des Kurzlebigen und des Ephemereren, der Online-Publikationen durchaus noch anhaften kann, nicht leichtfertig abtun – zumal die Edition generell als Publikationsart, welche überhaupt erst die Grundlagen für die meisten historischen Wissenschaften schafft, indem sie mit erheblichem, andere wissenschaftliche Publikationen meist weit übersteigendem Aufwand Primärtexte vorlegt, folgerichtig und allein schon aus ökonomischen Gründen die langlebigste Publikationsgattung sein müsste. Die berechtigte Forderung nach Zuverlässigkeit und Dauerhaftigkeit, die an zitierfähige Texteditionen gestellt wird, wird bei digitalen Editionen aktuell noch durch die Tatsache konterkariert, dass keine digitale Edition bisher den Nachweis erbringen konnte, über mehrere Jahrzehnte lang zuverlässig ihre Aufgabe zu erfüllen. Digitalen Editionen wird, wo auch immer sie vorangetrieben werden, ein Vertrauensvorschuss entgegengebracht, ihre Förderer müssen stärker als etwa bei klassischen Druckkostenzuschüssen in Vorleistung treten. Das Risiko großer Ressourcenverschwendung wäre untragbar, wenn nicht glaubhaft gemacht werden könnte, dass digitale Editionen einen Mehrwert haben (oder das Potenzial eines Mehrwerts), der die Risiken einer solchen paradigmatischen Neuentwicklung aufwiegt. Ein Teil dieses Mehrwerts rührt dabei von den informationslogistischen Vorteilen von Open Access und restriktionsfreier

weltweiter Erreichbarkeit, und auch dieser praktische Aspekt der Informationsversorgung sollte als intrinsischer Wert gewürdigt werden.

Eine Universitätsbibliothek, die als Dienstleister für die Wissenschaft agiert, sollte jedoch die Güte ihrer Dienstleistung nicht nur an Äußerlichkeiten der ›Lieferwege‹ messen, sondern in dem Grad, in dem sie selbst an der Methodik der wissenschaftlichen Leistung beteiligt ist, auch die wissenschaftsinhärenten Qualitäten der von ihr bereitgestellten Dienste auf den Prüfstand stellen, also fragen, inwiefern ein Dienst dem wissenschaftlichen Erkenntnisgewinn förderlich ist. Nach bisheriger Erfahrung ist die methodische Involvierung unserer Bibliothek in die digitalen Editionsprojekte, die wir als Partner der Wissenschaft betreiben, beträchtlich, da wir stärker als bei Druckpublikationen häufig am gesamten ›Lebenszyklus‹ einer digitalen Edition beteiligt sind, nicht zuletzt durch die Erfordernisse der Datenmodellierung an editionsmethodischen Entscheidungen partizipieren und diese durch vorbereitende Beratung und laufende Unterstützung mitprägen. Eine Infrastruktur wie heiEDITIONS, die vielen editorischen Gegenständen und Ansätzen ein Obdach bieten will, könnte ihre Aufgabe aber nicht auf skalierbare, zukunftsorientierte und letztlich wirtschaftliche Weise erfüllen, wenn sie nicht über einen einheitlichen konzeptuellen Rahmen verfügen würde, der die Heterogenität der Gegenstände und Ansätze in einer einheitlichen Matrix aufnehmen könnte. Eine Editionsinfrastruktur für viele ließe sich nicht betreiben, wenn jede Edition *a class of its own* wäre. Dadurch sind wir als Anbieter für die konzeptuelle Gestaltung des Dienstes verantwortlich, den wir für digitale Editionen entwickeln, und müssen solche Konzepte, die nicht selten genuin wissenschaftliche editionsmethodische Konsequenzen mit sich bringen, selbst kritisch hinterfragen.

Digitale Editionen haben die Chance, sich im Vergleich zu herkömmlicher Editorik in einem inhärent anspruchsvolleren Editionsparadigma zu verorten, das nicht nur zu einem besseren Arbeitsprozess des Edierens führt, sondern das neuen und höheren Ansprüchen zu genügen verhilft, die

an Editionen von den sie nutzenden Wissenschaften heute und künftig herangetragen werden. Mit ›digitaler Edition‹ meine ich nicht in erster Linie die Publikationsweise im *World Wide Web*, sondern beziehe mich vor allem auf die Freiheiten der digitalen Arbeitsweise sowie auf die Flexibilität und Ausdrucksmächtigkeit digitaler Datenmodelle, die sich nicht oder nicht ausschließlich von einer Endpräsentation ableiten, weder von einer druckorientierten typographischen Gestaltungsform noch von der einen oder anderen digitalen Anzeigemöglichkeit: Mein Fokus liegt auf einem Selbstverständnis der Edition, das sich primär an den Erfordernissen der Materie ausrichtet, also an der Frage, wie das zu Edierende am sachgerechtesten dokumentiert, erfasst, wiedergegeben, durchdrungen, erschlossen und natürlich letztendlich auch präsentiert werden kann – das Letztere aber nur als ein Aspekt von vielen. Sicher sind die *interfaces*, die Präsentationsformen, entscheidend für die Rezeption der Ergebnisse einer Edition, jedoch muss das Edieren – anders als vielfach bei der herkömmlichen Erstellung nur gedruckter Ausgaben – nicht von vornherein durch sie vorgegeben werden. Digitale Datenmodelle für Editionen können mehr fassen, als jede einzelne Präsentationsform wiedergeben kann. Das digitale Edieren eröffnet einem Herausgeber mehr als zuvor die Chance, sich bei der Konzeption einer Edition primär durch die Beschaffenheit der edierten Materie leiten zu lassen, indem womöglich mehr Dimensionen des Edierten zu ihrer vollen Geltung kommen können, als wenn eine vorgegebene Präsentationsform von Anfang an die editorische Methodik steuert. Wenn die New Philology in den um 1990 mehr erhofften als tatsächlich verfügbaren Möglichkeiten der Computertechnik begeistert eine Wende begrüßte, durch die es möglich würde, die Materialmenge einer Edition ohne einen durch den Druck vorgegebenen Auswahlzwang vollumfänglich zu präsentieren (also etwa alle Textzeugen in Digitalisaten und Transkripten),⁶ so erkennen wir heute deutlicher, dass die Ausdrucksfähigkeit digitaler Datenmodelle Editionen neue qualitative Möglichkeiten eröffnet, ihrem Gegenstand umfassender gerecht zu werden.

Bei der Erstellung herkömmlicher gedruckter Ausgaben bestimmt häufig eine für lineare Lektüre zu optimierende, nach semantisch-logischen Einheiten anzuordnende Textgestaltung die editorische Arbeitsweise, wobei ein primär als sinntragender Sprachausdruck aufgefasster Text vom visuellen und materiellen Substrat seiner historischen Überlieferung abzulösen ist und Letzteres allenfalls beschreibend in die Edition Eingang finden kann. Alternativ kann auch in herkömmlicher Arbeitsweise – von der Faksimilierung einmal abgesehen – ›diplomatisch‹ ediert werden, wobei viel von den graph(emat)ischen und auch visuellen Eigenschaften der historischen Quelle abgebildet werden kann, die semantisch-logische Textebene aber unweigerlich zu kurz kommt. Auch eine linguistische Erschließung ist für herkömmliche druckorientierte Editionen etwa über (ggf. lemmatisierte) Konkordanzen möglich; die Benutzung bleibt dabei aber freilich meist statisch auf eine alphabetische Listenanordnung der Konkordanz angewiesen. Für inhaltliche Erschließung von *named entities*, intertextuellen Bezügen und Sachbegriffen haben sich in Druckeditionen Register als nützliches Werkzeug etabliert; ihre Funktionalität beschränkt sich aber vor allem auf eine Nutzungsrichtung (Suche im Register als Weg zu Fundstellen im Text), und die Tragweite von Registern kann kaum über ein einzelnes ediertes Werk hinausreichen. Für die Abbildung der Überlieferungsvarianz verfügen herkömmliche Editionen über das Mittel des *apparatus criticus*, dessen Grenzen bekannt sind: Sollen sie ihren Nutzer nicht durch unüberblickbaren Umfang erdrücken, werden sie zu Schaukästen ausgewählter Lesarten, bei deren Rezeption der Leser dem Herausgeber einen erheblichen Vertrauensvorschuss gewähren muss; das Nachvollziehen einer Lesart im Originalkontext eines Textzeugen ist schwer möglich.

Die gerade aufgezählten Ebenen des Edierten – neben der semantisch-logischen Ebene der Sinneinheiten die visuell-materielle Ebene des Textträgers, die linguistische Ebene, die extensionale Ebene der Verweise auf textexterne Dinge (›Extrareferenzialität‹) sowie die Ebene der Überlieferungspluralität und somit der Varianz – sind auch in herkömmlichen

Editionen bekannt. Die traditionellen Lösungen für den Umgang mit ihnen werden ihnen aber jeweils nur partiell gerecht, weil es weder in einem herkömmlichen Editionsmanuskript des Herausgebers noch in einer einzigen statischen Druckform möglich ist, ihnen jeweils umfassend Rechnung zu tragen.

Digitale Editionsdatenmodelle können vollwertige Räume für mehrdimensionale Texterschließung schaffen: Die Qualität einer Editionsinfrastruktur, die für viele Editionsgegenstände und -paradigmen ein angemessenes Zuhause bieten soll, wird sich daran messen müssen, wie sachgemäß und ausdrucksfähig die Lösungen sind, welche die Infrastruktur für einen integrierten Umgang mit den genannten Ebenen bereitstellen kann, sowohl in der Datenmodellierung als auch in der Präsentation und nicht zuletzt in der Datenerarbeitung, also beim *authoring* der Edition, in Hilfsmitteln für den Arbeitsprozess der Herausgeber.

Die Heidelberger Editionsinfrastruktur heiEDITIONS wird mit Blick auf solche Ansprüche entwickelt, wobei für einige der genannten Aspekte bereits Lösungen existieren oder in Sicht sind, andere noch nicht über Experimente (wie den Umgang mit der linguistischen Annotation) oder erste Schritte (Erleichterungen beim *authoring*) hinausreichen.

Ein grundlegendes Problem ist die Unterscheidung der genannten (traditionell primären) semantisch-logischen Textebene, in der ein Text meist als eine mehr oder weniger hierarchisch strukturierte, geordnete Abfolge von Sinneinheiten wie Kapiteln, Strophen, Prosaabsätzen oder Versen aufgefasst wird (wobei in Verstexten die Situation zusätzlich dadurch verkompliziert wird, dass formal-poetische Einheiten wie Verse auch nicht zwingend mit Sinneinheiten zusammenfallen), von der Ebene der visuellen Einheiten von Textträgern, die im Wesentlichen durch Seitenflächen, Layoutbereiche und Zeilen konstituiert wird. Diese beiden Ebenen interagieren miteinander auf vielfältige Weise. Häufig gehen sie miteinander Symbiosen ein (Beginn eines Kapitels auf einer neuen Seite; Absetzung von Versen) und machen gegenseitig voneinander ›Gebrauch‹, in anderen Fällen schei-

nen sie sich neutral zueinander zu verhalten und überlappen einander ohne erkennbare Rücksicht, gelegentlich laufen sie aber auch einander zuwider (Zeilenüberläufe, bei denen der Text eines inhaltlichen Segments in einer ›fremden‹ Zeile fortgesetzt wird; Briefe, deren Text nach der letzten Seite auf den eigentlich vorangegangenen Seiten fortgeschrieben und am Rand der ersten Seite abgeschlossen wird). Aus mittelalterlichen Handschriften zur Genüge bekannt sind auch diverse Formen von Marginalien, welche die Linearität des Textstroms durch Anmerkungen unterbrechen oder durch Einschübe bewirken, dass der inhaltlich zu lesende Text visuell immer wieder von einem Layoutbereich in einen anderen springt.

Wenn die visuelle Struktur des Textträgers ernst genommen werden soll, um etwa Seite für Seite – in Korrelation zum Digitalisat – eine diplomatisch zeilengetreue Textanzeige zu bewerkstelligen oder um überhaupt erst einen Rohtext mithilfe von Texterkennungssystemen⁷ zu erarbeiten und das dabei gewonnene Text-Bild-Alignment auch bei weiterer editorischer Textaufbereitung nicht zu verlieren, braucht es Mechanismen, mit denen die visuelle Struktur unabhängig von der Sinnstruktur erfasst werden kann. Auch für die editorische Neuordnung der zu lesenden Textabfolge im potenziellen Widerspruch zur Textreihenfolge der Quelle sind Lösungen notwendig.

Die Datenmodellierung in heiEDITIONS basiert zum allergrößten Teil auf der Text Encoding Initiative (TEI), bei der die semantisch-logische Textstruktur grundsätzlich (und auch historisch) Vorrang hat (vgl. Burnard 2013, Abschnitt 13). Um eine von Sinnstrukturen unabhängige Erfassung visueller Einheiten zu ermöglichen, sind in heiEDITIONS TEI-konforme Erweiterungen vorgesehen, die eine vollständige Trennung der Ebenen ermöglichen. Darauf basierend sieht auch die Visualisierung in heiEDITIONS zwei grundverschiedene Ansichten (*views*) der Editionstexte vor (die selbstredend aus einer einzigen Datenbasis erzeugt werden): In der ›Leseansicht‹ wird ein zur Lektüre sinnvoll angeordneter und nach Sinneinheiten strukturierter Text präsentiert, der keinerlei Layoutaspekte der histori-

schen Quelle nachbildet, nur punktuell die Seitenanfänge der Quelle im Text anzeigt und zu diesen im Digitalisat verlinkt. Die Navigation im Text orientiert sich hier an inhaltlichen Abschnitten wie Kapiteln. In der ›Quellenansicht‹ hingegen bestimmen die visuellen Strukturen des Textträgers die Navigation und die Textanzeige. Der Leser ›blättert‹ durch das Digitalisat des Textträgers und bekommt seitenweise einen Text präsentiert, der wie in der Quelle in Layoutbereichen und Zeilen untergebracht ist, einschließlich der positionsgetreuen Anordnung von Spalten und Marginalien.

Die jeweils spezifische Leistungsfähigkeit der ›Quellenansicht‹ und der ›Leseansicht‹ geht weit darüber hinaus, den Text einmal mit Originalumbruch und einmal ohne diesen einzublenden, und der Unterschied zwischen den beiden Ansichten hat auch wenig damit zu tun, dass der Text auf Zeichenebene einmal originalgetreu und einmal z. B. normalisiert oder mit übernommenen editorischen Korrekturen erscheint. Die Textdarstellung auf Zeichenebene ist (je nach editorischer Erschließungstiefe) grundsätzlich für beide ›Ansichten‹ dynamisch einstellbar. Die ›Quellenansicht‹ und die ›Leseansicht‹ unterscheiden sich vielmehr in der Makroanordnung des Textes oberhalb der Zeichenebene, und sie sind in der Lage, den Text nicht nur jeweils anders hierarchisch strukturiert, sondern auch in einer jeweils anderen Reihenfolge, ggf. auch komplett umgeordnet, zu präsentieren. Somit wird es künftig auch für Editionen, die mit Texterkennungssystemen arbeiten und dabei zeilenweise Koordinateninformationen im Digitalisat gewinnen, möglich sein, diese enge Kopplung an das Digitalisat der Quelle beizubehalten und den Text bei Bedarf dennoch nach inhaltlichen Gesichtspunkten editorisch neu zu sortieren.

Für die inhaltliche Erschließung der in Texten erwähnten Personen, Orte, Ereignisse, Werke, Begriffe u. Ä. bietet heiEDITIONS eine Registerfunktionalität, die an herkömmliche Buchregister angelehnt ist, aber im Vergleich zu ihnen wesentliche Mehrwerte mit sich bringt. Zum einen ist die Nutzungsrichtung nicht auf eine Suche nach Textfundstellen in einer Liste beschränkt, da annotierte Textstellen auch als solche in der visuali-

sierten Edition direkt sichtbar sind und beim Anklicken weiterführende Informationen bieten (perspektivisch auch direkt eine Auflistung weiterer Fundstellen derselben Sache, die momentan noch lediglich über einen ›Listeneinstieg‹ des Registers zugänglich ist). Zum anderen können heiEDITIONS-Register zu einer Entität viel mehr als nur eine Ansetzungsform des Namens enthalten: Personen können mit Lebensdaten, Orte mit Lageinformationen, jeder Eintrag außerdem mit beliebigen Erläuterungen, Links (etwa zu Wikipedia oder zu speziellen Internetressourcen), bibliographischen Angaben und Bildern versehen werden, sodass solche Register auch zu kleinen Nachschlagewerken oder Glossaren werden und bei geographischen Daten die Form von Gazetteers annehmen können. Darüber hinaus sind Verknüpfungen zu Normdateien wie der [GND](#) oder [GeoNames](#) möglich und erwünscht, über die weitere Angaben aus solchen kollaborativ erarbeiteten zentralen *knowledge bases* zur Anzeige gebracht werden können und die inhaltliche Erschließung von Editionstexten künftig auch für projektübergreifende Fragestellungen verwertet werden kann.

Dem Aspekt der Überlieferungspluralität edierter Texte begegnet heiEDITIONS mit zwei Ansätzen, die beide an traditionelle typographische Lösungen anknüpfen, dabei aber deutliche funktionale Verbesserungen mitbringen: Synopsen und textkritischer Variantenapparat.

(1) Die synoptische Nebeneinanderstellung mehrerer Volltexte, die an sich in der Schriftkultur sogar noch vor der Erfindung des Codex als Mittel zur Darbietung von Textversionen und zur Erleichterung der Rezeption von Textvarianz angewandt wurde – man denke an die ›Hexapla‹ des Origenes –, ist auch in der modernen editorischen Druckwelt weit verbreitet, bleibt aber durch physische Formate und die Notwendigkeit, statisch endgültig festzulegen, was miteinander visuell korrelieren und daher nebeneinander gesetzt werden soll, eingeschränkt. Die digitale Datenmodellierung macht es möglich, die Korrelation zwischen Teilen von Textversionen präziser zum Ausdruck zu bringen, etwa dann, wenn ›asymmetrische‹ Korrelationen zwischen jeweils ungleicher Anzahl von Sinneinheiten (z. B.

ein Vers korreliert mit einem Verspaar) beschrieben werden sollen oder wenn es darum geht, die hypothetische Position eines Textsegments in einem Textzeugen, in dem das betroffene Textsegment fehlt, im Vergleich zu anderen Textversionen sinnvoll zu verorten. Die dynamische digitale Präsentation gestattet es wiederum überhaupt erst, einerseits die überlieferte Textabfolge eines Textzeugen in originaler Anordnung zu präsentieren und andererseits auf Nutzerwunsch für jede Sinneinheit die damit in anderen Textversionen jeweils korrelierenden Stellen direkt nebeneinander anzuzeigen – für Texte mit stark variabler Anordnung wie den ›Armen Heinrich‹⁸ (aber grundsätzlich für jede Mehrfachüberlieferung mit Umstellungen) mit Mitteln statischer Drucksynopsen ein Ding der Unmöglichkeit.

heiEDITIONS bietet für diese Herausforderung eine Lösung in Form der sog. ›Fokuszeile‹.⁹ Der Nutzer wählt in einer beliebigen Textversion in einer synoptischen Spalte ein Textsegment, woraufhin sich alle anderen Versionen (in anderen Spalten) an diesem Segment mit ihren jeweils korrelierenden Textstellen vertikal ausrichten (einschließlich identifizierter vermuteter Fehlstellen) und so gleichsam eine Zeile bilden, in der die gegenseitige Korrelation gewährleistet ist, während oberhalb und unterhalb dieser ›Fokuszeile‹ die Textversionen in ihrer jeweiligen Originalanordnung bleiben, sodass sich bei Umstellungen, Hinzufügungen oder Fehlstellen außerhalb der Fokuszeile nicht mehr zwingend eine Korrelation im Nebeneinander ergibt. Die Festlegung der ›Auflösungsstufe‹ der Synopse (metaphorisch in Analogie zur Bildauflösung), also der hierarchischen Ebene der Sinneinheiten, deren gegenseitige Korrelation durch die Synopse ausgedrückt werden soll (z. B. Verse versus Strophen), obliegt dem Herausgeber.

(2) Auch für die digitale Modellierung des Variantenapparats wäre eine präzise Beschreibung der Korrelation möglichst auf Tokenebene die eigentlich idealerweise anzustrebende Lösung, denn dann wäre für jedes beliebige Textsegment in jedem Textzeugen aus den Editionsdaten abzulesen, welche Tokens in jedem anderen Textzeugen die jeweilige Entsprechung

bilden. Allerdings wäre so nur gesagt, was einander entspricht und nicht, wie beschaffen eine Entsprechung ist: Die eigentlich relevante Frage nach der Art der etwaigen Varianz wäre damit noch nicht beantwortet. Für die digitale Modellierung einer derart feinen Korrelationsbeschreibung zwischen Textversionen auf Tokenebene wäre der Ansatz des *variant graph*¹⁰ auch für heiEDITIONS vielversprechend, da erst das Graphmodell (im Gegensatz zu tabellarischen Strukturen) die erforderliche Ausdrucksmächtigkeit angesichts von Umstellungen in Aussicht stellt. Die Art der Entsprechung und damit die Beschaffenheit der Varianz zu beschreiben, ist aber schwieriger; dies könnte auf Tokenebene bis zu einem gewissen Grad durch linguistische Annotation möglich gemacht werden (was eine maschinelle Auswertung der Varianz nach linguistischen Kategorien wie der Lexik oder *part of speech* erlauben würde). Eine vollwertige semantische Beschreibung der Varianz – sei es von größeren variierenden Wortgruppen oder in inhaltlichen Nuancen – ist so aber kaum vorstellbar (wobei diese Einschränkung den Wert der linguistischen Annotation nicht mindern soll, die auch in heiEDITIONS perspektivisch angestrebt wird).

Für die greifbare Zukunft realistischer erscheint uns für den digitalen editorischen Umgang mit Varianzaufbereitung ein an herkömmliche Formen angelehnter Variantenapparat, der einerseits von den bekannten typographischen Konventionen des *apparatus criticus* ausgeht, sich andererseits aber die digitale Verfügbarkeit der vollständig transkribierten Paralleltexte (neben einem zum Leittext gewählten Textzeugen) voll zunutze macht – natürlich nur dann, wenn eben tatsächlich die miteinander kollationierten Texte als Volltranskripte vorliegen (und somit in der Synopse auf einer größeren Ebene gewählter Sinneinheiten sozusagen schon ein positiver Variantenapparat gegeben ist, z. B. versweise). Wenn man dann nach erfolgter philologisch-intellektueller Kollation als Herausgeber die für die Präsentation im Apparat (nach eigenen Kriterien) ausgewählten Lesarten angäbe, würde man diese Lesarten nicht als Zeichenketten aus dem Paralleltext abtippen oder kopieren, sondern man würde eine Verknüpfung

zwischen der als Lemma festgelegten Passage des Leittextes und einem Segment des Paralleltextes setzen.

Daraus ergäbe sich nicht nur eine effiziente Arbeitsweise, die gleichzeitig unnötige Fehler von selbst vermeide, sondern die so mit dem Leittext verknüpften Lesarten könnten ebenso dynamisch anpassbar zur Anzeige gebracht werden wie der Leittext selbst (mit Darstellungsoptionen wie dem einstellbaren Umgang mit Abkürzungen). Die angezeigten Lesarten könnten zudem in einer Online-Visualisierung mit Verlinkungen in die Volltexte der Parallelversionen unterlegt werden, sodass man beim Anklicken einer Lesart zur entsprechenden Stelle im Volltranskript des abweichenden Textzeugen gelänge.¹¹

Da die Qualität der TEI-basierten Datenmodelle in heiEDITIONS direkt mit der Eignung der Infrastruktur zusammenhängt, die skizzierten editionsrelevanten Textdimensionen editorisch abzubilden, kommt der Datenmodellierung und dem Datenmodellmanagement ein hoher Stellenwert zu. Zudem müssen Aufgaben der Datenmodellierung auch den ökonomischen Anforderungen an Beherrschbarkeit und Skalierbarkeit selbst bei großer Heterogenität editorischer Paradigmen und Gegenstände standhalten. Schließlich sollen die in Editionsprojekten erarbeiteten TEI-Daten zuverlässig und robust visualisiert werden können, ohne dass für jedes einzelne Projekt *ad hoc*-Lösungen erfunden werden müssten – solche sind vielmehr nach Möglichkeit zu vermeiden, zugunsten einer zentralen und allen Projekten gemeinsamen technischen Infrastruktur. Nur eine Plattform, deren modulare Komponenten mit ihren Verarbeitungs- und Visualisierungstools vergleichbaren Prinzipien folgen, kann selbst bei wachsender Komplexität und Anzahl der Editionen beherrschbar bleiben. Dank einer möglichst homogenen Technik werden einzelne Projekte auch von künftigen Weiterentwicklungen des Systems profitieren und ggf. in neue Formate migriert werden können. Die Einheitlichkeit der Systeme erfordert dann ihrerseits eine Einheitlichkeit der Daten, die sich der Datenmodellierung verdankt.

In heiEDITIONS wird für diese Aufgabe ein dediziertes Set an Tools eingesetzt, mit deren Hilfe ein TEI-konform aufgebautes und für die eigenen Bedarfe angepasstes XML-Schema (eine sog. *customization* der TEI: heiEDITIONS Schema), ein als Ontologie im Sinne des Semantic Web strukturiertes Kategoriensystem (heiEDITIONS Concepts) sowie Anleitungen und eine technische Dokumentation (heiEDITIONS Dokumentation) gepflegt werden.¹² Zum Einsatz kommen dabei in erster Linie die von der TEI selbst bereitgestellten Werkzeuge; eine wesentliche Erweiterung stellt die Einbindung des auf RDF (Resource Description Framework) und OWL (Web Ontology Language) basierenden Kategoriensystems heiEDITIONS Concepts dar, das in den konzeptuellen Modellen [CIDOC CRM](#) und [FRBRoo](#)¹³ verankert ist und für eine breite Palette überlieferungsbezogener und editorischer Phänomene normierte Definitionen und Identifikatoren bietet, auf die in TEI-Dokumenten direkt Bezug genommen wird. Somit wird das in TEI-Daten zu verwendende taxonomische System an einer zentralen Stelle gepflegt; bei immer wieder erforderlichen Ergänzungen werden neue Konzepte in das bestehende Kategoriensystem so eingepflegt, dass sie sich möglichst in den bestehenden Rahmen fügen.

Ziel der Datenmodellmanagement-Komponente von heiEDITIONS ist eine interne Interoperabilität¹⁴ innerhalb des Hauses, damit einmal entwickelte Verarbeitungs- und Präsentationswerkzeuge unter Vermeidung unnötiger Redundanzen über Projektgrenzen hinweg genutzt werden können und Neuerungen bei kontinuierlicher Weiterentwicklung möglichst vielen Projekten zugute kommen.

heiEDITIONS ist an der Universitätsbibliothek Heidelberg eingebettet in heiRIS (Heidelberg Research Infrastructure) – ein breit gefächertes Ökosystem miteinander lose gekoppelter Dienste, die von der DOI-Vergabe über die Speicherung audio-visueller Medien bis hin zur etwaigen Buchpublikation, zur bibliothekarischen Katalogisierung einer Edition und zur Langzeitarchivierung der Editionsdaten reichen (zu heiEDITIONS im Kontext von heiRIS vgl. Effinger [u. a.] 2019). Alle Editionen nutzen für die Er-

werbung, Verwaltung und Präsentation von Digitalisaten das bewährte, anfangs bereits genannte Heidelberger Digitalisierungssystem DWork.

Nicht alle Aspekte von heiEDITIONS sind für die Altgermanistik relevant, da die Infrastruktur ein breites Spektrum editorischer Gegenstände bedient, zu denen auch moderne Gelehrtenkorrespondenzen oder Editionen aus dem Umfeld der Kunstgeschichte gehören. Dennoch hat die Altgermanistik in heiEDITIONS eine besondere Stellung inne, und die Weiterentwicklungsperspektiven der Infrastruktur orientieren sich – meist induktiv in strategisch ausgewählten Pilotprojekten – ausdrücklich auch am Bedarf des Faches.

Anmerkungen

- 1 Zum ersten Digitalisierungsprojekt ›Digitalisierung spätmittelalterlicher Handschriften aus der Bibliotheca Palatina‹, das mit DFG-Förderung in den Jahren 2001–2002 stattfand, und den damaligen Perspektiven vgl. Effinger [u. a.] 2000; Saurma-Jeltsch/Effinger 2001; sowie Effinger [u. a.] 2003.
- 2 ›Bibliotheca Laureshamensis – digital: Virtuelle Klosterbibliothek Lorsch‹.
- 3 ›Bibliotheca Palatina – digital: Virtuelle Rekonstruktion der einst berühmtesten Büchersammlung Deutschlands‹; vgl. Probst 2017.
- 4 Im altgermanistischen Kontext sind ferner die Sammlungen der Editionsprojekte ›Kaiserchronik – digital‹, ›Der arme Heinrich – digital‹, ›Lübisches Recht – digital‹, und ›Nikolaus von Jeroschin – digital‹ zu nennen. Im letztgenannten Fall handelt es sich sogar eher um eine Autorsammlung, da neben der ›Krönike von Pruzinlant‹ auch das Fragment der ›Adalbert-Vita‹ zum Angebot gehört.
- 5 So etwa bei Fr1, Fr2 und Fr8 der ›Krönike von Pruzinlant‹ des Nikolaus von Jeroschin (<https://digi.ub.uni-heidelberg.de/nvjd/jeroschin/handschriften.html>).
- 6 Vgl. Cerquiglini 1989, S. 112–116; kritisch dazu Stackmann 1993, S. 8f., und ders. 1994, S. 418.
- 7 Besonders relevant für Editionen sind aktuell im Bereich der Texterkennung (OCR – *optical character recognition* bzw. HTR – *handwritten text recognition*) die Programme [Transkribus](#) und [eScriptorium](#). Die UB Heidelberg erwägt perspektivisch eine Integration der Open-Source-Software eScriptorium in ihre Texterkennungsworkflows und betreibt eine Testinstanz des Systems. Nicht

zuletzt die Gewinnung des Text-Bild-Alignments (Koordinaten von Textsegmenten auf digitalen Bildern) halten wir dabei auch im Editions-kontext für attraktiv.

- 8 Vgl. Fernández Riva/Millet 2022.
- 9 Unser Konzept der ›Fokuszeile‹ wurde unter maßgeblicher Beteiligung von Leonhard Maylein, der auch für die technische Umsetzung verantwortlich zeichnete, im Rahmen der Edition ›Der arme Heinrich – digital‹ erstmalig entwickelt und wird derzeit weiter verbessert.
- 10 Zum *variant graph* s. Schmidt/Colomb 2009. In abgewandelter Form ging dieses Graphmodell in die Softwareprojekte CollateX und Stemweb ein, vgl. dazu z. B. Dekker [u. a.] 2014 und Andrews 2014. Dabei werden Tokens, die quer über mehrere Textzeugen als gleich betrachtet werden, jeweils durch einen Knoten repräsentiert; die Kanten drücken aus, in welchen Textzeugen die durch eine Kante verbundenen Knoten vertreten sind; die Richtung der Kanten steht für den Textfluss. Gewisse Schwächen des Modells in den bisher vorgestellten Implementierungen bestehen aus meiner Sicht darin, dass bei asymmetrischer Korrelation von Tokens (wenn jeweils eine ungleiche Anzahl von Tokens einander entspricht) nicht präzise zum Ausdruck gebracht wird, welche Tokens bzw. Tokengruppen einander entsprechen, wenn in den verglichenen Textversionen in direkter Nachbarschaft zu einer Korrelationsstelle gleichzeitig auch unbeteiligte Hinzufügungen oder sonstiges Fremdmaterial (Interpunktion, Glossen, Paratexte) vorlägen; ferner dass keine ›Schachtelung‹ innerhalb miteinander korrelierender Textsegmente abgebildet werden kann (Korrelation auf mehreren Ebenen bzw. in unterschiedlichen Präzisionsstufen); und schließlich dass noch nicht an die Abbildung von Varianz unterhalb der Tokenebene gedacht wurde (vgl. hypothetisch ›hellblau‹ vs. ›hellrot‹ vs. ›dunkelblau‹ vs. ›dunkelrot‹). Dessen ungeachtet gibt ein *variant graph* die Variation zwischen Textversionen in ihrem Textfluss sehr gut wieder.
- 11 Die skizzierte Form des Variantenapparats soll im Rahmen des Editionsprojekts ›Boners *Edelstein* – digital‹ realisiert werden, das die UB Heidelberg zusammen mit Gerd Dicke als wissenschaftlichem Herausgeber plant. Die Digitalisate aller Textzeugen sind bereits im Open Access verfügbar.
- 12 Die sukzessive erarbeitete Dokumentation von heiEDITIONS ist in Teilen bereits unter <https://heieditions.github.io> zugänglich.
- 13 CIDOC CRM ist ein konzeptuelles Referenzmodell für die Beschreibung von Gegenständen des kulturellen Erbes; FRBRoo ist eine mit CIDOC CRM harmonisierte Form des bibliothekarischen Referenzmodells FRBR (›Functional Re-

quirements for Bibliographic Records»). U. a. die FRBR-Kategorien ›Werk‹, ›Expression‹, ›Manifestation‹ und ›Exemplar‹ erweisen sich auch bei der Modellierung von Editionsdaten als operational nützlich, etwa wenn es zu unterscheiden gilt zwischen singulären Phänomenen einer materiellen Quelle und dem, was einen Text auch unabhängig von einem Träger ausmacht, oder wenn multiple Textversionen einem gemeinsamen ›Werk‹ zugeschrieben werden. Freilich entscheidet immer letztlich der Herausgeber, was er editionspraktisch für ein ›Werk‹ hält.

- 14 Die Interoperabilität (direkte technische Nachnutzbarkeit) der heiEDITIONS-gemäß gestalteten TEI-Daten wird tatsächlich nur im Rahmen der hauseigenen Infrastruktur bezweckt, darüber hinaus ist sie kaum realistisch, weil die TEI große Spielräume bei konkreter Anwendung zulässt und Anpassungen durch ihre Nutzer geradezu vorsieht. Gleichzeitig legen wir großen Wert darauf, dass die TEI-Daten von heiEDITIONS mit anderen Interessierten austauschbar sind, d. h. dass sie transparent und verständlich sind und damit von Dritten mit vernünftigen Aufwand wiederverwendet werden könnten. Zur Diskussion der Begriffe *interoperability* und *interchange* im TEI-Kontext vgl. Holmes 2016.

Literaturverzeichnis

Sekundärliteratur

- Andrews, Tara L.: Analysis of variation significance in artificial traditions using Stemmaweb, in: Digital Scholarship in the Humanities 31 (2016) [im Advance Access bereits 2014], S. 523–539 ([online](#)).
- Burnard, Lou: The evolution of the Text Encoding Initiative. From research project to research infrastructure, in: Journal of the Text Encoding Initiative 5 (2013) ([online](#)).
- Cerquiglini, Bernard: Éloge de la variante. Histoire critique de la philologie, Paris 1989.
- Dekker, Ronald Haentjens/van Hulle, Dirk/Middell, Gregor/Neyt, Vincent/van Zundert, Joris: Computer-supported collation of modern manuscripts. CollateX and the Beckett Digital Manuscript Project, in: Digital Scholarship in the Humanities 30 (2015) [im Advance Access bereits 2014], S. 452–470 ([online](#)).
- Effinger, Maria/Maylein, Leonhard/Pietzsch, Eberhard/Spyra, Ulrike: Per Mausklick ins Spätmittelalter. Digitalisierung und Erschließung spätmittelalterlicher Bilderhandschriften aus der Bibliotheca Palatina, in: BIT online 6 (2003), S. 235–247 ([online](#)).

- Effinger, Maria/Maylein, Leonhard/Šimek, Jakob: Von der elektronischen Bibliothek zur innovativen Forschungsinfrastruktur. Digitale Angebote für die Geisteswissenschaften an der Universitätsbibliothek Heidelberg, in: *Bibliothek Forschung und Praxis* 43 (2019), S. 311–323 ([online](#)).
- Effinger, Maria/Saurma-Jeltsch, Lieselotte E./Pietzsch, Eberhard: Deutsche Forschungsgemeinschaft fördert Projekt »Digitalisierung spätmittelalterlicher Bilderhandschriften aus der Bibliotheca Palatina«, in: *Theke* (2000), S. 47–50.
- Fernández Riva, Gustavo/Millet, Victor: ›Verschiedenheit‹ der Handschriften. Über Varianz in der Überlieferung des ›Armen Heinrich‹ Hartmanns von Aue. Mit einer vollständigen Verskonkordanz, in: *ZfdA* (2022), S. 291–321 ([online](#)).
- Holmes, Martin: Whatever happened to interchange?, in: *Digital Scholarship in the Humanities* 32 (2017) [im Advance Access bereits 2016], S. i63–i68 ([online](#)).
- Probst, Veit: Digitization at the Heidelberg University Library. The digital Bibliotheca Palatina project, in: *Digital Philology* 6 (2017), S. 213–233 ([online](#)).
- Saurma-Jeltsch, Lieselotte E./Effinger, Maria: Forschung per Mausclick, in: *Ruperto Carola* 3 (2001), S. 4–12.
- Schmidt, Desmond/Colomb, Robert: A data structure for representing multi-version texts online, in: *International Journal of Human-Computer Studies* 67 (2009), S. 497–514 ([online](#)).
- Stackmann, Karl: Die Edition – Königsweg der Philologie?, in: Bergmann, Rolf/Gärtner, Kurt (Hrsg.): *Methoden und Probleme der Edition mittelalterlicher deutscher Texte*, Tübingen 1993 (Beihefte zu editio 4), S. 1–18.
- Stackmann, Karl: Neue Philologie?, in: Heinze, Joachim (Hrsg.): *Modernes Mittelalter. Neue Bilder einer populären Epoche*, Frankfurt a. M./Leipzig 1994, S. 398–427.

Online-Ressourcen

- arthistoricum.net. Fachinformationsdienst Kunst: <https://www.arthistoricum.net>.
- Bibliotheca Laureshamensis – digital: Virtuelle Klosterbibliothek Lorsch: <https://www.bibliotheca-laureshamensis-digital.de>.
- Bibliotheca Palatina – digital: Virtuelle Rekonstruktion der einst berühmtesten Büchersammlung Deutschlands: <https://digi.ub.uni-heidelberg.de/de/bpd/index.html>.
- Boners *Edelstein* – digital: <https://doi.org/10.11588/edition.bed>.
- CIDOC CRM (Conceptual Reference Model): <https://www.cidoc-crm.org>.
- Der arme Heinrich – digital: <https://doi.org/10.11588/edition.ahd>.
- DWork (Heidelberger Digitalisierungsworkflow): <https://www.ub.uni-heidelberg.de/helios/digi/dwork.html>.

eScriptorium: <https://escripta.hypotheses.org>.

FID4SA (Fachinformationsdienst Südasiens): <https://www.fid4sa.de>.

FRBRoo (Functional Requirements for Bibliographic Records): <https://cidoc-crm.org/frbroo>.

GeoNames: <https://www.geonames.org/>.

GND (Gemeinsame Normdatei): <https://gnd.network>.

heiBOOKS (Heidelberger E-Books): <https://books.ub.uni-heidelberg.de/heibooks>.

heiEDITIONS (Heidelberger digitale Editionen): <https://heieditions.github.io>.

heiUP (Heidelberg University Publishing): <https://heiup.uni-heidelberg.de>.

Iwein – digital: <https://doi.org/10.11588/edition.iwd>.

Kaiserchronik – digital: <https://doi.org/10.11588/edition.kcd>.

Lübisches Recht – digital: <https://doi.org/10.11588/edition.lrd>.

Nikolaus von Jeroschin – digital: <https://doi.org/10.11588/edition.nvjd>.

Propylaeum. Fachinformationsdienst Altertumswissenschaften:
<https://www.propylaeum.de>.

TEI (Text Encoding Initiative): <https://tei-c.org>.

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

Welscher Gast digital: <https://doi.org/10.11588/edition.wgd>.

Anschrift des Autors:

Dr. Jakub Šimek

Universitätsbibliothek Heidelberg

Plöck 107–109

69117 Heidelberg

E-Mail: simek@ub.uni-heidelberg.de