

# Zum gegenwärtigen Stand der digitalen Editorik in der Altgermanistik

## Einige Bemerkungen

Gabriel Viehhauser, Universität Stuttgart

ORCID: 0000-0001-6372-0337

### 1 Einleitung

In wohl kaum einem anderen Anwendungsbereich in den Geisteswissenschaften bietet sich der Einsatz digitaler Mittel derart an wie in der Editorik volkssprachiger mittelalterlicher Texte: Wie die einschlägige Forschung der letzten Jahrzehnte gezeigt hat, sind volkssprachige Texte, wie sie etwa in der Altgermanistik auftreten, notorisch unfest, liegen in unterschiedlichen Varianten und Textgestalten vor und manifestieren sich zudem in einer Überlieferung, deren Materialität und kulturgeschichtlicher Gehalt per se von Relevanz ist (vgl. hierzu etwa grundlegend Stackmann 1964 und Bumke 1996). All das lässt sich oft besser als im Druck im digitalen Medium aufbereiten, welches bekanntermaßen nicht dieselben räumlichen Einschränkungen mit sich bringt wie die Printedition, die zwischen zwei Buchdeckeln Platz finden muss (zusammenfassend zu den Eigenschaften digitaler Editionen vgl. Sahle 2013). Aufgrund der Offenheit und der potentiellen Unbegrenztheit des digitalen Mediums lassen sich in einer digitalen Edition daher nicht nur ein aus dem Reichtum der Überlieferung rekonstruierter Editionstext präsentieren, sondern auch Handschriftenfaksimiles, die den Blick auf die Materialität der Überlieferung eröffnen oder handschriftennahe Transkriptionen, die den Umgang mit dem Text im Laufe seiner Überlieferungsgeschichte widerspiegeln, sowie schließlich auch die synoptische Darstellung mehrerer Textzeugen, die den analytischen Blick auf Fassungsunterschiede und mutmaßliche Tendenzen der Textbearbeitung ermöglichen. Eigentlich, so ließe sich konstatieren, erlaubt gerade erst der Einsatz des Computers eine objektadäquate Darstellung altgermanistischer Texte, die im Printmedium noch gar nicht möglich war.

Vor diesem Hintergrund wäre also zu erwarten, dass sich die digitale Form des Edierens wesentlich auf die mediävistische Editorik ausgewirkt hätte. Gleichwohl haben einschlägige Bestandsaufnahmen der Editionslandschaft immer wieder nahegelegt, dass sich diese Erwartungen kaum erfüllt haben (vgl. zur Mediävistik allgemein etwa Robinson 2005, Porter 2103). Doch wie ist es überhaupt um den gegenwärtigen Stand der digitalen Editorik in der Altgermanistik bestellt? Im Folgenden sollen dazu ein paar Beobachtungen angestellt werden.

### 2 Findable?

Als Ausgangspunkt für solche Überlegungen stellt sich zunächst die Frage, welche digitalen Editionen es in der Altgermanistik derzeit eigentlich gibt bzw. in weiterer Folge, auf welcher Grundlage sich eine solche Frage überhaupt verlässlich beantworten lässt. Kommt man aus dem Digital-Humanities-Umfeld, wird man hierfür wohl zunächst an die beiden Online-Editionskataloge von Franzini et al. (2016–) und Sahle (2020ff.) als potentielle Quellen denken.

Der Katalog von Franzini et al. umfasst derzeit 323 Einträge von digitalen Editionsprojekten aus unterschiedlichen Philologien und Zeiträumen, der von Sahle sogar 871 Datensätze. In beiden Katalogen lassen sich die Ergebnisanzeigen umfassend filtern, etwa nach den Kategorien ‚Period‘ und ‚Language‘, die sich zur Eingrenzung altgermanistischer Editionen eignen sollten. Beschränkt man etwa im Katalog von Franzini et al. die von den Editionsprojekten in den Blick genommene ‚Period‘ auf ‚Middle Ages‘ und ‚Language‘ auf ‚gmh‘ (den ISO-Code für Mittelhochdeutsch), führt dies zu für die Altgermanistik einschlägigen Ergebnissen, jedoch lediglich zu vier Treffern, nämlich dem ‚Parzival-Projekt‘, der Edition der Göttinger Statuten (‚kundige bok

2'), dem „Armen Heinrich – digital“ und schließlich zur MGH-Edition der Konzilschronik des Ulrich von Richental. Ein weiterer Treffer lässt sich durch die Einstellung von ‚ger‘ (dem ISO-Code für Deutsch) als Sprache erzielen, nämlich die Sammlung Schweizerischer Rechtsquellen online (SSRQ online).

Eine etwas umfangreichere Basis bietet der Katalog von Sahle. Hier lässt sich für die vorliegenden Zwecke als Sprache ‚german‘ auswählen, unter ‚Period‘ ‚early‘, ‚high‘ und ‚late ma [= middle ages]‘ sowie (partiell) auch ‚early modern‘. Für ‚early ma‘ findet sich ein Treffer, für ‚high ma‘ finden sich acht und für ‚late ma‘ ganze 23 Einträge. In Sahles Katalog lassen sich die Ergebnisse jedoch noch durch die Kategorie ‚Subject‘ weiter einschränken, indem etwa geschichtswissenschaftliche von literaturwissenschaftlichen Editionen abgegrenzt werden können. Wählt man demnach als Filter ‚literature‘, reduzieren sich die Zahlen deutlich, nämlich auf null Einträge für ‚early‘, sieben Einträge für ‚high‘ und nur mehr vier Einträge für ‚late ma‘, was auf eine gewisse geschichtswissenschaftliche Grundausrichtung des Sahle-Katalogs hindeuten dürfte. Der Wert ‚early modern‘ für ‚Period‘ liefert dazu noch zwei Ergebnisse für das ‚Subject‘ ‚literature‘, das Narragonien-digital-Projekt, daneben aber auch Lessings Übersetzungen, die aber aus dem Rahmen der vorliegenden Fragestellung fallen. Die Zeitspanne ist für die ‚early modern‘-Kategorie also eher zu breit angesetzt, als dass sich trennscharf nach altgermanistischen Projekten suchen ließe.

Blickt man nun noch etwas genauer auf die erzielten Einträge, so zeigt sich rasch, dass bei Sahle sehr unterschiedliche editorische Unternehmungen verzeichnet sind. Neben umfangreichen Editionsprojekten wie etwa dem „Parzival-Projekt“ (und der daraus abgeleiteten Edition des „Rappoltsteiner Parzifal“), dem „Armen Heinrich – digital“, dem „Welschen Gast – digital“, der Editionsplattform zur „Lyrik des deutschen Mittelalters (LDM)“, dem poetischen Werk des Hugo von Montfort und der „Kaiserchronik – digital“, finden sich auch die Editionen bzw. Transkriptionen von Einzelhandschriften (Sankt Galler Nibelungenhandschrift, Berliner Eneas-Handschrift, Rostocker Liederbuch) oder eher Repositorien-artige Textplattformen wie das „Hartmann von Aue-Portal“. Auch das Alter der Projekte variiert stark, bei den älteren Projekten häufen sich entsprechend die ins Nichts führenden toten Links.

Geht man nach den Ergebnislisten der beiden Kataloge, scheint die Bilanz an digitalen altgermanistischen Editionen rein zahlenmäßig also nicht besonders üppig auszufallen; es stellt sich aber natürlich die Frage, ob in den beiden Verzeichnissen überhaupt alle altgermanistischen digitalen Editionen erfasst sind. Was hier auf dem Spiel steht, ist nichts weniger als die Kategorie der Findability, also der Buchstabe F aus dem mittlerweile vermutlich doch weithin bekannten Akronym der FAIR-Richtlinien zur Datenhaltung (Wilkinson et al. 2016), ohne deren versprochene Berücksichtigung wohl kaum ein Projektantrag mehr auskommt. Was die FAIR-Richtlinien dabei genau für Editionen bedeuten, ist derzeit Gegenstand der Diskussion (Gengnagel et al. 2022, Gengnagel et al. 2023).

Für die Findability weisen etwa Gengnagel et al. (2022) darauf hin, dass Editionsprojekte in den oben ausgewerteten Katalogen verzeichnet sein sollten. Hier zeigen sich aber wohl deutlich die Herausforderungen des interdisziplinären Zugangs der digitalen Geisteswissenschaften bzw. insbesondere der digitalen Editorik. Denn anders als Forscher:innen aus den Digital Humanities würden Altgermanist:innen bei einem Verzeichnis von Editionsprojekten wohl nicht in erster Linie an die beiden Kataloge denken, sondern an den „Editionsbericht“ (2023), ein ursprünglich in analoger Form an der Arbeitsstelle „Deutsche Texte des Mittelalters“ und jährlich in der Zeitschrift „Germanistik“ publiziertes Verzeichnis, das mittlerweile über den „Handschriftencensus“ (2023) auch online zugänglich ist (zur Geschichte der Ressource vgl. Wolf 2009 und Breith 2010). Hier sind Stand heute ca. 370 Editionsprojekte zu mittelalterlichen deutschen Texten verzeichnet, die sowohl in analoger als auch in digitaler Form geplant sein können. Der Anteil der digitalen Editionen lässt sich aber nur durch ein ‚close reading‘ der Projektbeschreibungen ungefähr abschätzen, da sich der ‚Editionsbericht‘ in erster Linie als Auflistung versteht und nicht als facettiert durchsuchbare Datenbank; Angaben über die mediale Form der Edition sind dementsprechend fakultativ und nicht in einer eigenen Kategorie verzeichnet.

Bei einer Durchsicht der verzeichneten Einträge lassen sich ca. 40 Editionsprojekte eruieren, die eine digitale Komponente aufweisen dürften. Wie in den beiden Digital-Katalogen sind zum Teil recht unterschiedliche Projekte versammelt, zum Teil auch kleinere Nebenprodukte aus Dissertationen, retrodigitalisierte Editionen, Einzeltranskriptionen und sogar Projekte, bei denen noch nicht ganz entschieden ist, ob eine digitale Komponente beigegeben werden soll. Einige ambitionierte Vorhaben dürften zudem im Planungsstadium stecken geblieben sein und können daher nicht zum gegenwärtigen Stand hinzugerechnet werden.

Der ‚Editionsbericht‘ dürfte damit eher Vollständigkeit erreichen als die beiden Digital-Humanities-Kataloge: alle in den Digitalkatalogen aufgelisteten literaturwissenschaftlichen Projekte, die über Einzelhandschriften hinausgehen, sind auch im ‚Editionsbericht‘ zu finden; zudem begegnen auch einige prominente Projekte, die in den digitalen Katalogen fehlen, wie etwa jene zum österreichischen Bibelübersetzer, zur Wenzelsbibel, Hartmanns Iwein, Boners Edelstein oder das Projekt „Narrative Vermittlung religiösen Wissens. Edition und Kommentierung geistlicher Vers- und Prosatexte des 13. bis 16. Jahrhunderts.“

Ob Daten ‚findable‘ sind, scheint also in gewissem Maß auch Community-spezifisch zu sein. Editionen, die im Gesamtkontext der Digital Humanities ‚unter dem Radar‘ laufen, können im spezifischen Fach durchaus bekannt und auch in einschlägigen Verzeichnissen auffindbar sein. Für einen Bereich wie die digitale Editorik, der eine Zwischenstellung zwischen spezifischer inhaltlicher Domänenausrichtung und allgemeinem Anspruch einnimmt, stellt sich mithin die Aufgabe, Repositorien miteinander zu kombinieren, die von ihrem Generalitätsanspruch her quer zueinander liegen können.

Zudem stellt sich die Frage, ob sich bei einer Bestandsaufnahme von altgermanistischen Editionsprojekten die Grenze zwischen digital und analog immer sinnvoll ziehen lassen kann. Obwohl sich zwischen unterschiedlichen Spielarten in diesem Feld mit guten Gründen unterscheiden lässt (etwa retrodigitalisierte von genuin digitalen Editionen oder bloße Handschriftentranskriptionen von erschließenden Editionen, vgl. hierzu Sahle 2013), wird es vermutlich, wie in den Geisteswissenschaften üblich, an den Rändern der Typologie immer zu Ausfransungen kommen und prototypische Beispiele eher randständigen Phänomenen gegenüberstehen. Das gilt letztlich auch für zeitliche Kategorisierungsfragen, worauf etwa schon die quer zu traditionellen Einteilungen der Altgermanistik liegende Periodisierung von ‚early modern editions‘ im Katalog von Sahle hinweist.

### 3 Infrastrukturen bottom-up?

In dem 2018 von Henning Lobin, Roman Schneider und Andreas Witt herausgegebenen Band über ‚Digitale Infrastrukturen für die germanistische Forschung‘ ist ein Kapitel auch den ‚Digitalen Forschungsinfrastrukturen für die Germanistische Mediävistik‘ gewidmet. Andrea Rapp (2018) hat darin kenntnisreich die durchaus in nicht geringer Zahl vorhandenen Ressourcen zusammengetragen, die Altgermanisten zur Verfügung stehen, und diese auch in Kategorien gegliedert: Neben ‚Quellen und Nachweisinstrumenten‘, ‚Wörterbücher und Sprachatlanten‘ sowie ‚Werkzeugen‘ gibt es auch ein zusammengefasstes Kapitel zu ‚Korpora und Editionen‘, das sich vor allem den sprachwissenschaftlich bedeutsamen Ressourcen wie dem mhd. Textarchiv, den Referenzkorpora Alt- und Mittelhochdeutsch, Deutsch Diachron Digital und auch dem Klassiker TITUS, also dem Thesaurus Indogermanischer Text- und Sprachmaterialien, widmet. Es scheint mir nun auffällig zu sein, dass die Beschreibung der Ressourcen für digitale Editionen in dem Abschnitt demgegenüber lediglich exemplarisch erfolgt, nämlich über die Erwähnung des Berner Parzival-Projekts als digitales Vorzeigeprojekt: Eher als über großangelegte Infrastrukturen lässt sich die altgermanistische digitale Editorik derzeit über Einzelprojekte beschreiben. Dabei wird jedoch meines Erachtens deutlich, dass sich viele dieser Einzelprojekte in letzter Zeit erstaunlich generisch entwickelt haben und damit sozusagen aus vielleicht unerwarteter Richtung, nämlich bottom-up, Infrastruktureffekte mit sich bringen.

Blickt man etwa auf die Ergebnis-Liste aus dem Sahle-Katalog und zieht man die etwas editonsferneren Projekte wie das „Hartmann von Aue-Portal“ oder das Eneit-Faksimile ab, dann lassen sich die Einträge grob in vier Gruppen gliedern: Die erste Gruppe bildet das Parzival-Projekt, das nicht nur an der Edition des Wolfram’schen Romans arbeitet, sondern auch den Rappoltsteiner Parzifal sozusagen *en passant* mitediert und dafür Synergieeffekte mit der eigenen Plattform genutzt hat. Die zweite Gruppe hat sich um den „Welschen Gast“ an der UB Heidelberg formiert: Die Digitalausgaben der Kaiserchronik und des „Armen Heinrich“ nutzen wie der Welsche Gast die Infrastruktur der heiEditions-Plattform, dies gilt ebenso für die nur im ‚Editionsbericht‘ verzeichneten Projekte zu Boners Edelstein, zum Lübischen Recht, zum Gregorius, zum Iwein, zum Erec und zu Nikolas von Jeroschin. Gruppe drei geht aus vom ‚Lyrik des deutschen Mittelalters‘-Projekt (LDM), das zunächst angetreten ist, die Gedichte einzelner Anthologien neu zu edieren, sich mittlerweile aber als umfassende Plattform zur Liederedition etabliert hat, auf der man wohl jedes kurze strophige Textgebilde des deutschen Mittelalters einpassen könnte. Genutzt wird die LDM-Plattform beispielsweise von den Projekten zur Edition der in den Tönen Regenbogens überlieferten Lieder. Schließlich könnte man noch die Hugo-von-Montfort-Edition als frühes Aushängeschild einer vierten Gruppe nennen, die insgesamt aber sozusagen mehr top-down über die Infrastruktur gebildet wird, nämlich die Editionen, die von der GAMS, dem Geisteswissenschaftliches Asset Management System der Uni Graz, gehostet werden. Dieser Plattform bedienen sich weitere, nicht immer nur literaturwissenschaftlich ausgerichtete mediävistische Editionen wie die Wenzelsbibel oder das CoReMa-Projekt zu mittelalterlichen Kochrezepten. Darüber hinaus beherbergt die GAMS (ähnlich wie heiEditions) auch Editionen nicht-mittelalterlicher Texte.

#### 4 Exemplarische Bestandsaufnahme

Alle genannten Projektplattformen sind technisch auf durchaus hohem Niveau, auch was die Reflexion von Nachhaltigkeitsfragen und Standardisierung betrifft. Ich möchte im Folgenden mit dem auf der heiEditions-Plattform veröffentlichten Ausgabe des „Armen Heinrich“ (2018) ein einzelnes Editionsprojekt herausgreifen, um exemplarisch ein Bild zu entwerfen, welche Funktionalitäten und Formen derzeit in altgermanistischen Editionen zur Anwendung kommen können (zur generellen Konzeption von heiEditions vgl. Simek 2023).

Mit drei Handschriften und vier Fragmenten bietet der arme Heinrich eine recht überschaubare Überlieferung, die jedoch ein relativ hohes Maß an Varianz und Textunterschieden aufweist. In den konventionellen, dem Printparadigma folgenden Editionen wurden diese Textunterschiede zumeist zugunsten der Erstellung eines rekonstruierten Editionstextes nivelliert. Die digitale Edition zielt demgegenüber darauf ab, die überlieferungsgeschichtliche Vielfalt des armen Heinrich wieder in ihr Recht zu setzen (vgl. den Einleitungstext zur Edition).

Der Zugang in die sowohl in Hinblick auf Funktionalitäten als auch auf Erstellungsprinzipien auf der Webseite gut dokumentierte Online-Edition ist auf zweierlei Art möglich, nämlich über die Menüpunkte ‚Editionstexte‘ und ‚Handschriften‘. Unter ‚Handschriften‘ finden sich die Überlieferungszeugen des „Armen Heinrich“ in einer Liste zusammengestellt, welche kurze Eckdaten über die jeweilige Handschrift und ein Vorschaubild auf das Faksimile bietet. Klickt man auf einen Eintrag der Liste, gelangt man zu einer generischen Präsentationsansicht, die von der UB Heidelberg für Ihre umfangreichen Handschriftendigitalisate bereits vor der Entwicklung der heiEditions-Plattform eingerichtet wurde. Die Edition nutzt hier also sozusagen die eigene, bestehende Infrastruktur nach. Da der „Arme Heinrich“ aufgrund seiner Kürze nur in Sammelhandschriften überliefert wurde, gerät damit zunächst die gesamte Handschrift mit ihren Kontexten in den Blick. Die Handschriftenfaksimiles sind auf der UB-Plattform zumeist strukturell

nach ihrem Inhalt erschlossen, weshalb man auch jeweils die Einzeltexte einer Handschrift in einer Menüleiste anwählen und sich die entsprechenden Digitalisate aufschlagen lassen kann.<sup>1</sup> Über die normale Darstellung der UB hinaus geht in der Edition nun der Umstand, dass man für die Teile der Handschriften, die den „Armen Heinrich“ enthalten, zusätzlich eine Transkription anzeigen lassen kann, die auf einer TEI-Datei beruht, die auch downloadbar ist. Im Online-Interface kann man sich den Transkriptionstext in unterschiedlich normalisierter Weise anzeigen lassen, auch individuelle Einstellungen sind möglich. Die Transkription lässt sich dem Faksimile der Handschriftenseite gegenüberstellen ebenso wie einer Überblicksansicht, die Thumbnails aller Handschriftenseiten enthält, und einer Spalte mit Metadaten. Alle diese Ansichten sind variabel ein- und ausschaltbar. Schließlich lassen sich noch eigens für die vollständigen Textzeugen erstellte Übersetzungen und, hier wieder nur für den Cpg 341, eine Visualisierung der Lagenstruktur aufrufen.

Die Darstellung über den „Handschriften“-Zugang ist also sehr faksimileorientiert und am Digitalisat ausgerichtet, was der Herkunft der Editionsplattform aus dem System der Heidelberger UB entspricht.

Daneben ist über den ‚Editionstexte‘-Zugang aber auch die Anzeige einer synoptischen Darstellung möglich, bei der mit einer farblichen Markierung korrespondierende Verszeilen in den unterschiedlichen Textversionen deutlich gemacht werden können. Hier lassen sich wieder flexibel einzelne Fenster mit Textversionen zu- und wegschalten sowie schließlich aus der synoptischen Darstellung heraus die Faksimiles aufrufen, deren Anzeige aufgrund des gedrängten Platzes der Synopsenansicht dann aber schon recht unübersichtlich ausfällt.

Unter dem Menüpunkt ‚Materialien‘ lassen sich schließlich noch zusätzliche Informationen abrufen, etwa ein Literaturverzeichnis, aber auch eine Verskonkordanz, die einen groben Blick über die Abweichungen der Textzeugen im Versbestand erlaubt (vgl. hierzu Fernández Riva/Millet 2022).

Der Fokus der Edition richtet sich demnach vor allem auf die möglichst variable Darstellung der Überlieferung sowie auf deren synoptischen Abgleich. Eine geringere Rolle spielt demgegenüber die Herstellung eines kritischen Editionstextes.

Der „Arme Heinrich“ folgt damit einer Blaupause, die sich auch für andere Editionsprojekte nutzen lässt, etwa für die im Aufbau befindliche digitale Iwein-Edition. Auch diese wird durch den doppelten Zugang von Handschriftenkatalog und synoptischer Editionsansicht erschlossen werden, allerdings liegen hier die Transkriptionen noch nicht vor. Als Zusatzmaterialien wird auch für den Iwein eine Verskonkordanz angeboten, zudem lässt sich jetzt schon eine weitergehende Auswertung abrufen, die der „Arme Heinrich“ nicht bietet, nämlich eine Netzwerkdarstellung der Mitüberlieferung des Textes.

## 5 Perspektiven

„Der arme Heinrich – digital“ demonstriert damit also recht eindrücklich, welche Funktionalitäten in einer digitalen altgermanistischen Edition sinnvoll zum Einsatz gebracht werden können. Die Edition ist um Zugänglichkeit und eine ausführliche Dokumentation bemüht (über einen allerdings etwas versteckten Button auf der Seite lässt sich sogar eine englische Version der Editionsplattform abrufen), Aspekte der Nachhaltigkeit sind ebenso berücksichtigt wie eine potentielle Generalisierbarkeit der Interface-Lösung.

---

<sup>1</sup> Meine Beschreibung hier orientiert sich am ‚Idealfall‘, welcher in dem der UB Heidelberg zugehörigen Cpg 341 (Sigle Ba) gegeben ist. Jene Handschriften der Überlieferung, die aus anderen Bibliotheken und Portalen eingebunden sind, können in der Edition nur für den Hartmann-Teil präsentiert werden und sind nicht ähnlich tief strukturell erschlossen. De facto betrifft dies aber ohnedies nur Bb (Cod. Bodmer 72, Genf Coligny): Die dritte vollständige Handschrift (A, olim Straßburg, Stadtbibliothek), ist verbrannt und nur über den (im Portal vollständig präsentierten) Myller’schen Abdruck zugänglich, die restliche Überlieferung besteht aus Fragmenten.

Vor diesem Hintergrund stellt sich nun die Frage, ob und welche Perspektiven für digitale Editionslösungen in der Altgermanistik noch darüber hinaus entworfen werden können oder überhaupt sollten. Solch neue Perspektiven ergeben sich meiner Ansicht nach insbesondere dann, wenn man – wie in der Theoriediskussion zur digitalen Editorik immer wieder gefordert (vgl. z. B. Robinson 2004, Buzetti 2009, Boot/van Zundert 2011) – Editionen unter den digitalen Bedingungen noch radikaler neu denkt, als dies bislang zumeist der Fall ist. Dieses Neudenken drängt sich insbesondere dann auf, wenn man sich vor Augen hält, dass ein wesentliches Charakteristikum digitaler Editionen ihre Entgrenzung darstellt bzw. die Möglichkeit, im digitalen Medium alles mit allem zu verbinden. Diese Möglichkeit macht nun nicht an den traditionellen Werkgrenzen halt, etwa, indem man alle handschriftlichen Instanzen eines Textes miteinander verlinkt und in synoptische Relation stellt, sondern sie greift auch über die Grenzen des Werkes hinaus. Aus dem altgermanistischen Bereich haben beispielsweise jüngst Christian Griesinger und Michael Stolz diesen Sachverhalt in seiner Konsequenz zu Ende ausformuliert:

Als Konsequenz dieser Überlegungen ergibt sich, dass die Möglichkeiten und Perspektiven des digitalen Mediums unser Verständnis von Editionen verändern: Betrachtete man früher Editionen als isolierte Leistungen in der Regel einzelner Herausgeber, so sind sie heute kollektive Errungenschaften, die zum Teil von mehreren Arbeitsstellen in Kollaboration ediert werden. Setzt sich dieser Trend in Zukunft fort und gehen wir den nächsten logischen Schritt in der Vernetzung von Editionen mit anderen Ressourcen, dann werden digitale Editionen wichtige Teile komplexerer Forschungsumgebungen werden. [...] Wird diese Entwicklung in aller Konsequenz weitergeführt und geben Editionen ihre isolierte Stellung zukünftig auf, so werden sie in größere Zusammenhänge eingebettet – die Editionen werden Teil eines umfassenderen Textkorpus. (Griesinger/Stolz 2019, 117)

Wie könnte nun ein Workflow aussehen, der zu einer solchen Art von neuen Editionen beiträgt, bei denen die Grenzen zwischen Edition und Textkorpus verschwimmen? Ich möchte im Folgenden einige Stationen anführen, die sich in einen solchen Workflow einfügen lassen könnten. Dabei werde ich notwendigerweise stellenweise Bekanntes wiederholen, möchte aber insbesondere einige Aspekte aufgreifen, die mir für zukünftige Arbeit relevant erscheinen. Dabei soll der Workflow modular gedacht werden, die einzelnen Stationen stellen also Bausteine dar, die nicht notwendigerweise immer zum Einsatz kommen müssen und die durch weitere Bausteine ergänzt werden können. Die Liste kann und soll dementsprechend gar nicht Vollständigkeit beanspruchen und versteht sich als Diskussionsgrundlage.

## 5.1 Handschriftenerkennung

Lässt man Fragen der durch digitale Kataloge wesentlich erleichterten Handschriftenheuristik und der Bilderschließung beiseite, so stellt die Texterkennung einen ersten wichtigen, aber auch aufwändigen Schritt zur digitalen Edition dar. Durch die in den letzten Jahren stark verbesserten Deep-Learning-Verfahren haben sich die Möglichkeiten zur Automatisierung dieses Schritts stark verbessert (gängige Systeme stellen etwa Transkribus, OCR4all und eScriptorium dar). Einige der aktuellen Editionsprojekte in der Altgermanistik machen davon Gebrauch (beispielsweise das Projekt zum österreichischen Bibelübersetzer). Die automatische Handschriftenerkennung steht und fällt als maschinelles Lernverfahren mit der Menge an Trainingsdaten, die generiert werden können. Unter dem Aspekt der Nachhaltigkeit erscheint es daher von besonderer Relevanz, dass die in Editionsprojekten erarbeiteten Trainingsmodelle für die Allgemeinheit verfügbar gemacht werden, um so gute generische Modelle zu entwickeln, die gewinnbringend auch auf bislang untrainierte Handschriften angewandt werden können. Für die nächsten Jahre sollten hier erhebliche Synergie-Effekte zu erwarten sein.

Automatische Erkennungsverfahren werden freilich nie völlig fehlerfrei bleiben, doch lässt sich durch die schiere Menge, die zumindest einigermaßen korrekt erfasst werden kann, das Konzept

der Texterschließung insgesamt neu denken.<sup>2</sup> Denn anders als man es von philologisch genauen Editionen oder auch Handschriftentranskriptionen kennt und begreiflicherweise schätzt, könnte es etwa bei der Aufbereitung umfangreicher Überlieferungen eine durchaus pragmatische Lösung darstellen, ein gewisses Maß an Unschärfe zuzulassen. Dahinter steckt folgende Überlegung: als Grundlage für einen herkömmlichen Editionstext müsste das Ergebnis der automatischen Handschriftenerkennung manuell nachbearbeitet werden, was bei kleineren Editionsprojekten zwar aufwendig ist, aber immer noch eine erhebliche Arbeitserleichterung bedeutet. Im Fall umfangreicher Überlieferungen kann dieser Aufwand jedoch bald zu groß werden, ist aber vielleicht auch nicht immer nötig, denn dort wird man die Überlieferung ohnedies vorsortieren wollen und nicht jeden Textzeugen im Einzelwortlaut wiedergeben. In vielen Fällen ist es gar nicht von prioritärer Relevanz, die eher zufällige Varianz auf Wortebene in den Blick zu bekommen, sondern geht es eher um das Aufspüren von großräumigen Textabweichungen und die Gruppierung der Überlieferung. Zur bloßen Gruppierung der Handschriften reichen die automatisch erschlossenen Transkriptionstexte aber vermutlich durchaus aus.

Die digitalen Methoden ermöglichen hier also eine radikal neue Erschließungsform von Texten, die auf große Textmengen statt Details abzielt. Neben den kritischen Editionstext und den diplomatischen Handschriftenabdruck tritt damit gleichermaßen eine dritte Form, die unscharfe Transkription. Und auf dieselbe Weise, wie sich die beiden Formen Handschriftenabdruck und Edition nicht ausschließen müssen, sondern einander zuarbeiten können, kann auch die grobmaschige Erschließung der Überlieferung mit digitalen Methoden schließlich in eine Edition münden; allerdings lässt sich vorher besser abschätzen, welche Überlieferungsgruppen tatsächlich für die genauere editorische Darstellung in Frage kommen.

## 5.2 Gruppierung der Handschriften und Texte

Mithilfe digitaler Methoden eröffnet sich die Möglichkeit, (teil-)automatisch verwandte Handschriftengruppierungen zu erkennen. Recht ausgiebig erprobt wurden in diesem Kontext z. B. phylogenetische Methoden, die in der Bioinformatik zur Berechnung verwandter Genom-Sequenzen zum Einsatz kommen (Howe et al. 2001). Daneben wären aber auch andere Methoden, die in diesem Zusammenhang noch nicht eingesetzt wurden (etwa aus dem Information Retrieval oder der Text-Reuse-Forschung) einer näheren Untersuchung wert. Zu denken wäre hier etwa an die Berechnung von Cosinus-Distanzen zur Ermittlung der Dokumentähnlichkeit oder der Edit-Distance auf n-Gramm-Basis bzw. Word Embeddings zur Ermittlung bloß semantischer Übereinstimmungen.

Mit Hilfe dieser oder der phylogenetischen Methoden könnten sich selbst verworrene Überlieferungen einer Sortierung zuführen lassen, die sich nicht unbedingt, wie in der traditionellen Stemmologie üblich, auf ein genealogisches Modell beziehen muss, sondern schlicht Ähnlichkeiten von Versionen in den Blick bekommt, die sich für die Auswertung anbieten. Als Basis für eine solche Abschätzung von Ähnlichkeiten wären dabei möglicherweise bereits die im vorigen Abschnitt beschriebenen unscharfen Transkriptionen ausreichend.

Das Konzept der Ähnlichkeit ließe sich darüber hinaus aber ebenfalls über die Werkgrenzen hinweg denken: Die Untersuchung der Unterschiede zwischen zwei Textversionen muss nicht auf variante Textüberlieferungen beschränkt bleiben, sondern kann in einem als Textkorpus gedachten Verständnis von Editionen nominell unterschiedliche Texte betreffen (etwa die unterschiedlichen Versionen hagiographischer Legenden). Die Grenzen zwischen Textvarianz und Intertextualität ließen sich hier also überblenden.

---

<sup>2</sup> Ich führe hier einen Gedanken aus, der auf dem Bremer Workshop 2019 „Rudolf von Ems edieren“ in der Diskussion entstanden ist und für den ich folglich keine Urheberschaft beanspruchen kann und will.

### 5.3 Normalisierung und Lemmatisierung

Als Vorbereitung für solche automatischen Kollationen und auch für andere Arten der digitalen Auswertung größerer Textkorpora kann ein Preprocessing der Texte dienen, etwa die Berücksichtigung von Methoden zur automatischen Normalisierung und Lemmatisierung. Mit diesen Methoden könnte ein auf digitale Wörterbücher referenzierbarer Text erstellt werden, der sich auch für die wichtige Korpusbildung vergleichbarer mittelhochdeutscher Texte einsetzen ließe. Auch hier haben sich in jüngerer Zeit durch den Einsatz von Deep-Learning-Verfahren sowie durch die Verfügbarkeit von annotierten Korpora neue Möglichkeiten ergeben. Eine vielversprechende Perspektive für automatische Normalisierung und Lemmatisierung sowie Wortarten-Tagging bietet etwa der von Helmut Schmid (2019) erstellte RNN-Tagger, der auch ein auf dem Referenzkorpus Mittelhochdeutsch (Klein et al. 2016) trainiertes Modell für das Mittelhochdeutsche bereithält.

### 5.4 Community-basierte Standards (Quelldaten)

Die langfristige Speicherung der in einer Edition erstellten Quelldaten ist zwar wesentlich leichter zu bewerkstelligen als etwa die nachhaltige Zurverfügungstellung von Interfaces und deren Funktionalität, sie ist aber nach wie vor eine Forschungsaufgabe, die sich insbesondere bei der Bereitstellung von mittelalterlichen Texten stellt. Zwar gibt es die bekannten Standards für die Texterschließung (etwa die TEI), jedoch geben diese nur Rahmen vor, die z. B. von einzelnen mittelhochdeutschen Editionsprojekten unterschiedlich gefüllt werden. Von Nöten ist daher ein Diskussionsprozess über die Kodierung von Phänomenen, die insbesondere für mittelalterliche deutsche Textzeugen für eine interoperable Basis sorgen (z. B. bei der Kodierung von Korrekturen und Abkürzungszeichen). Dies ist keine rein technische Fragestellung, sondern erfordert den Einbezug der altgermanistischen Community. Dieser Einbezug der Community beschränkt sich dabei nicht nur auf den Dialog mit den Expert:innen auf dem Gebiet der digitalen Editorik, sondern umfasst auch einen Vermittlungsaspekt, da die im Dialog mit der Community erstellten Best-Practice-Richtlinien distribuiert und so in die Community zurückgespielt werden.

### 5.5 Semantic Enrichment

Neben der linguistischen Annotation von Texten mit Lemmata und POS-Tagging spielt die darüberhinausgehende Anreicherung von Editionen mit semantischen Annotationen eine immer größere Rolle. Dies können zum Teil noch strukturelle Informationen wie die Kennzeichnung von Reimwörtern und Metrum sein, aber auch etwa narrative Strukturmerkmale wie Erzählebenen oder Ereignisse, vor allem aber Entitäten, die in einem Text vorkommen, also etwa Orts- und Personennamen.

Die Auszeichnung solcher Entitäten kann von Hand geschehen, wünschenswert ist aber auch hier eine zumindest teilweise Automatisierung, etwa durch Methoden der Named Entity Recognition. Mit Hilfe solcher teilautomatischen Verfahren können größere Korpora erstellt werden, die der Community zur Verfügung gestellt werden sollen und, als Trainingsdaten, wiederum dazu genutzt werden können, entsprechende automatische Systeme zu verbessern.

Die Auszeichnung von Entitäten ermöglicht nicht nur die schnelle Erstellung von Indizes, sondern auch die Verlinkung über die einzelne Edition hinaus. Dazu ist es wichtig, die einzelnen Entitäten mit Normdaten abzugleichen, also mit Identifiern zu versehen, die im Idealfall die Verbindung zum Linked Open Data (LOD) des Semantic Web ermöglichen. Im Bereich von LOD spielt wiederum die Aufarbeitung der Daten mit Hilfe von standardisierten Ontologien bzw. Begriffshierarchien eine besondere Rolle.



## 5.6 Editions- und Präsentationsplattformen

Editionsplattformen sollten eine interoperable Präsentation ermöglichen, das heißt nicht nur die Editionstexte darstellen, sondern auch den Download der Quelldateien bzw. darüber hinaus deren maschinelle Abfrage durch APIs ermöglichen. Während der erste Punkt des Quelldatendownloads von den avancierteren der oben genannten Editionen abgedeckt wird, stellt das fehlende Bewusstsein für die Bedeutung von APIs einen großen Schwachpunkt der altgermanistischen (ebenso wie der allgemeinen) Editionslandschaft dar.

Bei Editions- und Präsentationsplattformen stellt sich das Problem der Nachhaltigkeit im besonderen Maße, da Funktionalitäten eines Interfaces zumindest in der Praxis zumeist nicht in derselben standardisierten Form beschrieben werden wie etwa Quelldateien, die mit TEI ausgezeichnet sind. Zwar lassen sich etwa die Verknüpfungen, die zwischen einzelnen Texttranskriptionen bestehen, im Quelltext auszeichnen, der Code, der die Präsentation etwa einer Textsynopse bewirkt, geht aber über den Quelltext hinaus und stellt meist einen Bestandteil dar, der als durchaus integral und essentiell für eine digitale Edition empfunden wird (vgl. zur Diskussion über die Priorität von Daten oder Interfaces Bleier et al. 2018).

Grundsätzlich gilt, dass je idiosynkratischer eine Digitaledition mit Funktionalitäten ausgestattet ist, sich die Erhaltung dieser Funktionalitäten desto schwieriger gestaltet. Abhilfe könnten hier also möglichst generische Lösungen schaffen. Freilich ergibt sich hier das übliche Problem bei Standards, dass man sich zumeist in ein Prokrustesbett eingezwängt fühlt, das genau nicht jene Funktionalitäten bereithält, die man gerade für seinen Spezialfall haben will. Standardisierung verhindert damit bis zu einem gewissen Grad Innovation (vgl. hierzu auch die bei Hausmann 2022 abgebildete Diskussion). Trotzdem ließe sich darüber diskutieren, ob man gerade für eine recht abgegrenzte Domäne wie altgermanistische Editionen bestimmte Funktionalitätsblöcke identifizieren könnte, die sich modular zum Aufbau eines generischen Interfaces verwenden lassen. Mögliche Funktionsblöcke wären etwa die variable synoptische Beiseitstellung von Textfenstern, die Gegenüberstellung von Transkription und Faksimile oder auch die Ein- und Ausblendung von handschriftlichen Normalisierungen und Besonderheiten.

## 5.7 Textanalyse und Visualisierung

Von der Frage nach den Editionsinterfaces ist es nur ein kleiner Schritt zu jener nach der Integration von Analyse- und Visualisierungstools, die im Sinne eines Distant-Reading-Ansatzes dabei helfen können, einen Überblick über die Inhalte einer Edition zu erlangen bzw. Tools zur Exploration bieten.

Möglich wären hier z. B. Netzwerkanalysen (wie sie sich etwa schon am Beispiel der Iwein-Digital-Ausgabe angedeutet haben) oder stärker computerlinguistisch ausgerichtete Verfahren wie Kollokationsanalysen und Topic Modelling bzw. der Einsatz von Distanzmaßen und Word Embeddings zur Aufdeckung von Textunterschieden bzw. -ähnlichkeiten oder für den Varianzabgleich. Bei der Ausweitung der Edition mit Hilfe dieser Verfahren stellt sich die Frage, ob eine proprietäre Integration der Tools in spezifische Editionsplattformen die beste Lösung darstellt. Flexibler und auch nachhaltiger dürfte der modulare Einsatz von Analysetools sein, die nicht nur für eine Edition, sondern generisch zum Einsatz gebracht werden können. Für eine solche modulare Architektur wird die bereits im vorigen Punkt genannte Einrichtung von APIs besonders wichtig.

## Referenzen

- Bleier, Roman, Martina Bürgermeister, Helmut W. Klug, Frederike Neuber und Gerlinde Schneider. Hrsg. 2018. *Digital Scholarly Editions as Interfaces*. Norderstedt: Books on Demand.
- Boot, Peter und Joris van Zundert. 2011. „The Digital Edition 2.0 and the Digital Library: Services, Not Resources.“ *Bibliothek und Wissenschaft* 44, 141–152.

- Breith, Astrid. 2010. „Mittelalter-Philologie im Internet. 36. Beitrag: Der Editionsbericht – Melde- und Schnittstelle für Editionsprojekte zu Mittelalterlichen Deutschen Texten.“ *ZfdA* 139, 132–133.
- Bumke, Joachim. 1996. „Der unfeste Text. Überlegungen zur Überlieferungsgeschichte und Textkritik der höfischen Epik im 13. Jahrhundert.“ In *„Aufführung“ und „Schrift“ in Mittelalter und Früher Neuzeit*, hrsg. von Jan-Dirk Müller, 118–29. Stuttgart/Weimar: Metzler.
- Buzzetti, Dino. 2009. „Digital Editions and Text Processing.“ In *Text Editing, Print, and the Digital World*, hrsg. von Marilyn Deegan und Kathryn Sutherland. Aldershot, Ashgate, 45–62.
- Editionsbericht. 2023. Editionsprojekte zu mittelalterlichen deutschen Texten. <https://editionsbericht.de>.
- Fernández Riva, Gustavo und Victor Millet. 2022. „‘Verschiedenheit‘ der Handschriften. Über Varianz im Versbestand in der Überlieferung des *Armen Heinrich* Hartmanns von Aue – mit einer vollständigen Verskonkordanz.“ *ZfdA* 151, 291–321.
- Franzini, Greta, Peter Andorfer und Ksenia Zaytseva. 2016. *Catalogue of Digital Editions: The Web Application*. <https://doi.org/10.5281/zenodo.1250796>.
- Gengnagel, Tessa, Frederike Neuber und Daniela Schulz. 2022. „Criteria for Reviewing the Application of FAIR Principles in Digital Scholarly Editions, Version 1.1.“ *RIDE: A Review Journal for Digital Editions and Resources*. <https://web.archive.org/web/20230120155729/https://ride.i-d-e.de/fair-criteria-editions/>.
- Gengnagel, Tessa, Frederike Neuber und Daniela Schulz. 2023. „EDITORIAL: FAIR Enough? Evaluating Digital Scholarly Editions and the Application of the FAIR Data Principles.“ *RIDE: A Review Journal for Digital Editions and Resources*. 16. <https://doi.org/10.18716/ride.a.16.0>.
- Griesinger, Christian und Michael Stolz. 2019. „Sprachwissenschaftliche Erschließungsmethoden für digitale Editionen mittelhochdeutscher Texte.“ *Das Mittelalter* 24 (1), 112–128.
- Handschriftencensus 2023. Eine Bestandsaufnahme der handschriftlichen Überlieferung deutschsprachiger Texte des Mittelalters. <https://handschriftencensus.de/>.
- Hartmann von Aue. 2018–2023. *Der arme Heinrich*. Textgeschichtliche elektronische Ausgabe, hrsg. von Gustavo Fernández Riva und Victor Millet, unter Mitarbeit von Jakub Šimek, mit Übersetzungen von Dietmar Peschel, Heidelberg: Universitätsbibliothek. <https://doi.org/10.11588/edition.ahd>.
- Hausmann, Albrecht. 2022. „Digitale Edition (Diskussionsbericht Sektion 2).“ In *Digitale Mediävistik. Perspektiven der Digital Humanities für die Altgermanistik*, hrsg. von Elisabeth Lienert, Joachim Hamm, Albrecht Hausmann und Gabriel Viehhauser. Oldenburg (Beiträge zur mittelalterlichen Erzählforschung 12), 77–78. <https://doi.org/10.25619/BmE20223197>
- Howe, Christopher J., Adrian C. Barbrook, Matthew Spencer, Peter Robinson, Barbara Bordalejo und Linne R. Mooney. 2001. „Manuscript Evolution.“ *TRENDS in Genetics* 17 (3), 147–152.
- Klein, Thomas, Klaus-Peter Wegera, Stefanie Dipper und Claudia Wich-Reif. 2016. Referenzkorpus Mittelhochdeutsch (1050–1350), Version 1.0, <https://www.linguistics.ruhr-uni-bochum.de/rem/>.
- Porter, Dot. 2013. „Medievalists and the Scholarly Digital Edition.“ *Scholarly Editing. The Annual of the Association for Documentary Editing* 34. [http://www.scholarlyediting.org/2013/essays/essay\\_porter.html](http://www.scholarlyediting.org/2013/essays/essay_porter.html).
- Rapp, Andrea. 2018. „Digitale Forschungsinfrastrukturen für die Germanistische Mediävistik. allen händen, die sich zum anbau dieses feldes anschicken, ist vollauf arbeit zugeordnet.“ In *Digitale Infrastrukturen für die germanistische Forschung*, hrsg. von Henning Lobin, Roman Schneider und Andreas Witt. Berlin/Boston: De Gruyter. <https://doi.org/10.1515/9783110538663>.

- Robinson, Peter. 2003. "Where We Are with Electronic Scholarly Editions, and Where We Want to Be." *Jahrbuch für Computerphilologie* 5.  
<http://computerphilologie.uni-muenchen.de/jg03/robinson.html>.
- Robinson, Peter. 2005. "Current issues in making digital editions of medieval texts – or, do electronic scholarly editions have a future?" *Digital Medievalist* 1.  
<https://doi.org/10.16995/dm.8>.
- Sahle, Patrick. 2013. *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels* (3 Bde.), Norderstedt: Books on Demand.
- Sahle, Patrick. 2020ff. *A Catalog of Digital Editions*. <https://www.digitale-edition.de>.
- Schmid, Helmut. 2019. "Deep Learning-Based Morphological Taggers and Lemmatizers for Annotating Historical Texts." In *DATECH, May 2019, Brussels, Belgium*. <https://www.cis.uni-muenchen.de/~schmid/papers/Datech2019.pdf>.
- Šimek, Jakub. 2022. „heiEDITIONS – eine Heidelberger Infrastruktur für Editionen (nicht nur) mittelalterlicher Texte.“ In *Digitale Mediävistik. Perspektiven der Digital Humanities für die Altgermanistik*, hrsg. von Elisabeth Lienert, Joachim Hamm, Albrecht Hausmann und Gabriel Viehhauser, 27–46. Oldenburg (Beiträge zur mittelalterliche Erzählforschung 12).  
<https://doi.org/10.25619/BmE20223193>.
- Stackmann, Karl. 1964. „Mittelalterliche Texte als Aufgabe.“ In *Festschrift für Jost Trier zum 70. Geburtstag*, hrsg. von William Foerster und Karl Heinz Borck, 240–267. Köln/Graz: Böhlau.
- Wilkinson, Mark D. et al. 2016: "The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data." *Sci Data* 3. <https://doi.org/10.1038/sdata.2016.18>.
- Wolf, Jürgen. 2009. „Wer ediert was, wo und wie? Der ‚Editionsbericht‘ als papiernes und elektronisches Hilfsmittel für den Editor.“ In *Wege zum Text. Überlegungen zur Verfügbarkeit mediävistischer Editionen im 21. Jahrhundert*, hrsg. von Wernfried Hofmeister und Andrea Hofmeister-Winter, 229–240. Tübingen: Niemeyer.