

Fachartikel aus:

Daniela Schulz / Marcus Baumgarten / Torsten Schaßan (Hg.): Digitales Edieren gestern, heute und morgen (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 7). 2026. DOI: [10.17175/sb007](https://doi.org/10.17175/sb007)

Titel:

Genauigkeit: Überlegungen zu Edition, digitaler Edition und künstlicher Intelligenz

Autor*in:

Georg Vogeler

Kontakt: georg.vogeler@uni-graz.at

Institution: Universität Graz

GND: [1035297884](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63884-p0000-00002-1726-1712) ORCID: [0000-0002-1726-1712](https://orcid.org/0000-0002-1726-1712)

Contribution (CRediT): [Conceptualization](#) | [Writing – original draft](#)

DOI des Beitrags:

[10.17175/sb007_007](https://doi.org/10.17175/sb007_007)

Nachweis im OPAC der Herzog August Bibliothek:

[1973150832](#)

Erstveröffentlichung:

12.08.2026

Lizenz:

Sofern nicht anders angegeben 

Letzte Überprüfung aller Verweise:

17.06.2026

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Künstliche Intelligenz](#) | [Digitale Edition](#) | [Maschinelles Lernen](#) | [Genauigkeit](#)

Empfohlene Zitierweise:

Georg Vogeler: Genauigkeit: Überlegungen zu Edition, digitaler Edition und künstlicher Intelligenz. In: Daniela Schulz / Marcus Baumgarten / Torsten Schaßan (Hg.): Digitales Edieren gestern, heute und morgen (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 7). Wolfenbüttel 2026. 12.08.2026. HTML / XML / PDF. DOI: [10.17175/sb007_007](https://doi.org/10.17175/sb007_007)

Georg Vogeler

Genauigkeit: Überlegungen zu Edition, digitaler Edition und künstlicher Intelligenz

Abstract

Der Beitrag diskutiert das Konzept der *Genauigkeit* als Prinzip wissenschaftlichen Edierens an der Schnittstelle traditioneller Editionsprinzipien, digitaler Arbeitsabläufe, *maschinellern Lernen* und *generativer KI*. Ausgehend von drei Kernprinzipien, der Einhaltung von Standards, dem Vetorecht der Quelle und wissenschaftlicher Sorgfalt, zeichnet er ihre historische Entstehung in Editionsrichtlinien und methodologischen Texten seit dem 19. Jahrhundert nach und erörtert den Einsatz *großer Sprachmodelle* in der digitalen Edition sowie Strategien zur Verringerung möglicher Fehler (»LLM as a judge«, »Expert in the loop« usw.). Argumentiert wird, dass die Editionswissenschaft ein qualitatives Verständnis von Genauigkeitsmaßen, ihren Verzerrungen und den Grenzen ihrer Übertragbarkeit entwickeln muss. Der Beitrag schließt mit der Forderung nach einer neuen digitalen Fehlerkultur, in der wir Genauigkeit als modellabhängig und statistisch beschreibbar und Sorgfalt als kritisch reflektierte Praxis verstehen.

The article examines the concept of *accuracy* in scholarly editing at the intersection of traditional editorial principles, digital workflows, *machine learning*, and *generative AI*. It takes as its starting point three core maxims – adherence to standards, the source's veto power, and scholarly meticulousness – and traces their historical formation in editorial guidelines and methodological writings since the 19th century. It discusses the use of *large language models* in digital editing and the strategies to mitigate errors (»LLM as a judge«, »Expert in the loop« etc.). The article argues that editorial scholarship must develop a qualitative understanding of accuracy measures, their biases, and their limits of transferability, and concludes by calling for a new digital error culture in which accuracy is understood as model-dependent and statistically describable and as meticulousness conceptualized as a critically reflected practice.

1. Einleitung

Dieser Text beruht auf einer Keynote mit dem Titel *Regeln und Praktiken digitalen Edierens – Zur Rolle maschinellen Lernens in der Editionsphilologie*, die ich im Herbst 2023 gehalten habe. Seitdem hat sich viel getan im Einsatz künstlicher Intelligenz in digitalen Editionen.¹ Schon diese Formulierung ist Ergebnis der Entwicklungen, denn 2023 waren Textgeneratoren, die heute als treibende Kraft künstlicher Intelligenz verstanden werden, Teil des Forschungsfelds maschinellen Lernens. Noch hatten wenige die Fähigkeit großer Sprachmodelle, aus wenigen Beispielen und mit natürlichsprachlichen Anweisungen Verfahren generalisieren zu können (*Few-Shot-Learning* und *In-Context-Learning*) und eine Vielzahl von Aufgaben zu übernehmen, identifiziert oder gar auf digitales Edieren angewendet.² Das Thema des damaligen Vortrags hat sich aber nur verschoben. Schon im Herbst 2023 wurde diskutiert, ob automatische Verfahren überhaupt richtige Ergebnisse liefern können. Ich möchte hier den Gedankenfaden weiterspinnen, den ich 2023 noch unter dem Eindruck von *HTR* (*Handwritten Textrecognition*) und *NER* (*Named Entity Recognition*) als Trainingsaufgaben maschinellen Lernens begonnen habe. Wie verändern sich die Praktiken des Edierens mit den verschiedenen digitalen Verfahren?

[1]

Die Praxis des Edierens unter dem Apell »Ediere richtig!« wird in drei Maximen realisiert, die auf eine Übereinstimmung mit in der Forschungsdiskussion entwickelten Standards, eine Übereinstimmung mit der Überlieferungsrealität und auf Verantwortlichkeit zielen:

[2]

1. Ediere so wie alle anderen! (Standards)
2. Ediere nichts, was nicht belegt ist! (Vetorecht der Quelle)

¹ Vgl. z. B. zum Stand von 2024: Pollin et al. 2025 und eine eher zufällige Auswahl einschlägiger Publikationen: Beshero-Bondar 2023; Schonhardt 2025; Laeger / Möller 2025; Henzel 2025 und natürlich auch den Beitrag von Gerrit Brüning (2025) in diesem Band.

² Vgl. Brown et al. 2020; Wei et al. 2022.

3. Ediere nichts, was Du nicht (wissenschaftlich) verantworten kannst, und scheue keine Mühe! (Akribie)

Ich möchte im Folgenden diese Maximen in der digitalen Edition unter dem Begriff der *Genauigkeit* reflektieren. Diese kann man unterschiedlich verstehen. Im Sinne der ersten und zweiten Maxime heißt Genauigkeit, dass Editor*innen nachvollziehbare und dem Text angemessene Terminologien und Kategorien verwenden sollen, also *präzise* und *exakt* zu sein. Sie sollen auch *genau* in dem Sinne sein, dass sie anerkannte Repräsentationen bilden, die dem Wissensstand über Paläographie, Kodikologie, Grammatik, Philologie und Geschichte nicht widersprechen, also *richtig* sein. Schließlich sind Editor*innen mit feinen Unterschieden und komplexen Zusammenhängen konfrontiert, die an jeder Stelle sorgfältig und gewissenhaft zu beurteilen sind. Begriffe wie Sorgfalt oder Gewissenhaftigkeit rücken deshalb in das Blickfeld. Der Beitrag ist geleitet von einer Skizze, wie Editor*innen sich seit dem 19. Jahrhundert zwischen Standards, regelgeleiteter Verallgemeinerung, individuellen Forschungsfragen und Problemstellungen in der Praxis positioniert haben.

[3]

2. Vordigitale Algorithmen: Editionsprinzipien

Edition ist eine »question del metodo«³ und diese schlägt sich in den Regelwerken und Kategoriensystemen der Editionswissenschaft nieder. Eine für die Editionswissenschaft spezifische Textgattung sind die Editionsprinzipien, die individuell einzelnen Editionen vorangestellt werden oder als Vereinbarungen von Personengruppen oder Orientierung von Institutionen für ihre Editionsreihen veröffentlicht worden sind. So hat z. B. Heinz Angermeier im Vorwort des Abteilungsleiters zum ersten Band der Reihe *Deutsche Reichstagsakten – Reichsversammlungen 1556–1662* die Editionsprinzipien der Reihe vorgestellt.⁴ Regionale und nationale historische Vereine haben seit Anfang des 20. Jahrhunderts Editionsrichtlinien publiziert, die inzwischen auch als Handreichungen im Internet zu finden sind.⁵ Die britische Pipe Roll Society hatte schon in der Mitte des 19. Jahrhunderts die Regulierung sogar so weit getrieben, dass sie Vorgaben für Drucktypen, die sogenannte *Record Type*, gemacht haben, um die Befunde in den Vorlagen präzise abbilden zu können.⁶ Diese Regelwerke formulieren sehr konkrete Abbildungsfunktionen von Befunden in den Vorlagen. Es geht dabei meistens um paläographische Befunde, aber auch philologische Probleme wie Interpunktion oder Notation von textkritischen und inhaltlichen Beobachtungen. Genauigkeit ist hier die konsistente und korrekte Anwendung der Regeln.

[4]

Diese Regelwerke folgen dem Muster »Wenn Du X beobachtest, dann schreibe Y in der Edition« und sind also algorithmisch formuliert. Ihnen stehen methodische Texte von Karl Lachmann⁷ oder Joseph Bédier⁸ und Einführungen in die Editionstechnik von Paul Maas oder David C. Greetham gegenüber.⁹ Sie beschreiben Vorgehensweisen und begründen die philologischen Beobachtungen. Sie sind also weicher formuliert als die Regelwerke. Die im Jahr 2000 erschienene *Ars edendi* von R.B.C. Huygens, deklariert das Edieren als Kunst (»Editing texts is an art«),¹⁰ vermeidet Regeln ganz und weist ausdrücklich auch auf persönliche Eigenschaften der Editor*innen hin. Sie sollen fähig sein, mit Ausnahmen umgehen zu können. Sie sollen

[5]

³ Pratesi 1992.

⁴ Vgl. Angermeier 1988.

⁵ Diese etwas lange Fußnote möge als ein bibliografischer Beitrag zu dieser Textgattung gelten, die eine systematische Auswertung verdient hätte: Arbeitsgemeinschaft außeruniversitärer historischer Forschungseinrichtungen 1981; Besch 1994; Besch 1973 / 1975; Beyer 2018; Carli 1967; Cencetti 1957; Commission Internationale de Diplomatique 1984; Commission Royale d'Histoire 1985; Consejo Superior de Investigaciones Científicas 1944; Ecole Nationales des Chartes 2001; Feigl 1955; Heckmann 2013; Heinemeyer / Gesamtverein der Deutschen Geschichts- und Altertumsvereine 1978; Heinemeyer 1978; Kuttner 1959; Masai 1950; Norme generali per la pubblicazione 1902; Norme per la stampa 1906; Pratesi 1964; Schultze 1930; Schultze 1966; Schultze 1978; Schwäbische Forschungsgemeinschaft und Schwäbische Forschungsstelle Augsburg der Kommission für bayerische Landesgeschichte bei der Bayerischen Akademie der Wissenschaften 2001; Wolff 1957.

⁶ Vgl. Condon / Hallam 1984.

⁷ Nicht als eigenständiger methodischer Text veröffentlicht, sondern als Teil einer Edition: Lachmann 1850, S. 3–15.

⁸ Vgl. Bédier 1928a; Bédier 1928b.

⁹ Vgl. Maas 1960; Greetham 1994.

¹⁰ Huygens 2000, S. 21.

»talent« besitzen oder »familiar with« sein.¹¹ Huygens nennt sogar psychologische Risiken ihrer Tätigkeit, wenn er von editorischen Lösungen spricht, die »striking« sind – und damit verführerisch und deshalb riskant oder gar falsch.¹² Genauigkeit ist hier eher eine Funktion von persönlichem Gespür und Intuition.

3. Algorithmisches Vorgehen in digitalen Editionen

Die Praxis digitalen Edierens baut besonders auf der Abbildungsfunktion der Editionsregeln auf, denn im Digitalen entfallen die Beschränkungen der Druckedition für die Kodierung philologischer Beobachtungen. Die Arbeit der **Text Encoding Initiative (TEI)** zielt so auf ein ausdifferenziertes Kodierungssystem für eine große Bandbreite philologischer Phänomene, die Textstrukturen, textuelle Varianz, paläographische Beobachtungen, Textredaktion oder inhaltliche Kommentierung digital abbilden. Eine andere Art der Abbildungsfunktion ist die Suche der digitalen Edition nach formalen Abbildungsregeln von Buchstabenbefunden und Kodierung. Die **Medieval Unicode Font Initiative (MUFI)** arbeitet an dafür angemessenen Zeichensätzen. Abbildungsregeln sind auch Teil der hyperdiplomatischen Editionsverfahren von Helmut Klug¹³ oder der XProc-Pipeline von Gioele Barabucci und Elisa Cugliana zur Transformation einer hyperdiplomatischen Transkription in eine normalisierte Edition.¹⁴ Sie erwecken die Vorstellung, dass sich eindeutige Abbildungsfunktionen zwischen Textvorlage und digitaler Präsentation zumindest im Bereich der Paläographie aufstellen lassen.

[6]

Die Eindeutigkeit ist jedoch nicht gegeben, wenn die Regeln verallgemeinert werden. Die TEI arbeitet, wie die oben erwähnten Editionsregelwerke, darauf hin, Abbildungsregeln editionsübergreifend zu formulieren. Standardisierte Regelwerke müssen jedoch weit genug für viele Einzelfälle sein und verringern damit Genauigkeit. Die TEI bietet z. B. explizit fünf verschiedene Möglichkeiten an, die Funktion eines kritischen Apparats zu kodieren.¹⁵ Textuelle Varianz kann aber auch mit weiteren Lösungen kodiert werden: choice, sourceDoc, @cor resp sind alles Kodierungsformen, für die man gute Gründe finden kann, sie anzuwenden, um Textalternativen zu kodieren, auch wenn mit jeweils unterschiedlicher Semantik. Lou Burnard hat deshalb darauf hingewiesen, dass die TEI nur die Grundlagen für eine Spezialisierung der Kodierungsregeln in einer Edition bildet.¹⁶

[7]

Eine ähnliche Diskussion wird für die Anzeige der Kodierung geführt. Kann die Kodierung mit eindeutigen Abbildungsregeln in eine Benutzungsoberfläche überführt werden? Das *TEI Processing Model* beschreibt z. B., wie man ein Regelwerk speichern könnte, das die Abbildung von bestimmten TEI-Kodierungen auf HTML-Elemente ermöglicht.¹⁷ TEI Boilerplate¹⁸ und CETEItan¹⁹ sind technologisch andere Lösungen mit dem gleichen Ziel. Helmut Klug und Selina Galka haben unter der Bezeichnung »Minimal Edition« Verfahren vorgestellt, die hyperdiplomatische Transkriptionen in Lesetexte umwandelt.²⁰ Alle diese technischen Verfahren sind Ergebnis dessen, dass Regelwerke prinzipiell als Algorithmen ausgedrückt werden können. Sie sind aber alle auf die Bedingungen einer individuellen Edition anzupassen. In der TEI werden diese individuellen editorischen Praktiken z. B. in einem *TEI ODD* (One Document Does It All) dokumentiert. Die Einbindung des TEI Processing Models in TEI ODD zeigt nun, dass auch die Individualisierung in der digitalen Edition als algorithmisches

[8]

¹¹ Huygens 2000, S. 21.

¹² Huygens 2000, S. 50.

¹³ Vgl. Böhm / Klug 2020.

¹⁴ Vgl. Cugliana / Barabucci 2021.

¹⁵ Vgl. TEI 2026; zur Codierung textueller Varianten [hier](#).

¹⁶ Vgl. Burnard 2019.

¹⁷ Vgl. Cummings et al. 2015; Turska et al. 2016.

¹⁸ Vgl. Walsh et al. 2012–2023.

¹⁹ Vgl. Cayless / Viglianti 2018.

²⁰ Vgl. Galka / Klug 2023.

Problem verstanden werden kann. Es entsteht also zusätzlich zum Algorithmus zur Abbildung der Textvorlage in der Edition, ein Algorithmus, der die rein digitalen Editionspraktiken abbildet (»Wenn Du X kodiert hast, dann zeige Y an.«).

Man kann diese Algorithmisierung weitertreiben und versuchen, für alle Schritte des digitalen Edierens Algorithmen zu finden. Tara Andrews hat z. B. 2019 vorgeschlagen, mehr als nur die Transformation von kodierter Annotation in die Anzeige als Algorithmus aufzufassen. Sie unterscheidet dabei zwischen kategorischen und individuellen Urteilen: kategorische Urteile sind systematische Repräsentationen, die in Iterationen und Mustern ausgedrückt werden können – und damit der klassische Ort für algorithmische Verarbeitung. Individuelle Urteile werden gefällt, wenn der*die Editor*in einen einzelnen Datenpunkt erstellt oder ändert. Die Entscheidung für die eine oder die andere Lesart eines Textes in den verschiedenen Überlieferungen kann also kategorisch fallen, wenn man sich z. B. dafür entscheidet, im Zweifelsfall der älteren Überlieferung den Vorrang zu geben. Sie fällt individuell, wenn man jede einzelne Textstelle prüft und entscheidet. Dieser zweite Vorgang wird nun dadurch algorithmisch, dass er sich im Sinne der Software-Versionierung als Patch zum bestehenden Datensatz ausdrücken lässt. Praktische Auswirkung hat das insbesondere für wechselseitige Abhängigkeiten der Algorithmen. Die Edition, die **T-Pen** als Transkriptionswerkzeug verwendet, kann mit den Verfahren der automatisierten Integration von Code-Änderungen (Continuous Integration) verarbeitet werden. Jede Transkriptionsänderung im **T-Pen**-Frontend erzeugt JSON und TEI-Repräsentationen, die für Tokenisierung, automatische Kollation mit **CollateX** und stemmatologische Analyse mit **StemmaWeb** ausgewertet werden kann.²¹ Hier werden die individuellen editorischen Entscheidungen zusammen mit allgemeinen Algorithmen zusammengeführt. Die zunächst nicht algorithmischen individuellen Entscheidungen einer Edition werden damit auch zu einfachen Befehlszeilen. Es entsteht eine Art *Meta-Algorithmus* der digitalen Edition, der nicht mehr die editorischen Entscheidungen selbst, sondern die Organisation editorischer Arbeit im Digitalen abbildet. Editionsarbeit gewinnt hier Genauigkeit, weil detailliert reproduzierbar.

[9]

In der digitalen Edition ist also zum heutigen Stand der Forschung die Idee erhalten geblieben, *richtig* im Sinne von Regelmäßigkeit zu edieren. Allgemeine Regelwerke wie die Guidelines der TEI können mehr Fälle kodieren, als die der vor-digitalen Zeit. Editorische Aktivitäten können nicht nur als kategorische Verfahren, sondern auch als individuelle Änderungen der Datengrundlage in Algorithmen ausgedrückt werden. Dieses Forschungsprogramm der Algorithmisierung von Edition ist noch nicht abgeschlossen, wie z. B. die Formalisierungsvorschläge für die Versionierung von Editionen von Martina Bürgermeister belegen.²² Genauigkeit entsteht in all diesen Verfahren durch die korrekte Formulierung und Anwendung des Algorithmus.

[10]

4. Meta-Algorithmen und Wahrscheinlichkeiten: Maschinelles Lernen

Gleichzeitig werden aber seit ein paar Jahren auch Verfahren diskutiert, die Dominique Stutzmann 2019 auf dem IMC (International Medieval Congress) als »Distant Editing« bezeichnet hat. Er und ich haben in der Diskussion darüber dann eine Parallele dieser Verfahren zur 2017 von Mike Kestemont, Vincent Christlein und Dominique Stutzmann postulierten »Künstlichen Paläographie«²³ gezogen. Verfahren des maschinellen Lernens ermöglichen es inzwischen, viele Aktivitäten der Edition an Maschinen zu delegieren, die damit eine *künstliche Edition* erzeugen. Tobias Hodel hat in den vergangenen Jahren mehrfach auf das Potential dieser Entwicklungen hingewiesen.²⁴ Melissa Terras, Joe Nockels, Sarah Ames, Paul Gooding, Andy Stauder und

[11]

²¹ Vgl. Andrews et al. 2019.

²² Vgl. Bürgermeister 2025.

²³ Kestemont et al. 2017.

²⁴ Vgl. Hodel 2023.

Günter Mühlberger sehen die Angebote von Transkribus als Werkzeug zur Automatisierung des Edierens.²⁵ Ich selbst habe versucht, die künstliche Edition in ein konzentrisches Modell zwischen Proto-Editionen aus Bildern und Beschreibungen und vollständig von Menschen überprüften Editionen einzuordnen.²⁶

Die künstliche Edition beruht auf einem anderen Meta-Algorithmus als der Abbildung von Änderungen in der Datenbasis und ihrer Integration in deterministische Prozeduren, nämlich auf den Algorithmen des maschinellen Lernens. Dabei werden statistische Muster aus Datenvorlagen extrahiert und daraus in den einzelnen Knoten des neuronalen Netzes Gewichte für Verarbeitungsfunktionen abgeleitet. Diese erlauben wahrscheinliche Vorhersagen zu Daten, die nicht Teil des Trainingsmaterials waren. Diese Lernverfahren sind schon lange im Bereich der optischen Zeichenerkennung von gedruckten Texten und der linguistischen Annotation im Einsatz. Als neuronale Netze sind sie im Bereich der digitalen Paläographie schon 2009 vorgeschlagen worden,²⁷ haben aber erst mit den Erfolgen im Bereich der automatischen Handschriftenerkennung (Handwritten Text Recognition, HTR) im Rahmen der **READ-Coop** und dem **Transkribus-Service** breite Aufmerksamkeit im Feld der Editorik gefunden. [12]

Maschinelles Lernen ermöglicht die Imitation von menschlichen Tätigkeiten beim Edieren auch jenseits der Transkription. Inhaltliche und sprachliche Annotation einer digitalen Edition kann mit den zunehmend verfeinerten Methoden der Informationsextraktion und des *Part-of-Speech-Taggings* von maschinell gelernten Funktionen übernommen werden.²⁸ Im Bereich der automatischen Stemma-Erstellung haben sich die Verfahren des *DeepLearnings* noch nicht durchgesetzt. Für Kollation sind Diff-Algorithmen und für Stemma-Erstellung phylogenetische Methoden Stand der Forschung, auch wenn Armin Hoenen und Laura Mancini maschinelles Lernen als Werkzeug zur Erzeugung von Stemmata angeregt haben.²⁹ [13]

Diese Verfahren führen nun eine Eigenschaft in das Edieren ein, die mit dem eingangs angesprochenen Prinzip der *Genauigkeit* in Konflikt gerät, nämlich die Wahrscheinlichkeit. Die Verfahren des maschinellen Lernens sind alle Annäherungen an Regeln, die eine Vorhersage möglichst wahrscheinlich machen. Künstliche Editionen sind also wahrscheinlich genau, also in einem mehr oder weniger gut messbaren Grad genau. Damit sind sie aber grundsätzlich anders als die im vorangehenden Abschnitt beschriebenen digitalen kritischen Editionen, die ihre Genauigkeit durch deterministische Algorithmen und eindeutige Abbildungsregeln in Datenstandards zu erreichen versuchen. [14]

Wir brauchen also als Editor*innen und Benutzer*innen von Editionen ein neues Verständnis von Genauigkeit als statistischem Maß. Transkriptionen waren schon immer Interpretationen der Editor*innen. Sie waren jedoch relativ leicht zu verstehen, denn es gab eine Abbildungsregel von 1–n paläographischen Phänomenen auf 1 Kodierung / Notation (Buchstabe im Zeichensatz des Setzers). Die Abbildungsregel war angereichert durch interpretative Anmerkungen, die Beobachtungen verbalisierten, die von der Regel abwichen. Diese grundsätzliche Abbildungsregel war in den Transkriptionsprinzipien dokumentiert und in den publizierten generellen Editionsrichtlinien den Leser*innen zugänglich. Eine »Explikation der Auslegung ›Transkription«³⁰ im Sinne von Thomas Stäcker war also relativ einfach. Im digitalen Edieren wird diese Explikation durch ein digitales Bild in der Edition noch weiter erleichtert. [15]

Maschinelle Verfahren informieren dagegen über das Verhältnis zwischen Paläographie und Transkription in drei Formen: In der Benennung des verwendeten Modells, der für das Training verwendeten Referenzdaten (*Ground Truth*) und in einem Maß für die Genauigkeit des Modells, meistens der *Character Error Rate (CER)*.³¹ [16]

²⁵ Vgl. Terras et al. 2024.

²⁶ Vgl. Vogeler 2025.

²⁷ Vgl. Fusi 2009.

²⁸ Zu diesen Verfahren vgl. z. B. Mitkov 2022.

²⁹ Vgl. Hoenen 2022; Mancini 2022.

³⁰ Vgl. Stäcker 2020.

³¹ Vgl. zur CER Neudecker et al. 2021.

Aber was heißt eigentlich eine ›CER 2,3 %‹? Man möge sich die Zahlen einmal konkret vorstellen, um ein Gespür für die so beschriebene Genauigkeit zu bekommen: 2,3 Buchstaben von 100 entsprechen nicht den im Evaluationsdatensatz vorgegebenen Werten, also ist jeder 50. Buchstabe falsch, durchschnittlich mindestens in jeder 2. Zeile ein Fehler. Hier stellt sich als erstes die Frage, mit wieviel Fehlern können wir als Editor*innen in einer Transkription leben. 2018 haben Olga Beloborodova, Wout Dillen und Joshua Schäuble in einer Transkribus-Konferenz einmal eine Genauigkeit von 95 % als schlechtes Ergebnis für eine*n Editor*in postuliert und deshalb eine manuelle Überprüfung automatischer Transkriptionen eingefordert.³² Gleichzeitig können auch fehlerhafte OCR-Daten noch zu hinreichend guten Ergebnissen von Folgeaufgaben wie der Erkennung von benannten Entitäten führen.³³

Eine solche CER ist jedoch ein Durchschnittswert. Eine gute Bewertung dieser Genauigkeit muss sich der Vielfältigkeit der Fehler stellen. Mehrere Fehler können sich natürlich in einem Wort häufen. Auch Beschädigungen der Vorlage oder Passagen in einer Hand oder Schrift, die nicht zum Trainingsmaterial gehörte, führen zu einer Häufung der Fehler. Die Editor*innen müssen sich auch inhaltlicher Verzerrungen bei der Interpretation bewusst sein, wie z. B. dass sich die Fehler häufiger in Eigennamen verbergen.³⁴ Editor*innen müssen eine CER also qualitativ auswerten können. [17]

Die CER ist aber aus einem ganz anderen Grund noch eine ungewohnte Neuerung für die Editor*innen: Sie misst die Genauigkeit an einem bekannten Datenausschnitt, nicht die Genauigkeit der Anwendung des Modells auf unbekanntes Material. Das ist ja auch eigentlich nicht zu messen. Trotzdem wird sie in der Editionspraxis einerseits als Werkzeug eingesetzt, um Arbeitseinsatz in die Verbesserung der Trainingsdaten oder des HTR-Modells zu begründen, ohne dass der Effekt dieser Genauigkeit auf neue Bilder bekannt wäre. Wenn man sie als Indikator verwenden will, um für eine automatische Transkription von bislang ungekannten Texten ein gutes Modell auszuwählen, dann müssen die Editor*innen also die Ähnlichkeit zwischen Trainingsmaterial des Modells und Anwendungsfall ermitteln. [18]

Die Frage der Standards lebt in Zeiten des maschinellen Lernens neu auf in der Frage, ob die in den unterschiedlichen Projekten geschaffenen Transkriptionen mit ihren unterschiedlichen Zielen zu einer gemeinsamen Trainingsbasis zusammenkommen können. Das CATMUS-Projekt schlägt z. B. eine graphemische Transkription vor, die z. B. ›u‹ und ›v‹ auf ›u‹ reduziert.³⁵ Ist diese Vereinfachung der Genauigkeit im Sinne der Editor*innen zuträglich, weil sie das Trainingsmaterial und damit die Wahrscheinlichkeit einer korrekten Transkription erhöht, oder ist sie eine Reduktion der Genauigkeit, weil weniger Allographen abgebildet werden als in den Vorlagen verwendet? [19]

Eine Diskussion auf der jüngsten DHd in Wien über die Maße *Accuracy*, *Precision*, *Recall* und *F1* ist hier ebenso erhellend.³⁶ Die Maße beschreiben Genauigkeit als Anteil aller Vorhersagen an den erwarteten Werten (*Accuracy*), als positiver Vorhersagewert einer gesuchten Unterscheidung (*Precision*), als Sensitivität auf ein gesuchtes Phänomen (*Recall*) und als das gewichtete harmonische Mittel von *Precision* und *Recall* (*F1*). Bei der automatischen Identifikation von Raumkonstruktionen in romantischer Literatur ist die überwiegende Zahl der Wörter in einem Text irrelevant. Eine Klassierung in sechs inhaltliche Klassen und eine Klasse ›irrelevant‹ führt also automatisch zu einer ›Accuracy‹ jenseits von 99 %, weil der Anteil an korrekt als irrelevant markierten Wörtern immer sehr hoch ist. Stellen wir uns also Annotationen von sechs Raumbegriffen (›urban‹, ›ländlich‹, ›Innenraum‹, ›Gewässer‹, ›Gebirge‹ und ›Wald‹) in romantischer Literatur vor, dann ist offensichtlich, dass die überwiegende Zahl der Sätze keine solche Raummetaphorik verwenden. Sie werden in automatischer Annotation dementsprechend einfach als ›irrelevant‹ erkannt. Die wenigen mehr [20]

³² Vgl. Beloborodova et al. 2018.

³³ Vgl. Hamdi et al. 2023.

³⁴ Vgl. Ströbel et al. 2022; Romein et al. 2025.

³⁵ Vgl. Pinche et al. 2023; Clérice et al. 2024.

³⁶ In der Diskussion über das Doktoratsprojekt von Nils Kellner (Rostock): Automatische Erkennung von Spatial Frames und Emotionen in deutschen und spanischen Romanen der Romantik. Vgl. Kellner 2026.

oder weniger korrekten Zuweisungen zu den sechs Raumbegriffen fallen kaum noch ins Gewicht für die Berechnung einer allgemeinen »Accuracy«. Erst die Genauigkeitsmaße für einzelne Klassen, die das Verhältnis von richtig gefundenen Werten zu den fälschlicherweise einer Klasse zu geordneten Werten (Precision) oder das Verhältnis der gefundenen Werte zu allen erwarteten Werten (Recall) misst, ist dann informativ. Der F1-Wert, der gut eingeführt ist, ist auch nur einer von verschiedenen *F_β-Werten*, die man verwenden kann, um die Treffergenauigkeit eines Modells zu beschreiben, in dem das Verhältnis von Precision und Recall anders gewichtet wird. Andere Evaluationsmetriken fragen nach Aggregationsmöglichkeiten (*Macro / Micro F1*), zeigen an, wo Verwechslungen vorliegen (Confusion Matrix), oder messen Ähnlichkeiten an textlichen Überlappungen (*BLEU (= Bilingual Evaluation Understudy) / SacreBLEU (= shareable, comparable, and reproducible BLEU)*, *ROUGE (= Recall Oriented Understudy for Gisting Evaluation)*, *METEOR (= Metric for Evaluation of Translation with Explicit ORdering)*) oder semantischer Ähnlichkeit (*BertScore*). Lässt sich so editorische Genauigkeit messen?

Zwei Lehren sind aus diesen Entwicklungen für die Editionsphilologie zu ziehen: Wir müssen uns erstens damit beschäftigen, dass mit maschinellem Lernen Modelle trainiert werden, die unsere Benutzungsszenarien abbilden können. Die HTR-Technologien als Automatisierung von Transkription sind dafür das beste Beispiel. Zweitens müssen wir ein Verständnis von den Verfahren und den Maßen gewinnen, mit denen die Genauigkeit dieser Modelle beschrieben wird. [21]

5. Digitale Edition mit generativer KI

Was im Jahr 2026 unter der Bezeichnung *Künstliche Intelligenz (KI)* diskutiert wird, ist eine Weiterentwicklung des maschinellen Lernens, wie sie im vorherigen Abschnitt vorgestellt worden ist. KI entfernt sich aber von den Verfahren, die zum Durchbruch der HTR geführt haben. Der entscheidende Unterschied ist, dass seit Herbst 2022 technische Lösungen der Allgemeinheit zur Verfügung gestellt worden sind, die mit nur wenig Input sinnvolle Zeichensequenzen generieren können – oder anders gesagt: Die auf natürlichsprachliche Anfragen sinnvolle Antworten in Schriftsprache erzeugen können, die sogenannte *generative KI*. Diese Antworten können auch Code oder strukturierte Daten sein. Da die Texterzeugung auf einer sehr großen Menge an Trainingsdaten beruht, können sie Texte erzeugen, die das Wissen aus den Trainingsdaten repräsentieren – oder anders gesagt: Die sogenannten großen Sprachmodelle (*Large Language Model, LLM*) erzeugen Texte, die den Texten, mit denen sie trainiert wurden, ähnlich sind – und zwar auch in inhaltlicher Hinsicht. Die Größe der Sprachmodelle hat nun dazu geführt, dass sie auch für Aufgaben eingesetzt werden können, mit denen man nicht gerechnet hat, ihnen deshalb emergente Fähigkeiten zugeschrieben werden.³⁷ [22]

Seit der Veröffentlichung der ersten allgemeinen Textgeneratoren experimentiert die Editorik mit der neuen Technologie. Einen ersten etwas systematischen Einblick hat der *IDE-Workshop »Generative KI, LLMs und GPT bei digitalen Editionen«* im Frühjahr 2024, der im Rahmen der *DHd 2024* stattfand, gegeben, aus dem Christopher Pollin, Franz Fischer, Patrick Sahle, Martina Scholger und ich verallgemeinernde Ergebnisse gezogen haben.³⁸ Aus der allgemeinen Diskussion über den Einsatz von LLMs in digitalen Editionen schält sich nun zum einen eine ethische Kritik an den Systemen heraus, die auf den unverhältnismäßigen Ressourcenbedarf und die Abhängigkeit von kommerziellen Services hinweisen.³⁹ Ein zweiter Strang der Diskussion zielt auf die Genauigkeit: Die Systeme generativer KI enthalten eine Wahrscheinlichkeitsfunktion, die sprachlich richtige, kontextuell plausible, aber inhaltlich falsche Reaktionen auf den Input im *Prompt* geben können, die sogenannten Halluzinationen. Darauf wird mit vier Strategien reagiert: [23]

1. Reduktion des Wahrscheinlichkeitsraums der Textvorhersage im *temperature*-Parameter des Sprachmodells.

³⁷ Vgl. Wei et al. 2022. Die Anwendbarkeit des Begriffs der »Emergenz« zweifeln an Schaeffler et al. 2023; Krakauer et al. 2025.

³⁸ Vgl. Pollin et al. 2025.

³⁹ Vgl. Ries et al. 2025.

2. Integration des Genauigkeitsanspruchs in den Prompt.
3. Technische Validierungsmechanismen wie die Erzeugung deterministischen Programmcodes statt probabilistischer Vorhersage der Lösung oder Integration von algorithmischen Kontrollschritten, auch unter Einsatz von LLMs selbst (»LLM as a judge«).
4. Verlagerung der Verantwortung an eine*n »Expert in the Loop« (EITL).

Christopher Pollin hat in seinen Vorschlägen zum sogenannten *Promptotyping* eine andere Richtung eingeschlagen. Er zielt darauf, dass Modellierungs-, Kodierungs- und Programmieraufgaben durch eine konsistente Beschreibung der Aufgaben an die Sprachmodelle delegiert werden können, die den dafür nötigen Programmcode erzeugen. Er weist aber auch darauf hin, dass gerade für die Kodierungsaufgaben eine lokale Kontrolle durch Experten nötig ist.⁴⁰ Dieser EITL ist für die Weiterentwicklung editorischer Praxis unter dem Gesichtspunkt der Genauigkeit wohl der wichtigste Ansatz. Egal, welche Verfahren künstlicher Intelligenz eingesetzt werden, greift er die Beobachtung von Huygens auf, nämlich die Fähigkeit des*der Editors*in, überzeugende philologische Entscheidungen treffen zu können, die nicht alleine auf Regeln aufbauen. Der*die Editor*in ist die Person, die dem KI-System einen Auftrag gibt, die für seine Erfüllung relevante Kontexte auswählt, und die Ergebnisse des Systems überprüft. Diese Personen besitzen ganz im Sinne Huygens Erfahrung und Intuition – philologische und technische. Eine wichtige Zukunftsfrage für die Editionswissenschaft ist deshalb, wie diese Expert*innen ausgebildet werden.

[24]

Eine zweite Zukunftsfrage richtet sich an die Praxis der Zusammenarbeit zwischen technischem System und Expert*in. Die technischen Systeme von automatischer Texterkennung, automatischem Textvergleich und automatischer semantischer Annotation, automatischer TEI-Kodierung und Programmierung der Benutzungsumgebung, legen nahe, dass die Expert*innen von Routineaufgaben entlastet werden und sich auf die spannenden Problemfälle konzentrieren können. Unter dem Gesichtspunkt der Genauigkeit rückt die Kategorie der Akribie in den Vordergrund. Studien aus verschiedenen wissenschaftlichen Domänen zeigen, dass die Kombination aus menschlicher Arbeit mit KI-basierten Prozessen zu einer Verstärkung der im Training der KI enthaltenen Annahmen führen kann.⁴¹ Die kognitive Entlastung bei einfachen Aufgaben führt also nicht notwendig dazu, dass die kognitiv anspruchsvolleren Aufgaben besser gelöst werden. Es bedarf einer kritischen Haltung der Expert*innen gegenüber den Ergebnissen der KI, um das Ergebnis zu verbessern.⁴² Ob eine automatische Texterkennung richtig ist oder nicht, bedarf also einer Prüfung, die ein Feingefühl dafür besitzt, welche Arten von Fehlern das System machen könnte – und diese akribisch prüft.

[25]

6. Von der Notwendigkeit einer neuen Fehlerkultur

Es geht unter den Editor*innen natürlich immer noch das Gespenst um, ihre Arbeit würde durch Maschinen ersetzt werden. Der Gedankengang oben führt aber zu einer ganz anderen Perspektive: Edieren war schon immer eine soziale Praxis, ein Teil des Austausches von Menschen über ihre kritischen Erfahrungen mit Texten. Fehler waren und sind eine wichtige Kategorie in dieser Erfahrung. Sie informieren über das Verhalten von Kopistinnen und Kopisten, sind also Grundlage stemmatologischer Editionspraktiken. Fehler sind ein Risiko bei der Verwendung von Textausgaben, wenn sie weniger kritisch angelegt sind als es wissenschaftlichen Standards entspricht. Es ist aber nicht immer leicht, sie zu bestimmen, insbesondere nicht beim Einsatz von Methoden des maschinellen Lernens und generativer KI. Es braucht also eine neue *Fehlerkultur* in der Editorik. Ich stelle mir das als eine Fehlerkultur vor, die Akribie als kollektive Aufgabe versteht, die nicht dem*der einzelnen Editor*in aufgebürdet wird, sondern intersubjektiv ermittelt. Die digitale Fehlerkultur verwendet »99 prozentige Genauigkeit« nicht als Metapher, sondern als ein Modellierungsproblem. Sie versucht damit auf die Frage Antwort zu geben, was da eigentlich berechnet wird. Eine solche digitale Fehlerkultur macht sich über die statistischen und stochastischen Hintergründe solcher

[26]

⁴⁰ Vgl. Pollin 2023.

⁴¹ Vgl. Schroeder et al. 2025.

⁴² Vgl. Beck et al. 2025.

Zahlen Gedanken. Schließlich braucht es aber auch eine Fehlerkultur, die fähig ist, Entscheidungen zu fällen, die von Feingefühl für das editorische Problem, für den Text und für seine Rezeption zeugen. Viele Regeln und Praktiken digitalen Edierens können maschinell gelernt werden. Wir Editor*innen werden uns aber auch in Zukunft über das nötige Feingefühl, die am besten geeignete Statistik und die aufgewendete Akribie unserer Arbeit austauschen, die es für die von uns gewünschte Genauigkeit braucht.

Bibliografie

- Tara Lee Andrews / Anahit Safaryan / Tatevik Atayan: Continuous Integration Systems for Critical Edition: The Chronicle of Matthew of Edessa. In: Digital Humanities Conference 2019. Konferenzabstracts (Utrecht, 09.–12.07.2019). 2019. PDF / HTML. DOI: [10.34894/mz7fbi](#)
- Heinz Angermeier: Vorwort des Abteilungsleiters. In: Der Reichstag zu Speyer 1570. Bearbeitet von Maximilian Lanzinner. 2 Bände. Göttingen 1988. Teilband 1: Protokolle, S. 19–25. [[Nachweis im GVK](#)]
- Arbeitsgemeinschaft außeruniversitärer historischer Forschungseinrichtungen: Empfehlungen zur Edition frühneuzeitlicher Texte. In: Archiv für Reformationsgeschichte – Archive for Reformation History 72 (1981), S. 299–315. DOI: [10.14315/arg-1981-jg12](#)
- Jacob Beck / Stephanie Eckman / Christoph Kern / Frauke Kreuter: Bias in the Loop: How Humans Evaluate AI-Generated Suggestions. arXiv. 10.09.2025. DOI: [10.48550/arXiv.2509.08514](#)
- Joseph Bédier: La tradition manuscrite du Lai de l'Ombre. Réflexions sur l'art d'éditer les anciens textes. In: Romania 54 (1928), no. 214, S. 161–198. [[online](#)]
- Joseph Bédier: La tradition manuscrite du Lai de l'Ombre. Réflexions sur l'art d'éditer les anciens textes. In: Romania 54 (1928), no. 215 / 216, S. 321–356. [[online](#)]
- Olga Beloborodova / Wout Dillen / Joshua Schäuble: CATCH 2020 Project, Manuscript Genetics, Digital Scholarly Editing, Samuel Beckett (1906–1989). In: Transkribus User Conference. Konferenzabstracts (Wien, 08.–09.11.2018). 2018. PDF. [[online](#)]
- Werner Besch: Editionsprinzipien in interdisziplinären Abstimmungen. Annäherungen bei der Herausgabe deutscher Texte der frühen Neuzeit. In: Marlene Nikolay-Panter / Wilhelm Janssen / Wolfgang Herborn (Hg.): Geschichtliche Landeskunde der Rheinlande. Regionale Befunde und raumübergreifende Perspektiven. Georg Droege zum Gedenken. Köln u. a. 1994, S. 467–489. [[Nachweis im GVK](#)]
- Werner Besch: Zur Edition von deutschen Texten des 16. Jahrhunderts. In: Alemannisches Jahrbuch 1973 / 1975, S. 392–411. DOI: [10.57962/regionalia-18935](#)
- Elisa E. Beshero-Bondar: Declarative Markup in the Time of »AI«. In: Balisage: The Markup Conference 2023 (= Balisage Series on Markup Technologies, 28). Proceedings (Washington, 31.07.–04.08.2023). Rockville, US-MD 2023. HTML. DOI: [10.4242/BalisageVol28.Beshero-Bondar01](#)
- Burkhard Beyer: Praktische Tipps für die Edition landesgeschichtlicher Quellen (= Materialien der Historischen Kommission für Westfalen, 15). Münster 2018. [[Nachweis im GVK](#)]
- Astrid Böhm / Helmut W. Klug: Quellenorientierte Aufbereitung historischer Texte im Rahmen digitaler Editionen: Das Problem der Transkription in mediävistischen Editionsprojekten. In: Martin Fischer (Hg.): Digitale Methoden und Objekte in Forschung und Vermittlung der mediävistischen Disziplinen. Akten der Tagung Bamberg, 08.–10. November 2018. Bamberg 2020, S. 51–72. PDF. DOI: [10.20378/irb-56077](#)
- Tom B. Brown / Benjamin Mann / Nick Ryder / Melanie Subbiah / Jared Kaplan / Prafulla Dhariwal / Arvind Neelakantan / Pranav Shyam / Girish Sastry / Amanda Askell / Sandhini Agarwal / Ariel Herbert-Voss / Gretchen Krueger / Tom Henighan / Rewon Child / Aditya Ramesh / Daniel M. Ziegler / Jeffrey Wu / Clemens Winter / Christopher Hesse / Mark Chen / Eric Sigler / Mateusz Litwin / Scott Gray / Benjamin Chess / Jack Clark / Christopher Berner / Sam McCandlish / Alec Radford / Ilya Sutskever / Dario Amodei: Language Models are Few-Shot Learners. arXiv. 28.05.2020. Version 4 vom 22.07.2020. DOI: [10.48550/arXiv.2005.14165](#)
- Gerrit Brüning: Digitale Editionen von Goethes Werken seit 1998. Bilanz und Perspektiven in Zeiten Generativer KI. In: Daniela Schulz / Marcus Baumgarten / Torsten Schaßan (Hg.): Digitale Edieren gestern, heute und morgen (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 7). Wolfenbüttel 2025. 30.12.2025. HTML / XML / PDF. DOI: [10.17175/sb007_002](#)
- Martina Bürgermeister: Intentionen der Änderungen: Ein geisteswissenschaftliches Beschreibungsschema für Versionen digitaler Editionen. In: Nils Reiter / Thomas Haider / Daniel Kababgi / Hendrick Buschmeier / Patrick Helling (Hg.): DHd 2025. Under Construction. 11. Tagung des Verbands Digital Humanities im deutschsprachigen Raum. Konferenzabstracts (Bielefeld, 3.03.–7.03.2025). Bielefeld 2025, S. 330–333. DOI: [10.5281/zenodo.14943258](#)
- Lou Burnard: What is TEI Conformance, and Why Should You Care? In: Journal of the Text Encoding Initiative 12 (2019 / 2020). PDF / HTML. DOI: [10.4000/jtei.1777](#)
- Matilde Carli: Norme tecniche per l'edizione critica delle pergamene pisane dei secoli VIII–XII (= Biblioteca del Bollettino storico Pisano – Collana storica, 5). Pisa 1967. [[Nachweis im GVK](#)]
- Hugh Cayless / Raffaele Vigiante: CETEIcean: TEI in the Browser. In: Balisage: The Markup Conference 2018 (= Balisage Series on Markup Technologies, 21). Proceedings (Washington, 31.07.–03.08.2018). DOI: [10.4242/BalisageVol21.Cayless01](#)
- Giorgio Cencetti: Progetti di unificazione delle norme per la pubblicazione delle fonti medievali. In: Atti del Convegno di studi delle fonti del Medioevo europeo in occasione del 70° della fondazione dell'Istituto Storico Italiano (Rom, 14.04.–17.04.1953). Rom 1957, S. 25–34. [[Nachweis im GVK](#)]
- Thibault Clérice / Ariane Pinche / Malamatenia Vlachou-Efstathiou / Alix Chagué / Jean-Baptiste Camps / Matthias Gille Levenson / Olivier Brisville-Fertin / Federico Boschetti / Franz Fischer / Michael Gervers / Agnès Boutreux / Avery Manton / Simon Gabay / Patricia O'Connor / Wouter Haverals / Mike Kestemont / Caroline Vandeyck / Benjamin Kiessling: CATMuS Medieval: A Multilingual Large-Scale Cross-Century Dataset in Latin Script for Handwritten Text Recognition and Beyond. In: Elisa H. Barney Smith / Marcus Liwicki / Liangrui Peng (Hg.): ICDAR 2024. Document Analysis and Recognition. 18. International Conference. Proceedings (Athen, 30.08.–04.09.2024). Cham, CH 2024, S. 174–194. DOI: [10.1007/978-3-031-70543-4_11](#)
- Commission internationale de diplomatique / Robert-Henri Bautier (Hg.): Normes internationales pour la publication des documents financiers, administratifs et domaniaux du Moyen Age. In: Diplomatica et sigillographica. Travaux préliminaires de la CID et de la commission internationale de sigillographie pour une normalisation internationale des éditions de document et un Vocabulaire international de la diplomatique et de la sigillographie (= Folia Caesaraugustana, 1). Saragossa, ES 1984, S. 75–93. [[Nachweis im GVK](#)]
- Commission internationale de diplomatique / Robert-Henri Bautier (Hg.): Normes internationales pour la publication des lettres et correspondances. In: Diplomatica et sigillographica. Travaux préliminaires de la CID et de la commission internationale de sigillographie pour une normalisation internationale des éditions de document et un Vocabulaire international de la diplomatique et de la sigillographie (= Folia Caesaraugustana, 1). Saragossa, ES 1984, S. 65–74. [[Nachweis im GVK](#)]
- Commission internationale de diplomatique (Hg.): Projet de regles pour l'établissement de notices descriptives de sceaux. In: Diplomatica et sigillographica. Travaux préliminaires de la CID et de la commission internationale de sigillographie pour une normalisation internationale des éditions de document et un Vocabulaire international de la diplomatique et de la sigillographie (= Folia Caesaraugustana, 1). Saragossa, ES 1984, S. 95–109. [[Nachweis im GVK](#)]
- Commission Royale d'Histoire (Hg.): Instructions pour la publication des textes historiques. Brüssel 1985. [[Nachweis im GVK](#)]
- Margaret M. Condon / Elizabeth M. Hallam Smith: Government Printing of the Public Records in the Eighteenth Century. In: Journal of the Society of Archivists 7 (1984), H. 6, S. 348–388. DOI: [10.1080/00379818409514252](#)
- Consejo Superior de Investigaciones Científicas / Escuela de Estudios Medievales (Hg.): Normas de transcripción y edición de textos y documentos. Madrid 1944. [[Nachweis im WorldCat](#)]

- Elisa Cugliana / Gioele Barabucci: Signs of the Times: Medieval Punctuation, Diplomatic Encoding and Rendition. In: Journal of the Text Encoding Initiative 14 (2021). HTML / PDF. DOI: [10.4000/jtei.3715](https://doi.org/10.4000/jtei.3715)
- James Cummings / Sebastian Rahtz / Brian Pytlík-Zillig / Martin Mueller / Magdalena Turska: TEI Simple: Power, Economy, and a Processing Model for Encoders and Developers. In: ADHO – 2015. Global Digital Humanities. Proceedings (Sydney, 29.06.–03.07.2015). 2015. HTML. [\[online\]](#)
- Ecole Nationales des chartes (Hg.): Conseils pour l'édition des textes médiévaux (= Orientations et méthodes). 2 Bände. Paris 2001. [\[Nachweis im GVK\]](#)
- Heinrich Feigl: Die Grundsätze für die Edition der Oberösterreichischen Weistümer. Richtlinien für Editionen. In: Anzeiger der philosophisch-historischen Klasse der Österreichischen Akademie der Wissenschaften 92 (1955), S. 326–340. [\[Nachweis im GVK\]](#)
- Daniele Fusi: Aspects of Application of Neural Recognition to Digital Editions. In: Malte Rehbein / Torsten Schaßan / Patrick Sahle (Hg.): Kodikologie und Paläographie im Digitalen Zeitalter / Codicology and Palaeography in the Digital Age (= Schriften des Instituts für Dokumentologie und Editorik, 2). Norderstedt 2009, S. 175–195. PDF. [\[online\]](#)
- Selina Galka / Helmut W. Klug: Minimal Editing: Die Hyperdiplomatische Transkriptionsplattform. In: Anna Busch / Peer Trilcke (Hg.): DHd 2023. Open Humanities, Open Culture. 9. Tagung des Verbands Digital Humanities im deutschsprachigen Raum. Konferenzabstracts (Belval / Trier, 13.03.–17.03.2023). Trier 2023. DOI: [10.5281/zenodo.7715321](https://doi.org/10.5281/zenodo.7715321)
- David C. Greetham: Textual Scholarship: An Introduction (= Garland Reference Library of the Humanities, 1417). New York 1994. DOI: [10.4324/9780203821220](https://doi.org/10.4324/9780203821220)
- Ahmed Hamdi / Elvys Linhares Pontes / Nicolas Sidere / Mickaël Coustaty / Antoine Doucet: In-Depth Analysis of the Impact of OCR Errors on Named Entity Recognition and Linking. In: Natural Language Engineering 29 (2023), H. 2, S. 425–448. DOI: [10.1017/S1351324922000110](https://doi.org/10.1017/S1351324922000110)
- Dieter Heckmann: Leitfaden zur Edition deutschsprachiger Quellen (13.–16. Jahrhundert). In: Preußenland 3 (2013), S. 7–13. [\[Nachweis im GVK\]](#)
- Walter Heinemeyer / Gesamtverein der Deutschen Geschichts- und Altertumsvereine (Hg.): Richtlinien für die Edition landesgeschichtlicher Quellen. Marburg u. a. 1978. PDF. [\[online\]](#)
- Walter Heinemeyer: Richtlinien für die Edition mittelalterlicher Amtsbücher. In: Walter Heinemeyer / Gesamtverein der Deutschen Geschichts- und Altertumsvereine (Hg.): Richtlinien für die Edition landesgeschichtlicher Quellen. Marburg u. a. 1978, S. 17–28. PDF. [\[online\]](#)
- Katrin Henzel: Textual Scholarship, Artificial Intelligence, Corpora and Intelligent Editions (ESTS 2024). Tagung an der Eötvös Loránd University, Budapest, 2.–4. Oktober 2024. In: editio 39 (2025), H. 1, S. 231–233. DOI: [10.1515/editio-2025-0013](https://doi.org/10.1515/editio-2025-0013)
- Tobias Hodel: Konsequenzen der Handschriftenerkennung und des maschinellen Lernens für die Geschichtswissenschaft. Anwendung, Einordnung und Methodenkritik. In: Historische Zeitschrift 316 (2023), H. 1, S. 151–180. DOI: [10.1515/hzhz-2023-0006](https://doi.org/10.1515/hzhz-2023-0006)
- Armin Hoenen: Can the Language of the Collation be Translated into the Language of the Stemma? Using Machine Translation for Witness Localization. arXiv. 11.06.2022. DOI: [10.48550/arXiv.2206.05603](https://doi.org/10.48550/arXiv.2206.05603)
- Robert Buchard Constantijn Huygens: Ars edendi. A Practical Introduction into Editing of Medieval Latin Texts. Turnhout 2000. [\[Nachweis im GVK\]](#)
- Nils Kellner: Automatische Erkennung von Spatial Frames und Emotionen in deutschen und spanischen Romanen der Romantik. In: Silke Schwandt / Gabriel Viehhauser / Tara Andrews / Thomas Wallnig / Beatrice Nava / Patrick Helling: DHd 2026 Nicht nur Text, nicht nur Daten (DHd2026). Konferenzabstracts (Wien, 23.–27.02.2026). Wien 2026. DOI: [10.5281/zenodo.18696166](https://doi.org/10.5281/zenodo.18696166)
- Mike Kestemont / Vincent Christlein / Dominique Stutzmann: Artificial Paleography: Computational Approaches to Identifying Script Types in Medieval Manuscripts. In: Speculum 92 (2017), H. S1, S. 86–109. DOI: [10.1086/694112](https://doi.org/10.1086/694112)
- David C. Krakauer / John W. Krakauer / Melanie Mitchell: Large Language Models and Emergence: A Complex Systems Perspective. arXiv. 10.06.2025. DOI: [10.48550/arXiv.2506.11135](https://doi.org/10.48550/arXiv.2506.11135)
- Stephan Kuttner: Notes on the Presentation of Text and Apparatus in Editing Works of the Decretists and Decretalists. In: Traditio 15 (1959), S. 452–464. DOI: [10.1017/S0362152900008370](https://doi.org/10.1017/S0362152900008370)
- Stephan Kuttner: [addendum] Notes on the Presentation of Text and Apparatus in Editing Works of the Decretists and Decretalists. In: Traditio 26 (1970), S. 432. [\[online\]](#)
- Karl Lachmann: T. Lucretii Cari De rerum natura libros commentarius. Berlin 1850. [\[Nachweis im GVK\]](#)
- Moritz Laeger / Oke-Lukas Möller: Digitale Quelleneditionen und KI: Aktuelle Tendenzen, Herausforderungen und Probleme. Workshop an der Herzog August Bibliothek Wolfenbüttel, 10. / 11. April 2025. In: editio 39 (2025), H. 1, S. 255–60. DOI: [10.1515/editio-2025-0018](https://doi.org/10.1515/editio-2025-0018)
- Paul Maas: Textkritik. 4. Auflage. Leipzig 1960. [\[Nachweis im GVK\]](#)
- Laura Mancini: An Open Collaborative System for Digital Stemmatology. In: Alessio Maria Braccini / Stefano Za (Hg.): itAIS 2021. Digital Resilience and Sustainability: People, Organizations, and Society. 18. Conference of the Italian Chapter of AIS. Konferenzabstracts (Trient, 15.–16.10.2021). Trient 2021. [\[online\]](#)
- François Masai: Principes et conventions de l'édition diplomatique. In: Scriptorium 4 (1950), Nr. 2, S. 177–193. [\[online\]](#)
- Ruslan Mitkov (Hg.): The Oxford Handbook of Computational Linguistics. 2. Auflage. Oxford 2022. [\[Nachweis im GVK\]](#)
- Clemens Neudecker / Karolina Zaczynska / Konstantin Baierer / Georg Rehm / Mike Gerber / Julián Moreno Schneider: Methoden und Metriken zur Messung von OCR-Qualität für die Kuratierung von Daten und Metadaten. In: Michael Franke-Maier / Anna Kasprzik / Andreas Ledl / Hans Schürmann (Hg.): Qualität in der Inhaltserschließung. Berlin u. a. 2021, S. 137–166. DOI: [10.1515/9783110691597-009](https://doi.org/10.1515/9783110691597-009)
- Norme generali per la pubblicazione dei testi storici per servire alle edizioni della Regia Deputazione di storia patria per le antiche province e la Lombardia. In: Miscellanea di storia italiana 3 (1902), Bd. 7, S. XXXVII–LVI. [\[online\]](#)
- Norme per la stampa delle Fonti per la storia d'Italia. In: Bullettino dell'Istituto storico italiano 28 (1906), S. XI–XXI. [\[online\]](#)
- Ariane Pinche / Thibault Clérice / Alix Chagué / Jean-Baptiste Camps / Malamatenia Vlachou-Efstathiou / Matthias Gille Levenson / Olivier Brisville-Fertin / Federico Boschetti / Franz Fischer / Michael Gervers / Agnès Boutreux / Avery Manton / Simon Gabay / Patricia O'Connor / Wouter Haverals / Mike Kestemont / Caroline Vandyck / Benjamin Kiessling: CATMuS-Medieval: Consistent Approaches to Transcribing Manuscripts: A Generalized Set of Guidelines and Models for Latin Scripts from Middle Ages (8th–16th Century). 2023. Letzte Aktualisierung: 22.11.2024. [\[online\]](#)
- Christopher Pollin / Franz Fischer / Patrick Sahle / Martina Scholger / Georg Vogeler: When It Was 2024 – Generative AI in the Field of Digital Scholarly Editions. In: Zeitschrift für digitale Geisteswissenschaften 10 (2025). 10.07.2025. HTML / XML / PDF. DOI: [10.17175/2025_008](https://doi.org/10.17175/2025_008)
- Christopher Pollin: Workshopreihe »Angewandte Generative KI in den (digitalen) Geisteswissenschaften – AGKI«. Zenodo. 03.11.2023. Version v1.2.0 vom 21.01.2025. DOI: [10.5281/zenodo.10647754](https://doi.org/10.5281/zenodo.10647754)
- Alessandro Pratesi: Norme per la pubblicazione delle fonti documentarie nel »Codice diplomatico Barese«. In: Archivio Storico Pugliese 17 (1964), S. 3–16. [\[Nachweis im GVK\]](#)

- Alessandro Pratesi: Una questione di metodo: L'edizione delle fonti documentarie. In: Alessandro Pratesi: Tra carte e notai: Saggi di diplomatica dal 1951 al 1991 (= Miscellanea della società romana di storia patria, 35). Roma 1992, S. 7–31. [[Nachweis im GVK](#)]
- Thorsten Ries / Tara Andrews / Lisa Rosendahl / Patrick Sahle / Gabriel Viehhauser / Georg Vogeler / Philipp Hegel: Critical AI in der Digitalen Editorik. In: Nils Reiter / Thomas Haider / Daniel Kababgi / Hendrick Buschmeier / Patrick Helling (Hg.): DHd 2025. Under Construction. 11. Tagung des Verbands Digital Humanities im deutschsprachigen Raum. Konferenzabstracts (Bielefeld, 03.03.–07.03.2025). Bielefeld 2025, S. 101–103. DOI: [10.5281/zenodo.14943126](#)
- Anнемieke C. Romein / Andreas Rabus / Gundram Leifert / Phillip Benjamin Ströbel: Assessing Advanced Handwritten Text Recognition Engines for Digitizing Historical Documents. In: International Journal of Digital Humanities 7 (2025), H. 1, S. 115–134. DOI: [10.1007/s42803-025-00100-0](#)
- Rylan Schaeffer / Brando Miranda / Sanmi Koyejo: Are Emergent Abilities of Large Language Models a Mirage? In: Alice Oh / Tristan Naumann / Amir Globerson / Kate Saenko / Moritz Hardt / Sergey Levine (Hg.): Advances in Neural Information Processing Systems 36. 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Proceedings (New Orleans, 10.12.–16.12.2023). New Orleans 2023, S. 55565–55581. PDF. [[online](#)]
- Michael Schonhardt: Die digitale Edition als Schnittstelle: Editorische Workflows im KI-Zeitalter. In: Das Mittelalter 30 (2025), H. 1, S. 54–69. DOI: [10.17885/heiup.mial.2025.1.25120](#)
- Hope Schroeder / Deb Roy / Jad Kabbara: Just Put a Human in the Loop? Investigating LLM-Assisted Annotation for Subjective Tasks. In: Wanxiang Che / Joyce Nabende / Ekaterina Shutova / Mohammad Taher Pilehvar (Hg.): Findings of the Association for Computational Linguistics. ACL 2025. 63rd Annual Meeting of the Association for Computational Linguistics (Wien, 27.07.–01.08.2025). Wien 2025, S. 25771–25795. [[online](#)]
- Johannes Schultze: Grundsätze für die äußere Textgestaltung bei der Herausgabe von Quellen zur neueren Geschichte. In: Korrespondenzblatt des Gesamtvereins der Deutschen Geschichts- und Altertumsvereine 78 (1930), S. 37–45. [[Nachweis im GVK](#)]
- Johannes Schultze: Richtlinien für die äußere Textgestaltung zur neueren deutschen Geschichte. In: Blätter für deutsche Landesgeschichte 102 (1966), S. 1–10. [[Nachweis im GVK](#)]
- Johannes Schultze: Richtlinien für die äußere Textgestaltung bei Herausgabe von Quellen zur neueren deutschen Geschichte. In: Walter Heinemeyer (Hg.): Richtlinien für die Edition landesgeschichtlicher Quellen. Marburg u. a. 1978, S. 25–36. [[Nachweis im GVK](#)]
- Schwäbische Forschungsgemeinschaft / Schwäbische Forschungsstelle Augsburg der Kommission für bayerische Landesgeschichte bei der Bayerischen Akademie der Wissenschaften (Hg.): Editionsrichtlinien für landesgeschichtliche Quellen. Erstellt und beschlossen von der Ausschusssitzung der SFG am 19. Februar 2001. [[online](#)]
- Thomas Stäcker: »A Digital Edition is not Visible« – Some Thoughts on the Nature and Persistence of Digital Editions. In: Zeitschrift für digitale Geisteswissenschaften 5 (2020). 16.10.2020. HTML / XML / PDF. DOI: [10.17175/2020_005](#)
- Phillip Benjamin Ströbel / Martin Volk / Simon Clematide / Raphael Schwitter / Tobias Hodel / David Schoch: Evaluation of HTR Models Without Ground Truth Material. In: Nicoletta Calzolari / Frédéric Béchet / Philippe Blache / Khalid Choukri / Christopher Cieri / Thierry Declerck / Sara Goggi / Hitoshi Isahara / Bente Maegaard / Joseph Mariani / Hélène Mazo / Jan Odijk / Stelios Piperidis (Hg.): Proceedings of the Thirteenth Language Resources and Evaluation Conference (Marseille, 21.–23.06.2022). Marseille 2022, S. 4395–4404. [[online](#)]
- TEI: Guidelines for Electronic Text Encoding and Interchange. P5 Version 4.11.0. Letzte Aktualisierung: 18.02.2026. [[online](#)]
- Melissa Terras / Joe Nockels / Sarah Ames / Paul Gooding / Andy Stauder / Günter Mühlberger: On Automating Editions. The Affordances of Handwritten Text Recognition Platforms of Scholarly Editing. In: Scholarly Editing 41 (2024). DOI: [10.55520/W257A74E](#)
- Magdalena Turska / James Cummings / Sebastian Rahtz: Challenging the Myth of Presentation in Digital Editions. In: Journal of the Text Encoding Initiative 9 (2016). PDF / HTML. DOI: [10.4000/jtei.1453](#)
- Georg Vogeler: Die konzentrische Edition. Auswertungspotentiale von digitaler Tiefenerschließung, Proto-Editionen, künstlichen Editionen und digitalen Volleditionen der Reichshofratsprotokolle. In: Ulrich Rasche / Tobias Schenk (Hg.): Der kaiserliche Reichshofrat. Interdisziplinäre Perspektiven auf Organisation und Funktion eines frühneuzeitlichen Zentralgerichts. (= Quellen und Forschungen zur höchsten Gerichtsbarkeit im Alten Reich, 82). Köln 2025, S. 489–507. DOI: [10.7788/9783412530563.489](#)
- John Walsh / Grant Simpson / Saeed Moaddeli: TEI-Boilerplate / TEI-Boilerplate. GitHub. 2012–2023. Datenset. [[online](#)]
- Jason Wei / Yi Tay / Rishi Bommasani / Colin Raffel / Barret Zoph / Sebastian Borgeaud / Dani Yogatama / Maarten Bosma / Denny Zhou / Donald Metzler / Ed H. Chi / Tatsunori Hashimoto / Oriol Vinyals / Percy Liang / Jeff Dean / William Fedus: Emergent Abilities of Large Language Models. In: Transactions on Machine Learning Research 8 (2022). PDF. [[online](#)]
- Adam Wolff: Projekt instrukcji wydawniczej dla pisanych źródeł historycznych do połowy XVI w. In: Studia Źródłoznawcze 1 (1957), S. 155–184. [[Nachweis im GVK](#)]