

Zeterberg, Max-Ferdinand; Clausen, Lasse; Heyhusen, Lisa-Katharina Vom gedruckten Buch zur digitalen Analyse am Beispiel der Klaus-Mollenhauer-Gesamtausgabe

Oberdorf, Andreas [Hrsg.]: *Digital Turn und Historische Bildungsforschung. Bestandsaufnahme und Forschungsperspektiven*. Bad Heilbrunn : Verlag Julius Klinkhardt 2022, S. 49-64



Quellenangabe/ Reference:

Zeterberg, Max-Ferdinand; Clausen, Lasse; Heyhusen, Lisa-Katharina: Vom gedruckten Buch zur digitalen Analyse am Beispiel der Klaus-Mollenhauer-Gesamtausgabe - In: Oberdorf, Andreas [Hrsg.]: *Digital Turn und Historische Bildungsforschung. Bestandsaufnahme und Forschungsperspektiven*. Bad Heilbrunn : Verlag Julius Klinkhardt 2022, S. 49-64 - URN: urn:nbn:de:0111-pedocs-248525 - DOI: 10.25656/01:24852

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-248525>

<https://doi.org/10.25656/01:24852>

in Kooperation mit / in cooperation with:



<http://www.klinkhardt.de>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

*Max-Ferdinand Zeterberg, Lasse Clausen und
Lisa-Katharina Heyhusen*

Vom gedruckten Buch zur digitalen Analyse am Beispiel der Klaus-Mollenhauer-Gesamtausgabe

Klaus Mollenhauer gilt als einer der bedeutendsten Erziehungswissenschaftler der zweiten Hälfte des 20. Jahrhunderts (vgl. Niemeyer & Rautenberg 2006, 347f.). In den 1950er- und 60er-Jahren prägte er die Sozialpädagogik als erziehungswissenschaftliche Teildisziplin (vgl. Mollenhauer 1959, 1964; Brumlik 1998), später war er (Mit-)Begründer einer kritischen Erziehungswissenschaft, die sozialwissenschaftliche Aspekte der Kritischen Theorie für die Erziehungswissenschaft fruchtbar machen wollte (vgl. Tenorth 1999). Seit den 1980er Jahren öffnete er die Erziehungswissenschaft für kulturgeschichtliche und -theoretische Auseinandersetzungen und fokussierte sich in seinem Spätwerk immer stärker auf Fragen der ästhetischen Bildung (vgl. Dietrich & Müller 2000, 14). Obwohl sein Einfluss auf die Entwicklung der Erziehungswissenschaft in der Disziplin unbestritten ist, lässt sich eine breite Rezeption seiner Veröffentlichungen im aktuellen Fachdiskurs nicht ausmachen. Dies ist auch darin begründet, dass ein Großteil seiner Schriften – wenn überhaupt – nur noch antiquarisch oder weit versprengt über die Bibliothekslandschaft zu beziehen ist.

Die textkritische und kommentierte Klaus-Mollenhauer-Gesamtausgabe (KMG) soll diesem Umstand Abhilfe schaffen und über eine im Internet frei zugängliche Edition die Auseinandersetzung mit Mollenhauers Gesamtwerk in all seinen Facetten ermöglichen. Das Projekt der KMG wird seit April 2018 von der Deutschen Forschungsgemeinschaft (DFG) gefördert und nach zwei dreijährigen Projektphasen im März 2024 abgeschlossen sein. Das KMG-Verbundprojekt besteht aus vier Arbeitsgruppen: An der Georg-August-Universität Göttingen werden unter Leitung von Klaus-Peter Horn Mollenhauers Schriften aus dem Zeitraum 1948–74 sowie sein Nachlass bearbeitet. Eine weitere Arbeitsgruppe an der Universität Osnabrück wird geleitet von Hans-Rüdiger Müller und ist für die Schriften bis 1987 zuständig. Cornelia Dietrich leitet die Arbeitsgruppe an der Humboldt Universität zu Berlin, deren Feld die Schriften von 1988 bis 1998 sowie posthume Veröffentlichungen sind. Zudem befasst sich eine Arbeitsgruppe an der Staats- und Universitätsbibliothek (SUB) Göttingen mit den informationstechnischen und -wissenschaftlichen Aspekten.

In diesem Beitrag soll am Beispiel der KMG dargestellt werden, was es bedeutet, eine digitale Edition zu erstellen und welchen analytischen Nutzen diese für die historische Bildungsforschung haben kann. Hierfür werden zunächst die notwendigen Vorarbeiten sowie die genutzte Software vorgestellt. Im zweiten Abschnitt werden notwendige inhaltliche Entscheidungen des Projekts dargestellt: Die technischen Möglichkeiten einer digitalen Edition sind immens groß, jedoch nicht in jedem Editionsvorhaben gleichermaßen sinnvoll. Abschließend wird ein Verfahren zur computergestützten Analyse der edierten Texte vorgestellt, das den analytischen Mehrwert einer digitalen gegenüber einer gedruckten Edition deutlich macht.

1 Vorarbeiten und „Werkzeug“

Für die Erstellung einer Gesamtedition ist das Auffinden aller Veröffentlichungen der Autor:in unabdingbar. Hier konnte im Rahmen der KMG auf eine schon erstellte Bibliographie der Werke Klaus Mollenhauers zurückgegriffen werden (vgl. Dietrich & Müller 2000), auf deren Grundlage sich weitere Recherchen anschlossen. Insbesondere der Nachlass enthielt zahlreiche Hinweise auf weitere Veröffentlichungen, die bisher unentdeckt geblieben waren. Die ursprüngliche Liste beinhaltete 235 Einträge, inzwischen umfasst die Edition 500 Texte, davon 187 Erstveröffentlichungen. Bis heute werden Wiederveröffentlichungen und kleinere Texte wie etwa Rezensionen in das Textkorpus der KMG aufgenommen.

Nachdem Urheberrechtsfragen mit der Rechteinhaberin Susanna Mollenhauer geklärt wurden, konnten 2015 und 2016 alle Koautor:innen Klaus Mollenhauers mit der Bitte um Zustimmung zur Wiederveröffentlichung ihrer Texte im Rahmen der Edition kontaktiert werden. Es liegt in der Verantwortung des jeweiligen Editionsprojektes, die Kontaktdaten erschöpfend zu recherchieren und diesen Prozess sorgfältig zu dokumentieren, damit mit der Wiederveröffentlichung der Schriften keine Urheberrechtsverletzungen begangen werden.

Die Digitalisierung der Texte ist ein weiterer Arbeitsschritt, bevor sie dann tatsächlich bearbeitet werden können. Um eine automatische Texterkennung (OCR) durchführen zu können, wurden von allen Erstveröffentlichungen Mollenhauers hochwertige Scans angefertigt. Dabei beschränkt sich das Projekt auf Erstveröffentlichungen, da diese Texte als *Leittexte* fungieren und die Basis für die Edition bilden, von der aus alle Abweichungen zwischen verschiedenen Auflagen der Texte abgebildet werden (vgl. Sahle 2017, 238). Leider waren viele Digitalisate, die vor dem Projektbeginn 2018 erstellt worden waren, für eine Texterkennung nicht geeignet. Viele Scans wiesen Wölbungen oder Verzerrungen auf oder waren in zu geringer DPI-Zahl angefertigt worden, weshalb eine OCR sehr ungenau oder unmöglich war. Daher mussten viele Veröffentlichungen ein zweites Mal besorgt und gescannt werden, um schließlich die Einzelseiten als Bilddateien an einen externen Dienst übergeben zu können.

Die Texterkennungssoftware gab die Texte als XML-Dateien aus, in denen jedes einzelne Wort samt Positionierung als eigenes Element ausgezeichnet war. XML steht für *Extensible Markup Language* und ist eine deskriptive Auszeichnungssprache zur Beschreibung von Text und Informationen. Es ist maschinen-, aber bis zu einem gewissen Grad auch menschenlesbar. Dabei hat XML den Charakter einer Metasprache: Sie gibt selbst kein Vokabular vor, sondern lediglich die grammatikalischen Regeln, an die sich die verschiedenen XML-Dialekte halten müssen (vgl. Vogeler & Sahle 2017, 132ff.). Inzwischen sind XML und insbesondere die Richtlinien der *Text Encoding Initiative* (TEI) der de-facto-Standard für digitale Editionen (vgl. Sahle 2017, 245). In der Regel wird auf Basis der TEI-Richtlinien ein spezifisches Metadatenschema entwickelt (ebd., 247).

Das KMG-Team erstellte ein solches Schema mit Blick auf die speziellen Erfordernisse und Ansprüche der Edition von Mollenhauers Texten. Mit einem Script wurde die aus der OCR vorliegende, viel zu kleinteilige Auszeichnung bereinigt, wodurch nur noch relevantes Markup für Seitenumbrüche oder Absätze im Dokument bestehen blieb. Zudem wurden die Dokumente mit einem TEI-Header versehen, der die Metainformationen des Projektes sowie der jeweiligen Veröffentlichungen enthält. Damit im gesamten Projekt an den Texten gearbeitet werden kann, wurden schließlich alle Dokumente in die virtuelle Forschungsumgebung *TextGrid* (vgl. TextGrid Konsortium 2006–2014) importiert.

TextGrid stellt mit dem *TextGridLab* sowohl die Arbeitsumgebung zur Auszeichnung der Texte als auch mit dem *TextGridRepository* die Infrastruktur für die Langzeitarchivierung. Das Webportal, das auf SADE – der *Scalable Architecture for Digital Editions* – basiert, greift auf die Dokumente im *TextGridLab* zurück. *TextGrid* ist ursprünglich im Rahmen von DARIAH-DE, dem Verbund *Digital Research Infrastructure for the Arts and Humanities*, entwickelt worden (DARIAH-DE). Zwar ist das Projekt DARIAH-DE inzwischen abgeschlossen, aber *TextGrid* wird weiterhin im Rahmen von CLARIAH-DE und der Nationalen Forschungsdateninfrastruktur weiterentwickelt. Die SUB Göttingen bietet die technische und informationswissenschaftliche Unterstützung zur Erstellung von digitalen Editionen im Rahmen von *TextGrid*. Sie war und ist Partnerin in verschiedenen Editionsprojekten und bietet inzwischen einen Service für digitale Editionen (vgl. Niedersächsische Staats- und Universitätsbibliothek Göttingen o. D.) an.

Die Auszeichnung der XML-Dokumente erfolgt mit dem Editor *Oxygen*, der als kostenpflichtiges Tool in *TextGrid* zur Verfügung steht. Dieser XML-Editor hat bereits die allgemeinen TEI-Schemata integriert, womit die eingegebenen Daten validiert und serialisiert werden können. Hierbei handelt es sich inzwischen um ein Standardwerkzeug in den Digital Humanities (vgl. Vogeler & Sahle 2017, 144). Neben *Oxygen* wird im Projekt außerdem die freie Software *Zotero* zur Literaturverwaltung genutzt – sowohl zur Erfassung von Literatur, auf die Mollenhauer in seinen Texten verwiesen hat, als auch der Literatur, die projektintern

zur Kommentierung der Texte verwendet wird. Gegen Projektende wird diese Zotero-Bibliothek öffentlich zugänglich gemacht. Auf dem Webportal kann man sich nach Bedarf die in Zotero erfassten Daten einblenden lassen.

2 Entscheidungen – Was wird ausgezeichnet?

Liegt ein Text als XML-Dokument in TextGrid vor, sind, wie beschrieben, erst drei Strukturen rudimentär ausgezeichnet: der TEI-Header, Absätze (<p> – *paragraph*) und Seitenumbrüche (<pb/> – *page beginning*). Der TEI-Header enthält bereits Metadaten (u. a. den Titel des Textes), muss aber zum Teil spezifiziert werden, zum Beispiel mit Informationen zu den einzelnen Fassungen. Als Fassungen werden die Varianten eines Textes bezeichnet. Dabei werden mehrere textlich-identische Ausgaben zu einer Fassung zusammengefasst: Fassung #a wäre beispielsweise die Veröffentlichung eines Textes in einer Zeitschrift, Fassung #b die erneute Veröffentlichung in veränderter Form in einem Sammelband. Die unveränderten Neuauflagen des Sammelbandbeitrags wären keine neue Fassung, sondern würden unter Fassung #b subsumiert. Im TEI-Header wird definiert, welche Merkmale diese Fassungen aufweisen, also zum Beispiel die Art der Anführungszeichen, der Hervorhebungen oder ob es Fuß- oder Endnoten gibt (vgl. Hild & Zeterberg 2020, 10ff.).

Im Text müssen zum Teil Absätze (<p>), die durch OCR-Fehler nicht ausgezeichnet sind, ergänzt werden. Zudem wird der Text in Sinnabschnitte (<div> – *divisions*) unterteilt. Dadurch werden zum Beispiel Unterkapitel mit Zwischenüberschriften (<head> – *heading*) voneinander abgegrenzt. Auf diese Weise wird die Grundstruktur des Textes erfasst.

Die digitale Edition von Klaus Mollenhauers Gesamtwerk bietet somit die textlich originalgetreue Reproduktion der Originalpublikationen und reichert diese mit weiteren Informationen an. Das geschieht zum einen über eine Kommentierung, zum anderen werden Daten aus externen Datenbanken eingebunden: Dafür werden Entitäten mit dem Element <name> ausgezeichnet, das mittels des Attributs @type für Personen, Kunstwerke und anderes spezifiziert wird. Mit dem Attribut @ref (*reference*) wird auf den dazugehörenden Eintrag in der *Gemeinsamen Normdatei* (GND) verwiesen. Auf diese Weise werden Daten über Personen und andere Entitäten eingebunden, die Mollenhauer in seinen Texten erwähnt. Im Webportal wird nur ein Bruchteil der in der GND erfassten Daten abgebildet, aber der Originaldatensatz ist verlinkt.

Hiermit wird dem Linked-Data-Ansatz des *Web of Science* (World Wide Web Consortium) gefolgt. Die Idee dahinter ist, im Internet verfügbare Daten inhaltlich miteinander in Beziehung zu setzen. Dabei profitieren auch die Datenhaltenden von dieser Verknüpfung: Während der Auszeichnung der Texte Mollenhauers fallen fehlende, unvollständige oder doppelte Einträge in der GND auf. Diese werden zu einem späteren Zeitpunkt des Projektes an die GND zurückgemeldet und so deren Datensätze korrigiert und ergänzt.

Mollenhauers Spätwerk ist gekennzeichnet durch eine intensive Auseinandersetzung mit kulturellen Zeugnissen – Kunstwerken, Musikstücken, literarischen Werken. Um diese angemessen und mit Informationen angereichert zu erfassen, wird für die Kunstwerke der XML-Dialekt *LIDO* genutzt. *LIDO* steht für *Light-weight Information Describing Objects* und dient der Bereitstellung und dem Austausch von Daten über Museumsobjekte (ICOM – International Committee for Documentation). Die LIDO-Dateien enthalten Informationen zu den Kunstwerken, wie zum Beispiel in welchem Museum das Gemälde hängt oder wo eine Skulptur entstanden ist sowie, wenn möglich, Verlinkungen auf Webseiten der sie besitzenden Museen.

Die Auszeichnung von Musikstücken befindet sich derzeit in der Entwicklung, da es hierfür noch keine äquivalente Auszeichnungspraxis wie für die Kunstwerke gibt. Beide werden, sofern von Mollenhauer als Abbildung gegeben, als JPEGs ins Portal integriert. Literarische Werke werden wie die von Mollenhauer zitierte Fachliteratur in der Zotero-Datenbank festgehalten.

Ein weiteres wichtiges Auszeichnungselement ist die „bibliografic citation“ (<bibl>). Mit <bibl> werden bibliografische Angaben erfasst, wodurch zum Beispiel Zitate, die Mollenhauer nennt, mit der Quellenangabe verknüpft werden können. Diese Quellenangabe wird durch das Attribut **@corresp** (*corresponds*) mit dem dazugehörigen geprüften Eintrag in der Zotero-Datenbank verlinkt. Dadurch ergibt sich die Chance, lückenhafte Literaturangaben zu ergänzen. Um solche Ergänzungen für spätere User:innen des Webportals kenntlich zu machen, wird in der Auszeichnung differenziert: Die eindeutig zuordenbaren Literaturverweise, bei denen aber Informationen ergänzt werden müssen, werden mit einem <bibl type="KMG"> ausgezeichnet. Bei Literaturverweisen, die nicht eindeutig zuzuordnen sind, beispielsweise weil es mehrere Auflagen gibt oder kein Titel genannt ist, wird in einem Kommentar der Hinweis ergänzt, dass hier „vermutlich“ die folgende Angabe gemeint ist. Diese Literaturverweise werden mit <bibl type="commentary"> erfasst. Gerade in seiner späten Schaffensphase verwendet Mollenhauer oft Zitate ohne Herkunftsbenennung oder setzt (zum Teil einzelne) Worte in Anführungszeichen. Dieses Vorgehen erfordert ein gewisses Maß an Interpretation bei der Auszeichnung von Anführungszeichen: in Verbindung mit einem Zitat mit einem <quote> (*quotation*), ohne Zitation mit einem <q>-Element (*quoted*). Hier zeigt sich also eine weitere Herausforderung der Auszeichnung: Wann kennzeichnen Anführungszeichen ein Zitat, wann hingegen stellen sie Hervorhebungen oder Distanzierungen dar?

In Mollenhauers Texten kommen außerdem Tabellen oder Grafiken bzw. Schemata und Bilder vor. Diese werden ebenfalls ausgezeichnet mit dem Container-element <table> bzw. <figure>. Bei allen <figure>-Elementen wird jeweils auch das Originalbild als JPEG mit dem TEI-Dokument verknüpft, sodass es später im Portal an der gleichen Stelle wie im Originaltext angezeigt wird.

2.1 Textkritik und die Transparenz im Umgang mit Fehlern und Ungenauigkeiten im Original

Der Begriff „Textkritik“ bezieht sich auf Abweichungen zwischen (weitgehend) inhaltsgleichen Veröffentlichungen. Die KMG wird in Form einer *Leittext-Ausgabe* (vgl. Sahle 2017, 238) ediert. In deren Darstellung wird einerseits klar, welche Gemeinsamkeiten (Wieder-)Veröffentlichungen haben, andererseits können Varianzen verschiedener Fassungen in einem Apparat erfasst und dargestellt werden. So erkennen User:innen Änderungen verschiedener Fassungen auf einen Blick. Dies betrifft sowohl Inhalts- als auch Strukturvarianzen. Würde zum Beispiel ein Absatz in einer späteren Fassung gegenüber dem Leittext hinzugefügt oder weggelassen, können sich User:innen dies in der interaktiven textkritischen Ansicht anzeigen lassen.

Für die Klaus-Mollenhauer-Gesamtausgabe wird bezüglich der Genauigkeit der Textwiedergabe davon ausgegangen, dass das Erscheinungsbild der ursprünglichen Ausgaben (z. B. Schriftart und -größe, Zeilenumbrüche usw.) nicht von inhaltlicher bzw. von interpretatorischer Bedeutung ist (vgl. Sahle 2013, 217). Die inhaltliche Wiedergabe sollte jedoch so genau wie möglich erfolgen, auch in Bezug auf inhaltliche Änderungen zwischen der Ausgabe *erster Hand* (also dem Leittext) und allen weiteren Fassungen. Hierzu wurde ein komplexes Vorgehen entwickelt, die verschiedenen Wiederabdrucke miteinander zu vergleichen. Die so betitelte *Kontrolllesung* beinhaltet das Kontrollieren der Abweichungen der Textzeilen (inklusive aller Satzzeichen und Absätze etc.).

Das Vorgehen variiert dabei in Bezug auf die Textgattung: Bei Aufsätzen wurde immer ein Wort-für-Wort-Vergleich durchgeführt. Für Monografien gibt es ein gestaffeltes Verfahren. Verknüpft zusammengefasst: Zeigten sich nach einem zehnteiligen Wort-für-Wort-Vergleich keine Abweichungen, wurden nur Anfangs- und Endworte von Doppelseiten kontrolliert. Wurde dabei eine Abweichung festgestellt, musste die gesamte Monografie kontrolliert werden.

Die so festgestellten Abweichungen werden in der Auszeichnung in einem **<app>** (*apparatus entry*) festgehalten, der einerseits im **<lem>** (*lemma*) den Basistext – also den Inhalt des Leittextes – enthält sowie andererseits ein **<rdg>** (*reading*)-Element, in dem die Varianz festgehalten wird. Dabei können innerhalb eines **<app>**s so viele verschiedene **<rdg>**-Elemente stehen, wie es Fassungen des Textes gibt. Damit nachvollziehbar bleibt, auf welche Fassung sich das **<lem>** und das **<rdg>** beziehen, wird diese mit dem Attribut **@wit** (*witness*) festgehalten (vgl. Hild & Zeterberg 2020, 9ff.).

Im Editionsportal können User:innen in der interaktiven textkritischen Ansicht auf ein als verändert markiertes Wort klicken und bekommen dann in einer Seitenleiste die Veränderungen gegenüber dem Leittext angezeigt. Um es in einem einfachen Beispiel zu zeigen: das Wort „besonderer“ in Fassung #a wird zu „aller“ in Fassung #b.

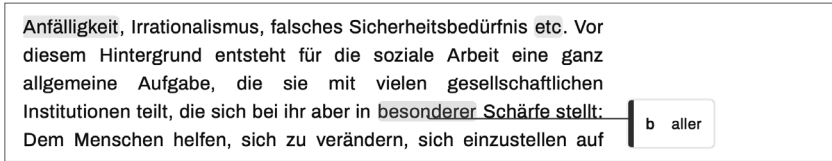


Abb. 1: Darstellung der Textkritik im Editionsportal

Auch die transparente Korrektur von Fehlern im Original (z. B. Schreibversehen wie Buchstabendreher oder fehlende Wörter) ist ein Anliegen der KMG. Um auch hier der Nachvollziehbarkeit und Transparenz treu zu bleiben, wird vermerkt, wer die Korrektur durchgeführt hat (vgl. Sahle 2017, 241).

2.2 Können vs. Brauchen – (Fast) alles ist möglich, doch was ist nötig?

Notwendige Entscheidungsprozesse

Theoretisch ist in der Auszeichnung mit TEI fast alles möglich. Es gibt jedoch auch vieles, was in bestimmten Editionen gar nicht benötigt wird. Bisher wurde nur gezeigt, was nach vielen Diskussionen und Entscheidungsprozessen für die Klaus-Mollenhauer-Gesamtausgabe als sinnvoll erachtet wurde. Im Folgenden sollen Beispiele für Entscheidungen gegen Auszeichnungsoptionen dargestellt werden. Eine dieser Diskussionen lief unter dem Titel „Raum-und-Zeit-Diskussion“:

In Sätzen wie „*Seit Karl Philipp Moritz 1785 seine Autobiographie geschrieben hat,...*“ (Mollenhauer 1983, 14) oder „*Am Pfingstsonntag des Jahres 1828 beobachten Nürnberger Bürger...*“ (ebd., 80) sind recht eindeutige Datumsangaben enthalten (1785 bzw. 25. Mai 1828). In den folgenden Beispielen hingegen, „*Am Beginn der nachantiken europäischen Geschichte der Bildung und Erziehung...*“ (ebd., 31) oder „*Pädagogik, so könnte man sagen, ist seit 300 Jahren die Welt ‚noch einmal‘,...*“ (ebd., 53; der Zusammenhang hier ist die Veröffentlichung des *Orbis Pictus* 1658 von Comenius), wird die Schwierigkeit erkennbar, ein konkretes Datum oder einen Zeitraum zu definieren. Da nach intensiven Diskussionen im Projekt kein Mehrwert der Datumsauszeichnung für mögliche Analysevorhaben erkennbar war, wurde beschlossen, von einer Auszeichnung etwaiger Datumsangaben abzusehen.

In Bezug auf Ortsangaben ist es teilweise ebenfalls nicht einfach, Konkretes festzuhalten. Als Beispiel dient hier der oben zitierte Satz: „*Am Beginn der nachantiken europäischen Geschichte der Bildung und Erziehung...*“ (ebd., 31). Hier könnte Europa aus heutiger Sicht oder eben in nachantiker Definition ausgezeichnet werden. Darüber hinaus hat auch hier der fehlende Mehrwert zur Entscheidung geführt, Ortsangaben nicht mit einer Auszeichnung zu versehen. Bei beispielsweise der Edition von *Reisetagebüchern* wäre die Sachlage eine komplett andere. Um noch zwei weitere Beispiele für Entscheidungen in der KMG zu nennen: *Zeilenumbrüche* werden nicht ausgezeichnet. Dafür spräche zwar, dass der Textfluss

des Originals auf dem Portal sichtbar würde; jedoch ergäbe sich durch eine Auszeichnung der Zeilenumbrüche, die ja in den Druckfassungen von willkürlichen Vorgaben (Satzspiegel, Schriftart und -größe) abhängen, kein inhaltlicher Mehrwert, weshalb sich der Aufwand der Auszeichnung nicht rechtfertigen lässt. Als Ausnahme stehen hier lediglich solche Textstellen, an denen die Zeilenumbrüche von inhaltlicher Bedeutung sind (z. B. Gedichte oder Tabellen). Und: In der KMG werden keine *verlagspezifischen Layouts* ausgezeichnet, also beispielsweise die Hervorhebung von gesamten Überschriften. Eine Ausnahme hier wäre jedoch die Markierung von Abweichungen, wenn beispielsweise eine Zwischenüberschrift kursiv gesetzt, in allen anderen Fällen dieser Hierarchieebene dieses Textes aber der normale Satz gesetzt wurde.

Entscheidungsprozesse in Bezug auf die Auszeichnung stehen immer im Zusammenhang mit der Frage: Was will die Edition bezwecken? Die Klaus-Mollenhauer-Gesamtausgabe will als textkritische und kommentierte Edition die Texte Mollenhauers in ihrem originalen Inhalt darstellen, Varianten genau aufzeigen und sie mit zusätzlichen Informationen (zum Beispiel zu Personen und bibliografischen Referenzen) anreichern. Hierzu mussten, wie zuletzt gezeigt, nicht nur *handwerkliche* (also welche Auszeichnungen sind möglich), sondern auch viele *inhaltliche* Entscheidungen (welche Auszeichnungen sind davon sinnvoll) getroffen werden. Eine große Herausforderung ist hierbei auch die Tatsache, dass die KMG als Verbundprojekt an vier Standorten arbeitet. Durch die große Anzahl der Mitarbeitenden und die Aufteilung der Texte auf die verschiedenen Arbeitsgruppen, muss ein hohes Maß an Abstimmung dafür sorgen, dass eine einheitliche Auszeichnung erfolgt. Alle gefassten Beschlüsse werden transparent in den projekteigenen Editi-onsrichtlinien auf dem zukünftigen Webportal einsehbar sein.

3 Computergestützte Analyse

Die Anreicherung der Texte mit zusätzlichen Informationen und die Verknüpfung mit externen Datenbanken dient in erster Linie der interaktiven Darstellung im Webportal. Darüber hinaus ermöglichen diese Daten weitergehende Analysen der Texte. Um zu verdeutlichen, welche analytischen Chancen eine digitale Edition bietet, wird am Beispiel der *Vergessenen Zusammenhänge* (vgl. Mollenhauer 1983, im Folgenden *VZ*) gezeigt, wie man die in den TEI-Dateien enthaltenen Informationen gezielt abfragen und auswerten kann. Ein Fokus liegt dabei auf der Darstellung der Methodik. Dies ist die Bedingung für eine Bewertung der Aussagekraft der Ergebnisse (vgl. Wiedemann & Lemke 2016, 405f.).

Das Werk *VZ* gehört zu den bekanntesten Veröffentlichungen Mollenhauers. Zu einer internationalen Rezeption haben die Übersetzungen ins Niederländische (vgl. Mollenhauer 1986), Japanische (vgl. Mollenhauer 1987), Norwegische (vgl. Mollenhauer 1996) und Englische (vgl. Mollenhauer 2014) beigetragen. Inhalt-

lich gelten die *VZ* als das Werk, das im Gesamtwerk Mollenhauers den Übergang von der emanzipatorischen Pädagogik der 1960er und 1970er Jahre zu einer kulturwissenschaftlich ausgerichteten Erziehungswissenschaft in den 1980er und 1990er Jahren markiert (vgl. Winkler 2002, 66; Groß 2010, 13f.; Niemeyer & Rautenberg 2011, 331). Zwar gibt es über die Härte des Bruchs bzw. Übergangs im Gesamtwerk unterschiedliche Einschätzungen¹, aber im Fachdiskurs herrscht Konsens, dass Mollenhauer sich in den *VZ* für ihn größtenteils neuen Forschungsgegenständen zuwendet. Hat Mollenhauer zwar schon zuvor literarische und autobiographische Fragmente in seinen Texten herangezogen (vgl. z. B. Mollenhauer u. a. 1975, 176ff.), stellt die Breite der Auseinandersetzung und die Ausführlichkeit der Analysen in Kombination mit der Betrachtung bildlicher Erzeugnisse – Holzschnitte, Selbstportraits etc. – eine Neuerung in seinem Werk dar. Anhand solcher Quellen entwirft er eine „grobe Skizze für das, was ‚Allgemeine Pädagogik‘ heute sein könnte“ (Mollenhauer 1983, 20).

Mit der im Folgenden vorgestellten Methode wird überprüft, ob sich Mollenhauer tatsächlich vorwiegend kulturellen Erzeugnissen zugewandt hat – insbesondere mit Blick auf die verwendete Literatur und genannte Personen. Dahinter steht die Annahme, dass die Nennung von Personennamen und die Literaturverweise Rückschlüsse über den Inhalt eines wissenschaftlichen Textes zulässt. Wenn die Analyse ergibt, dass sich Mollenhauer in signifikanter Häufigkeit auf Künstler:innen, Musiker:innen und Schriftsteller:innen sowie die entsprechende Literatur bezogen hat, unterstützt dieses Ergebnis die These, dass sich Mollenhauer in den *VZ* in besonderem Maße kulturellen Erzeugnissen zugewandt hat. In diesem Zuge kann zudem geprüft werden, ob es diesbezüglich Unterschiede zwischen der ersten und zweiten Fassung des Buchs gibt.

3.1 Methodik

Wie im zweiten Abschnitt beschrieben, sind die für die Zählung der Personennamen und Literaturverweise benötigten Informationen in den XML-Dateien explizit enthalten: Personennamen werden mittels `<name type="person">` mit den entsprechenden Daten in der GND, die Literaturverweise mittels `<bibl>` mit Einträgen in der Zotero-Datenbank verknüpft. Insofern muss lediglich ausgezählt werden, wie oft welche Referenzen in Mollenhauers Text enthalten sind. Insbesondere mit Blick auf den Gesamtkorpus kann dies maschinell sehr viel schneller und zuverlässiger geschehen als durch händisches Auszählen.

¹ Schon im zeitgenössischen Diskurs wurde ihm ein Bruch mit dem emanzipatorischen Anliegen der Kritischen Erziehungswissenschaft vorgeworfen (vgl. Wünsche 1984; Mollenhauer & Schulze 1987, 69). Mollenhauer selbst hielt dies für ein Missverstehen seines Anliegens (vgl. Häsing u. a. 1984, 46) und markierte den Übergang zur Beschäftigung mit Kultur als einen notwendigen Schritt zurück, eine Rückkehr zum erziehungswissenschaftlichen Kern, um schließlich „dahin zu kommen, einen substanziellen Begriff von ‚Emanzipation‘ zu gewinnen“ (Mollenhauer & Schulze 1987, 64).

Da hier bereits bekannte Informationen nutzbar gemacht werden, lässt sich dieses Verfahren in den Bereich des Information Retrieval beziehungsweise Data Mining einordnen (Heyer u. a. 2006, 4; Klinkke 2017, 268; Schöch 2017, 279). Üblicherweise wird dieser Bereich als Gegensatz zu Methoden wie Frequenz- und Kookkurrenzanalysen, Stilometrie, Named Entity Recognition und Topic Modeling dargestellt, die dem Distant Reading beziehungsweise der „Makroanalyse“ (DARIAH-DE 2015, 9) oder dem „Text Mining“ (Heyer u. a. 2006, 3) zugeordnet werden. Diese scheinbar eindeutige Unterscheidung muss aber insofern relativiert werden, als es sich durchaus um eine „quantitative Textanalyse“ (DARIAH-DE 2015, 9) handelt, was den Kriterien des Distant Reading entspricht (vgl. Thaller 2017, 18), da die Forscher:innen, die diese Analyse durchführen, nicht diejenigen sein müssen, die den Text gelesen und ausgezeichnet haben. Im Gegensatz zu den genannten Methoden des Distant Reading hat Information Retrieval den Vorteil, dass es sich gewinnbringend auf einzelne Texte anwenden lässt, weshalb es sich zum derzeitigen Stand des KMG-Projekts anbietet, solch ein Verfahren vorzustellen.

Bei der Informationsabfrage sind drei Punkte zu beachten: *Erstens* sollen nur die Referenzen gezählt werden, die einer Nennung von Personennamen oder einem Literaturverweis in der Originalpublikation entsprechen. Referenzen im TEI-Header, in den editorischen Kommentaren sowie in Korrekturen des Editonsteams würden das Ergebnis verfälschen, da sie erst im Zuge der Erarbeitung der KMG hinzugefügt wurden. Dementsprechend müssen sie bei der Zählung herausgefiltert werden. *Zweitens* sollen die Referenzen für jede Fassung getrennt gezählt, also die textkritische Auszeichnung beachtet werden. *Drittens* soll im Fall der Literaturverweise zwischen dem eigentlichen Text und dem Literaturverzeichnis differenziert werden. Üblicherweise listet ein Literaturverzeichnis alle zitierte Literatur auf. Einerseits kann dies die Ergebnisliste unnötigerweise aufblähen, andererseits besteht die Gefahr einer Verzerrung bei einem unvollständigen Literaturverzeichnis.

Als erstes müssen die relevanten Informationen aus den XML-Dateien extrahiert werden. Hierfür wurde die XML-Datei der VZ vom TextGrid-Laboratory in eine lokale eXist-Datenbank kopiert. Diese Open Source Datenbank eignet sich für diesen Anwendungsfall besonders gut, da sie speziell für XML-Dateien entwickelt wurde und eine native Unterstützung der XML-Abfragesprache XQuery bietet. Mit XQuery-Funktionen wird die XML-Datei nach `<name>`- und `<bibl>`-Elementen durchsucht. Dabei überprüfen die XQuery-Funktionen, ob diese Elemente die für die Analyse relevanten Attributswerte enthalten und an der richtigen Stelle in der Hierarchie des Dokuments stehen. Sind diese Bedingungen erfüllt, werden bei Personennamen der vollständige Name aus der GND und bei Literaturverweisen die vollständige Zitation, der Titel, die Namen der Autor:innen sowie das Publikationsjahr aus der Zotero-Datenbank abgerufen. Dieser Vorgang wird je Fassung einmal durchgeführt.

Das Ergebnis wird im JSON-Format ausgegeben. Mit Python-Skripten werden diese Daten über eine REST-Schnittstelle abgefragt und für eine menschliche Interpretation aufbereitet. Das Resultat sind Excel-Dateien: Für die Personennamen werden Tabellenblätter mit einer Auflistung aller Namensnennungen mit jeweils dem vollständigen Namen und dem GND-Identifer erstellt. Hinzu kommt eine Tabelle, in der für jeden Personennamen die Häufigkeit der Nennungen pro Fassung angegeben ist (siehe Abb. 3).

Für die Literaturverweise wird eine ähnlich aufgebaute Excel-Datei erstellt (siehe Abb. 2). Auch sie enthält Tabellenblätter, in denen alle Literaturverweise aufgelistet sind. Neben der vollständigen Zitation ist jeweils die Zotero-ID, Autor:in(nen), der Titel und das Publikationsjahr aufgelistet. Für jede Fassung gibt es zwei Tabellenblätter mit diesen Listen, da hier zwischen dem Text und dem Literaturverzeichnis unterschieden wird. Neben weiteren Tabellenblättern gibt es schließlich auch in dieser Datei eine nach Fassung differenzierte Auflistung der Häufigkeiten.

3.2 Ergebnisse und Interpretation

Die Zählung der Literaturverweise in den VZ erbringt folgendes Resultat: Insgesamt gibt es hier Verweise auf 102 unterschiedliche Literaturtitel. Zwischen den beiden Fassungen der VZ besteht in diesem Fall kein Unterschied.

#	Zitation	Fassung a	Fassung b
1	Comenius, Johann Amos. 1910. <i>Orbis sensualium pictus</i> . Hrsg	14	14
2	Oelkers, Jürgen, und Thomas Lehmann. 1983. <i>Antipädagogik</i> .	11	11
3	Benner, Dietrich. 1978. <i>Hauptströmungen der Erziehungswiss</i>	10	10
4	Ginzburg, Carlo. 1979. <i>Der Käse und die Würmer. Die Welt ei</i>	6	6
5	Kafka, Franz. 1960. <i>Brief an den Vater</i> . München: Piper.	6	6
6	Hauser, Kaspar. 1979. <i>Ich möchte ein solcher werden, wie... M</i>	6	7
7	Montaigne, Michel Eyquem de. 1953. <i>Essais</i> . Michel de Mont	5	5
8	Büffelkind Langspeer. 1958. <i>Häuptling Büffelkind Langspeer e</i>	5	5
9	Mannoni, Maud. 1978. <i>Ein Ort zum Leben. Die Kinder von Bo</i>	5	5
10	Pleißner, Helmuth. 1965. <i>Die Stufen des Organischen und der</i>	5	5
11	Müller, Rudolf Wolfgang. 1981. <i>Geld und Geist. Zur Entstehu</i>	5	5
12	Augustinus, Aurelius. 1955. <i>Confessiones. Bekenntnisse. Latei</i>	4	4
13	Moritz, Karl Philipp. 1960. <i>Anton Reiser. Ein psychologischer</i>	4	4
14	Pestalozzi, Johann Heinrich. 1932. <i>Brief an einen Freund über</i>	4	4
15	Alberti, Leon Battista. 1962. <i>Über das Hauswesen. Della fami</i>	4	4
16	Giedion, Sigfried. 1982. <i>Die Herrschaft der Mechanisierung. E</i>	4	4

Abb. 2: Die häufigsten Literaturverweise in den Vergessenen Zusammenhängen

Wie man in Abbildung 2 sieht, verweist Mollenhauer am häufigsten auf den *Orbis sensualium pictus* von Comenius, die *Antipädagogik* von Oelkers und Lehmann sowie die *Hauptströmungen der Erziehungswissenschaft* von Benner. Mehr als dreimal verweist er außerdem auf Schriften von Ginzburg, Kafka, Kaspar Hauser (hrsg. von

Hörisch), Montaigne, Büffelkind Langspeer, Mannoni, Plessner, Müller, Augustinus, Moritz, Pestalozzi, Alberti und Giedion. Auf fünf Titel verweist er dreimal, auf 18 Titel zweimal und auf 64 Titel wird von Mollenhauer nur einmal verwiesen. Die von ihm genutzte Literatur lässt sich in drei Kategorien² einteilen:

1. Kulturelle Erzeugnisse und kulturwissenschaftliche Literatur: Dazu gehören 48 Titel, zum Beispiel von Comenius, Brecht, Rilke, Augustinus und Homer sowie Asemissen, Eco und Barthes, auf die Mollenhauer insgesamt 112-mal verweist.
2. Erziehungswissenschaftliche und pädagogische Literatur: Hierzu lassen sich 35 Titel mit 64 Verweisen zählen. Dabei entfällt fast ein Drittel aller Verweise auf die Bücher von Oelkers und Lehmann sowie Benner. Ansonsten gehören u. a. Werke von Blankertz, Schleiermacher und Lessing zu dieser Kategorie.
3. Die restlichen 19 Titel (35 Verweise) lassen sich am ehesten als philosophische und soziologische Literatur klassifizieren. Hierzu gehören unter anderem Bücher von Plessner, Foucault, Adorno und Wittgenstein.

Die Übersicht über die Namensnennung bietet ein ähnliches Bild: Mollenhauer erwähnt an 527 Textstellen insgesamt 142 verschiedene Personen, die einen GND-Eintrag haben. Hinzu kommen Personen ohne GND-Eintrag wie zum Beispiel Mollenhauers Sekretärin Hannelore Heuer, fiktive Personen aus literarischen Werken und Kinder aus dem Beobachtungsbericht von Jürg Jegge, die in die Kategorie „keine Referenz“ fallen (siehe Abb. 3).

#	Name	Fassung a	Fassung b
1	Keine Referenz	51	51
2	Augustinus, Aurelius, Heiliger	37	37
3	Comenius, Johann Amos	35	35
4	Hauser, Kaspar	20	20
5	Velázquez, Diego Rodríguez de Silva y	19	19
6	Pestalozzi, Johann Heinrich	17	17
7	Flaubert, Gustave	16	16
8	Sartre, Jean-Paul	15	15
9	Gogh, Vincent van	13	13
10	Dürer, Albrecht	13	13
11	Rousseau, Jean-Jacques	13	13
12	Kleist, Heinrich von	12	12
13	Fichte, Johann Gottlieb	11	11
14	Kafka, Franz	10	10
15	Piaget, Jean	10	10

Abb. 3: Die häufigsten Namensnennungen in den Vergessenen Zusammenhängen

2 Bei der hier, aufgrund des Umfangs sowie der Menge des bisher analysierbaren, d. h. vollständig ausgezeichneten Materials, sehr begrenzten Analyse, wurde auf eine feinere Kategorisierung verzichtet (z. B. die Unterscheidung zwischen „Quellen“ und „Literatur“). Dennoch sollten die grundsätzlichen Möglichkeiten des Information Retrieval im Folgenden deutlich werden.

Besonders häufig nennt Mollenhauer Augustinus, Comenius und Kaspar Hauser gefolgt von Velázquez, Pestalozzi, Flaubert, Sartre, van Gogh, Dürer, Rousseau, Kleist, Fichte, Kafka und Piaget. Von den restlichen 128 Personen wurden 76 nur einmal erwähnt.

Auch die Personen lassen sich in drei Kategorien einteilen:

1. 53 Maler, Schriftsteller und andere Künstler. Insgesamt gibt es 186 Nennungen dieser Personen. Darunter besonders häufig Velázquez, Flaubert, van Gogh, Dürer, Kleist und Kafka.
2. 51 Pädagogen, Erziehungswissenschaftler, Philosophen, Bildungsreformer, Soziologen, Kunsthistoriker und Psychoanalytiker:innen die Mollenhauer 180-mal erwähnt.
3. 38 literarische, biblische, mythologische oder historische Personen wie Augustinus, Comenius oder Kaspar Hauser, die Mollenhauer 161-mal nennt.

Die Personen der ersten und dritten Kategorie lassen sich der kulturellen Sphäre zuordnen. Auffällig ist die Dominanz dieser Sphäre in beiden Abfragen: Mollenhauer verweist häufiger auf kulturelle Erzeugnisse und kulturwissenschaftliche Titel, als auf erziehungswissenschaftliche Literatur. Er nennt auch häufiger Künstler als Erziehungswissenschaftler und Pädagogen. Insofern bestätigen diese Ergebnisse die Annahme, dass sich Mollenhauer in den VZ vor allem kulturellen Erzeugnissen zugewandt hat.

3.3 Bewertung der Methode

Diese Ergebnisse zeigen dreierlei: In Bezug auf die Deutung der VZ bestätigen die Ergebnisse des Information Retrieval die aus der Literatur bekannte These und unterfüttern sie mit quantitativen Daten. Dass sie sinnvolle Ergebnisse produziert, zeigt wiederum, dass die vorgestellte Methode funktioniert. Schließlich hat sich auch die Vorannahme bestätigt, dass sich inhaltliche Schwerpunkte in der Nennung von Personennamen und Literaturverweisen ausdrücken.

Auch der Vergleich zum Leseindruck stützt die Ergebnisse, da sich in den Namenshäufigkeiten die Kapitelstruktur des Buchs widerspiegelt: In der Unterscheidung von „Präsentation“, „Repräsentation“, „Bildsamkeit“ und „Selbsttätigkeit“, sowie den von ihm ausgeführten „Schwierigkeiten mit Identität“, entfaltet Mollenhauer anhand kultureller Zeugnisse – autobiographischen Dokumenten, Briefen und Bildern – seine Theorie der Allgemeinen Pädagogik. Die für „Präsentation“ unter anderem ausgewählten Auszüge sind aus Augustinus' *Bekenntnissen* (397–401 n. Chr.), die Mollenhauer als „Teil des ersten großen Dokuments der Reflexion eines Menschen auf seine Bildung“ (Mollenhauer 1983, 23) hervorhebt – das erste große Beispiel in seiner Argumentation führt also dazu, dass Augustinus insgesamt am häufigsten genannt wird. Am zweithäufigsten wird Johann Amos Comenius genannt, dessen *Orbis sensualium pictus* (1658) Mollenhauer zur Entfaltung seines Repräsentationsbegriffes dient. An dritter Stelle der Häufigkei-

ten steht Kaspar Hauser, dessen Lebensgeschichte folgerichtig zur Explikation des Konzeptes „Bildsamkeit“ herangezogen wird. In der Auszählung folgen schließlich Velázquez, Pestalozzi, Flaubert, mit etwas Abstand Kafka und Büffelkind Langspeer, die ebenfalls in den ersten Kapiteln des Buches als Beispiele dienen, jedoch zur Unterstützung der zuvor genannten häufigeren Namen. Eine Ausnahme ist Kafka. Dessen *Brief an den Vater* (1919, veröffentlicht 1952) steht in den VZ ganz zu Beginn und dient Mollenhauer zur Darlegung der Problemstellung, wie Menschen ihre erlebte Erziehung retrospektiv zu entschlüsseln versuchen.

Andererseits zeigt eine genaue Lektüre die Grenzen des Information Retrieval: Neun der elf Verweise auf die *Antipädagogik* betreffen dort genutzte autobiografische Passagen, stellen also Sekundärzitationen von Autobiographien dar, die Mollenhauer im Unterkapitel „Gehen“ des Kapitels „Selbsttätigkeit“ nutzt. Sechs von zehn Verweisen auf Benner wiederum sind Sekundärzitate von Fichteanern. Dies verweist auf Ungenauigkeiten des Information Retrieval: Die Ergebnisse hängen davon ab, wie die Informationen erfasst wurden. In diesem Beispiel müssen Forscher:innen wissen, dass in der TEI-Auszeichnung nicht erfasst wird, dass es sich um Sekundärzitate handelt. Daher ist neben der genauen Kenntnis der Datenstruktur die Auseinandersetzung mit dem Metadatenschema und den Editionsrichtlinien unabdingbar und ein wichtiger Teil des Forschungsprozesses. Die Aussagekraft der Ergebnisse lässt sich im Sinne des „blended reading“ (Wiedemann & Lemke 2016, 410) durch eine manuelle Überprüfung und hermeneutische Analyse einzelner Textstellen erhöhen und kontrollieren.

3.4 Ausblick

Trotz dieser Limitierungen bietet dieses Vorgehen einen umfassenden Überblick über die Nennung von Personennamen und die Literaturverweise. Daraus lassen sich inhaltliche Schwerpunkte in Mollenhauers Texten ableiten, auch im Längsschnitt durch das Gesamtwerk. Doch sollte man hier den interpretativen Aufwand nicht unterschätzen.

In der KMG soll die vorgestellte Methode zum einen für die Kommentierung der Werke Klaus Mollenhauers eingesetzt werden. Zum anderen kann das Gesamtwerk, sobald es in Gänze ausgezeichnet und mit Daten angereichert ist, analysiert werden. Themenverläufe, die sich durch das Gesamtwerk ziehen, können dadurch sichtbar werden. Ob dies im deutlich zeitaufwändigeren Close Reading in dieser Form möglich wäre, ist fraglich. Wir gehen davon aus, dass wir durch solche computergestützten Analysen bisherige Annahmen über das Werk „mit Zahlen“ stützen oder aber auch modifizieren können. Dazu gehört auch die These, dass die VZ einen Wendepunkt im Gesamtwerk Mollenhauers darstellen. Dies ließe sich über einen Vergleich der Mollenhauer'schen Texte vor und nach der 1983 erst erschienenen Monografie überprüfen.

Für die Werkforschung, die im Rahmen der KMG betrieben wird, werden noch weitere computergestützte Analyseverfahren eingesetzt. Dazu gehört die Auswer-

tung der zu den Kunstwerken erfassten Daten sowie Methoden des Distant Reading, insbesondere Topic Modeling und Kookkurrenzanalysen. Darüber hinaus wird es die Möglichkeit geben, sich die edierten Texte in verschiedenen Formaten herunterzuladen. Somit können interessierte Forscher:innen außerhalb des KMG-Teams computergestützte Analysen durchführen.

4 Zusammenfassung

In diesem Beitrag wurde am Beispiel der KMG gezeigt, wie eine digitale Edition erstellt wird. Dabei wurden zunächst die notwendigen Vorarbeiten beschrieben sowie die genutzte Software genannt. Im zweiten Abschnitt wurde anhand einiger grundsätzlicher Entscheidungen erläutert, welche interpretative Arbeit mit der Erarbeitung einer digitalen Edition verbunden ist. Der letzte Abschnitt zeigt den Mehrwert einer digitalen Edition gegenüber einer gedruckten Edition, der über die einfachere Verfügbarkeit und die interaktive Darstellung hinausgeht: Am Beispiel der *Vergessenen Zusammenhänge* wurde gezeigt, wie Informationen aus den TEI-Dateien mittels Information Retrieval analytisch genutzt werden können. Insofern lässt sich zusammenfassen: Digitale Editionen haben ein großes Potential für die historische Bildungsforschung, da sie Texte nicht nur einfach neu verfügbar machen, sondern dies in einem Format tun, das neue Möglichkeiten der Analyse bietet; sie sind jedoch in ihrem Erstellungsaufwand nicht zu unterschätzen.

Literatur

- Brumlik, M. (1998): Klaus Mollenhauer. Die Sozialpädagogik in der Einheit seines Werks. In: Zeitschrift für Erziehungswissenschaft 1 (3), 431–440.
- DARIAH-DE (Hrsg.) (2015): Handbuch Digital Humanities. Anwendungen, Forschungsdaten und Projekte: DARIAH-DE.
- DARIAH-DE (Hrsg.): DARIAH-DE. Digital Research Infrastructure for the Arts and Humanities. URL: <https://de.dariah.eu/> (Zugriff: 17.05.2021).
- Dietrich, C. & Müller, H.-R. (Hrsg.) (2000): Bildung und Emanzipation. Klaus Mollenhauer weiterdenken. Weinheim und München: Juventa.
- Groß, S. (2010): Zwischen Politik und Kultur. Pädagogische Studien zur Sache der Emanzipation bei Klaus Mollenhauer. Würzburg: Königshausen & Neumann.
- Häsing, H.; Mollenhauer, K. & Wünsche, K. (1984): Wie zukunftsfähig ist unsere Vergangenheit? Ein Gespräch mit Klaus Mollenhauer über sein Buch „Vergessene Zusammenhänge“. In: päd. extra (4), 46–49.
- Heyer, G.; Quasthoff, U. & Wittig, T. (2006): Text Mining: Wissensrohstoff Text. Konzepte, Algorithmen, Ergebnisse. Herdecke: W3L-Verlag.
- Hild, A. & Zeterberg, M.-F. (2020): Von der gedruckten Quelle zum bearbeiteten Digitalisat: Ausgabenanalysen, Auszeichnungen und Metadaten. In: Hild, A.; Kirchberg, A.; Dietrich, C.; Horn, K.-P. & Müller, H.-R. (Hrsg.): Die Klaus Mollenhauer Gesamtausgabe (KMG). Textkritische und kommentierte Edition. Ein Werkstattbericht. Göttinger Beiträge zur erziehungswissenschaftlichen Forschung, 39. Göttingen: Universitätsverlag, 9–15. <https://doi.org/10.17875/gup2020-1333>.
- Hild, A.; Kirchberg, A.; Dietrich, C.; Horn, K.-P. & Müller, H.-R. (Hrsg.) (2020): Die Klaus Mollenhauer Gesamtausgabe (KMG). Textkritische und kommentierte Edition. Ein Werkstattbericht. Göttingen: Universitätsverlag. <https://doi.org/10.17875/gup2020-1333>

- ICOM – International Committee for Documentation (Hrsg.): LIDO. URL: <http://cidoc.mini.icom.museum/working-groups/lido/what-is-lido/> (Zugriff: 17.05.2021).
- Klinke, H. (2017): Information Retrieval. In: F. Jannidis; H. Kohle & M. Rehbein (Hrsg.): *Digital Humanities. Eine Einführung*. Stuttgart: J. B. Metzler, 268–278.
- Mollenhauer, K. (1959): *Die Ursprünge der Sozialpädagogik in der industriellen Gesellschaft*. Weinheim: Beltz.
- Mollenhauer, K. (1964): *Einführung in die Sozialpädagogik. Probleme und Begriffe der Jugendhilfe*. Weinheim und Basel: Beltz.
- Mollenhauer, K. (1983): *Vergessene Zusammenhänge. Über Kultur und Erziehung*. München: Juventa.
- Mollenhauer, K. (1986): *Vergeten samenhang: over cultuur en opvoeding*. Meppel, Amsterdam: Boom.
- Mollenhauer, K. (1987): *忘れられた連関 – 「教える - 学ぶ」とは何か*. Tokyo: Misuzu Shobo.
- Mollenhauer, K. (1996): *Glemte sammenhenger: om kultur og oppdragelse*. Oslo: Ad Notam Gyldendal.
- Mollenhauer, K. (2014): *Forgotten Connections. On culture and upbringing*. London, New York: Routledge.
- Mollenhauer, K. & Schulze, T. (1987): Klaus Mollenhauer in Gespräch mit Theodor Schulze. In: H. B. Kaufmann (Hrsg.): *Kontinuität und Traditionsbrüche in der Pädagogik. Im Gespräch mit Wilhelm Flitner und Ulrich Herrmann, Otto Friedrich Bollnow und Klaus Giel, Klaus Mollenhauer und Theodor Schulze, Doris Knab und Hans Rauschenberger*. Münster: Comenius-Institut, 48–65.
- Mollenhauer, K.; Brumlik, M. & Wudtke, H. (1975): *Die Familienerziehung*. München: Juventa.
- Niedersächsische Staats- und Universitätsbibliothek Göttingen (Hrsg.) (o. D.): *Service Digitale Editionen*. <https://www.sub.uni-goettingen.de/digitale-bibliothek/service-digitale-editionen/> (Zugriff: 17.05.2021).
- Niemeyer, C. & Rautenberg, M. (2011): Klaus Mollenhauer (1928–1998). *Pädagogik als vergessener Zusammenhang*. In: B. Dollinger (Hrsg.): *Klassiker der Pädagogik. Die Bildung der modernen Gesellschaft* (3., durchges. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften, 331–352.
- Sahle, P. (2013): *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels. Teil 1: Das typografische Erbe*. Norderstedt: BoD.
- Sahle, P. (2017): *Digitale Edition*. In: F. Jannidis; H. Kohle & M. Rehbein (Hrsg.): *Digital Humanities. Eine Einführung*. Stuttgart: J. B. Metzler, 234–249.
- Schöch, C. (2017): *Quantitative Analyse*. In: F. Jannidis; H. Kohle & M. Rehbein (Hrsg.): *Digital Humanities. Eine Einführung*. Stuttgart: J. B. Metzler, 279–298.
- Tenorth, H.-E. (1999): *Die Zweite Chance. Oder: Über die Geltung von Kritikersprüchen „kritischer Erziehungswissenschaft“*. In: H. Sünker & H.-H. Krüger (Hrsg.): *Kritische Erziehungswissenschaft am Neubeginn?! Frankfurt am Main: Suhrkamp*, 135–161.
- TextGrid Konsortium (Hrsg.) (2006–2014): *TextGrid. Virtuelle Forschungsumgebung für die Geisteswissenschaften*. <https://textgrid.de/> (Zugriff: 17.05.2021).
- Thaller, M. (2017): *Digital Humanities als Wissenschaft*. In: F. Jannidis; H. Kohle & M. Rehbein (Hrsg.): *Digital Humanities. Eine Einführung*. Stuttgart: J. B. Metzler, 13–18.
- Vogeler, G. & Sahle, P. (2017): *XML*. In: F. Jannidis; H. Kohle & M. Rehbein (Hrsg.): *Digital Humanities. Eine Einführung*. Stuttgart: J. B. Metzler, 128–146.
- Wiedemann, G. & Lemke, M. (2016): *Auf dem Weg zu einer Best Practice? In: M. Lemke & G. Wiedemann (Hrsg.): Text Mining in den Sozialwissenschaften. Grundlagen und Anwendungen zwischen qualitativer und quantitativer Diskursanalyse*. Wiesbaden: Springer Fachmedien Wiesbaden, 397–419.
- Winkler, M. (2002): *Klaus Mollenhauer. Ein pädagogisches Porträt*. Weinheim und Basel: Beltz UTB.
- World Wide Web Consortium: *Semantic Web*. <https://www.w3.org/standards/semanticweb/> (Zugriff: 17.05.2021).
- Wünsche, K. (1984): *Kultur erzieht sich selbst. Rezension zu „Vergessene Zusammenhänge“*. In: *Die Zeit* (5), 26.