

Schriften des Instituts für Dokumentologie und Editorik — Band 11

Kodikologie und Paläographie im digitalen Zeitalter 4

Codicology and Palaeography in the Digital Age 4

herausgegeben von | edited by

Hannah Busch, Franz Fischer, Patrick Sahle

unter Mitarbeit von | in collaboration with

Bernhard Assmann, Philipp Hegel, Celia Krause

2017

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Digitale Parallelfassung der gedruckten Publikation zur Archivierung im Kölner Universitäts-Publikations-Server (KUPS). Stand 4. September 2017.

SPONSORED BY THE



Federal Ministry
of Education
and Research

Diese Publikation wurde im Rahmen des Projektes eCodicology (Förderkennzeichen 01UG1350A-C) mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) gefördert.

Publication realised within the project eCodicology (funding code 01UG1350A-C) with financial resources of the German Federal Ministry of Research and Education (BMBF).

2017

Herstellung und Verlag: Books on Demand GmbH, Norderstedt

ISBN: 978-3-7448-3877-1

Einbandgestaltung: Julia Sorouri, basierend auf Vorarbeiten von Johanna Puhl und Katharina Weber; Coverbild nach einer Vorlage von Swati Chandna.

Satz: Lua \TeX und Bernhard Assmann

Kodikologie und Paläographie zwischen Geisteswissenschaften und Informatik

Andrea Rapp, Celia Krause, Philipp Hegel

Zusammenfassung

Digitale Kodikologie und Paläographie, wie sie sich in der Tagungsreihe *Maschinen und Manuskripte* des Projekts *eCodicology* präsentierten, werden üblicherweise in interdisziplinären Projekten und multidisziplinären Forschergruppen realisiert. Ein solches Design hat Konsequenzen für die digitale Kodikologie und Paläographie. Unsere Einführung beschreibt und systematisiert, wie mit dieser Situation bei den in diesem Band präsentierten Vorhaben umgegangen wird. Dabei zeigen sich Momente, die als ›zusammengesetzte‹ oder ›ergänzende‹ Interdisziplinarität verstanden werden können.

Abstract

Digital Codicology and Paleography as they are presented at the conferences of the *Machines and Manuscripts* series organized by the *eCodicology* project are normally realized in interdisciplinary projects and multidisciplinary teams. This research design has consequences for the structure of digital codicology and palaeography. Our introduction describes and categorizes the ways in which this situation is managed in the research projects represented in this volume. Aspects of composite and supplementary interdisciplinarity can be found.

1 Hintergrund

Der vorliegende Band fußt zum Teil auf Beiträgen aus der Tagungsreihe *Maschinen und Manuskripte*, die im Rahmen des Projekts *eCodicology* in den Jahren 2014 bis 2016 in Trier, Karlsruhe und Darmstadt veranstaltet worden ist. Für die erste internationale Tagung konnten Referentinnen und Referenten aus den Bereichen Paläographie, Kodikologie, Diplomatik, Informatik und Bibliothekswesen gewonnen werden. Die zweite Konferenz war technisch orientiert und bot Expertinnen und Experten der automatischen Mustererkennung oder der Informationsvisualisierung Raum für einen interdisziplinären Austausch. Die Abschlusskonferenz mit dem allgemein gehaltenen Titel *Forschung mit Schriftquellen im digitalen Zeitalter* hatte ihren Fokus auf

computergestützten Verfahren für die Analyse von handgeschriebenem Dokumenten und gedruckten Büchern. Analoge und digitale Methoden wurden explizit gegenübergestellt, um einen offenen Dialog zwischen traditionell und digital Forschenden zu ermöglichen. Buchkundliche wie auch technische Aspekte wurden gleichermaßen beleuchtet und ihre Relevanz innerhalb der Digital Humanities aufgezeigt.

Das Forschungsprojekt *eCodicology* hatte die Layoutanalyse von Handschriften des europäischen Mittelalters zum Gegenstand. Ziel war es, automatisch Maße auf Handschriftenseiten zu erkennen. Ausgangspunkt waren etwa 170.000 um Metadaten ergänzte Bilddigitalisate aus dem Skriptorium der Benediktinerabtei St. Matthias in Trier. Grundlegende Elemente der Gestaltung wie Seitenfläche oder Text- und Bildanteile wurden in ihrer Ausdehnung, Anzahl und Position auf jeder Seite durch den Einsatz von Bildverarbeitung und Algorithmen zur Merkmalsextraktion ermittelt, die dann statistisch ausgewertet wurden, um Muster und Veränderungen innerhalb des Bestandes von St. Matthias aufzeigen zu können.

2 Zwischen den Disziplinen

Digitale Kodikologie und Paläographie stehen, wie schon ihre Namen verdeutlichen, nicht mehr ganz auf dem festen Boden geisteswissenschaftlicher Tradition. Sie schweben zumindest scheinbar zwischen den Medien und den Disziplinen. So verstanden passen sie gut zu jenem nicht mehr ganz neuen Schlagwort der Interdisziplinarität, das seinerseits interdisziplinär behandelt wird. Wer aber schwebt, hat mit der Schwere zu kämpfen. Disziplinen sind selbst soziokulturelle Gebilde, die sich im Laufe der Zeiten herausgebildet, gefestigt und institutionalisiert haben. Auch wenn man anerkennt, dass bestimmte Gegenstände, Fragestellungen und selbst Methoden nicht nur von einer einzelnen Disziplin behandelt, beantwortet oder angewendet werden, heißt dies nicht, dass ein Austausch von Ergebnissen und Verhandlungen über Verfahren problemlos ist. Es gibt gute Gründe, warum sich Disziplinen ausdifferenziert haben und es für den einzelnen Wissenschaftler oder die einzelne Wissenschaftlerin schwierig ist, einen hinreichend hohen Grad an Spezialisierung in mehreren Disziplinen zu erreichen. Diese Schwierigkeit variiert mit den Disziplinen, aber zwischen traditionell ausgebildeten Geisteswissenschaftlern und Geisteswissenschaftlerinnen einerseits und Informatikern und Informatikerinnen andererseits ist die Differenz zumindest nicht zu unterschätzen.

Auch an den Gegenständen des Buches und der Schrift haben verschiedene Disziplinen Interesse. An der Tagungsreihe *Maschinen und Manuskripte* waren Philologien, historische Wissenschaften, Musikwissenschaft und Informatik beteiligt, in einigen Beiträgen wird ferner deutlich, dass sich auch Physik und Chemie mit diesen Gegenständen beschäftigen. Interdisziplinäre Tagungen, Projekte und Sonderforschungsbe-

reiche bringen seit einiger Zeit Vertreter und Vertreterinnen dieser unterschiedlichen Disziplinen zusammen, um diesen Gegenständen näherzukommen.¹ Die Tagungen des Projektes *eCodicology* und der daraus entstandene Band wollen einige dieser Perspektiven auf das Buch und die Schrift vorstellen. Dabei sollen die disziplinären Grenzen nicht verwischt, aber jene interdisziplinären Schnittmengen betont werden, die sich zwischen ihnen ergeben.

Das Projekt *eCodicology* war selbst insofern interdisziplinär angelegt, als Vertreter und Vertreterinnen aus Philologie, Geschichte, Archäologie und Informatik zusammengearbeitet haben, um Methoden der Bild- und Layoutanalyse mit kodikologischen Interessen, statistischen Auswertungen und Visualisierungstechniken zu kombinieren. Aus disziplinärer und interdisziplinärer Sicht sind zwei Punkte von allgemeinerem Interesse.

Bekannt und sprichwörtlich ist zum einen, dass Geisteswissenschaften und Informatik nicht immer die gleiche Sprache sprechen. Um sich eindeutiger verständigen zu können, wurde ein hierarchisch angelegtes Glossar kodikologischer Fachbegriffe in SKOS angelegt. Es steht nun auch anderen Forschenden zur Verfügung, um zum Beispiel automatisch vermessene Bildbereiche genauer zu beschreiben.

Das Projekt ist zum anderen in einem engeren Sinn interdisziplinär, wenn damit gemeint ist, dass Arbeitsschritte, die in einer Disziplin vorgenommen werden, Resultate aus anderen Arbeitsschritten voraussetzen, die in der anderen Disziplin vorgenommen werden. Die digitale Bild- und Layoutanalyse, die in das Aufgabengebiet der Informatiker fällt, schuf die Voraussetzungen für die statistische Auswertung der quantitativ arbeitenden Kodikologen. Die Voraussetzung war nicht zwingend, aber das Auffinden und Zählen von einzelnen graphischen Elementen durch den Kodikologen wäre deutlich weniger effizient gewesen. In diesem Fall handelte es sich also nicht um eine mögliche, noch enger gefasste Form von Interdisziplinarität, bei der Arbeitsschritte, die in einer Disziplin vorgenommen werden, Resultate aus anderen Disziplinen sogar notwendig voraussetzen. Aber die Verzahnung von Arbeitsschritten und Forschungsergebnissen spricht wohl dafür, hier nicht mehr von einer ›nur‹ multidisziplinären Sicht auf einen Gegenstand zu sprechen. Es lässt sich jedoch nicht von Transdisziplinarität sprechen, wenn damit gemeint ist, dass die einzelnen Wissenschaftler Arbeitsschritte übernehmen, die traditionellerweise von Vertretern einer anderen Disziplin übernommen werden.² Die Zuständigkeiten für die einzelnen Arbeitsschritte

¹ Zu denken ist an die Sonderforschungsbereiche *Materiale Textkulturen* in Heidelberg und *Manuskriptkulturen* in Hamburg. Zu denken ist auch an Projekte wie *Digital Resource and Database for Palaeography, Manuscript Studies and Diplomatic (DigiPal)*, die nicht von den Geisteswissenschaften allein umgesetzt werden.

² Jürgen Mittelstraß sieht in der Transdisziplinarität die »Interdisziplinarität im recht verstandenen Sinne«. Vgl. Mittelstraß 1987, 156: »Sie läßt die disziplinären Dinge nicht einfach, wie sie sind, sondern stellt, und sei es nur in bestimmten Problemlösungszusammenhängen, die ursprüngliche *Einheit der*

blieben vielmehr klar den einzelnen Disziplinen zugeordnet.³ Die Verwendung von Resultaten aus dem Arbeitsschritt einer Disziplin in einem Arbeitsschritt der anderen verlangt aber ebenso eine Übersetzung wie die Formulierung von Anforderungen der einen Disziplin an die andere. Interdisziplinarität als Zwischenform zwischen Multi- und Transdisziplinarität kann diese Abhängigkeiten vielleicht ganz gut bezeichnen.

Die Verbindung mehrerer, disziplinär zugeordneter Arbeitsschritte und die Übersetzungsarbeit an den Schnittstellen zwischen diesen Schritten kann begrifflich präziser gefasst werden. Das Ineinander der disziplinär bestimmten Arbeitsschritte kann als Ausdruck einer ›zusammengesetzten‹ Interdisziplinarität mit kooperativen Momenten an den Schnittstellen gesehen werden. Zusammengesetzt heißt dabei, dass verschiedene Fähigkeiten zur Lösung eines Problems verwendet werden.⁴ Kooperativ heißt, dass an den Nahtstellen Teamwork erforderlich ist.⁵ Explizit wird das Ineinandergreifen von Disziplinen im Aufsatz von *Inga Behrendt, Jennifer Bain* und *Kate Helsen* anhand der Rolle des Musikwissenschaftlers bei der Arbeit mit einem digitalen Neumen-Wörterbuch erklärt. *Nanette Riffler-Pipka* nimmt in ihrem Aufsatz indirekt Bezug auf interdisziplinäres Arbeiten, indem sie darauf anspielt, dass sich der Blick der Geisteswissenschaft und der Blick der Informationstechnik auf digitale Bilder wesentlich voneinander unterscheiden.

Die interdisziplinäre Tätigkeit im Projekt *eCodicology* umfasste methodische wie theoretische Aspekte. Theoretische Interdisziplinarität findet sich bei der Erstellung eines gemeinsamen begrifflichen Rahmens in *eCodicology*, um die Ergebnisse der Bildanalyse in kodikologische Termini zu übersetzen.⁶ Darüber hinaus wurde die

Wissenschaft [...] wieder her.« In seinem Sinne wäre im Fall von *eCodicology* daher nicht von »rechter Interdisziplinarität«, sondern eher von Multidisziplinarität zu sprechen. Vgl. Mittelstraß 1998, 32: »Interdisziplinarität in Form von *Multidisziplinarität* läßt alles Fachliche oder Disziplinäre, wie es ist; man rückt nur auf Zeit, und ohne die eigenen fachlichen oder disziplinären Orientierungen irgendwie zur Disposition zu stellen, zusammen.« Der in dieser Einleitung verwendete Begriff der Interdisziplinarität soll anzeigen, dass ein Vorhaben als Ganzes mehr als nur multidisziplinär sein kann, obwohl alle grundlegenden Arbeitsschritte selbst disziplinär zugeordnet bleiben.

³ Dies lässt sich auch als eine ethische Maxime interdisziplinärer Arbeit verstehen, wie sie Ian Hacking in einer persönlichen Stellungnahme zum Thema ausgedrückt hat. Vgl. Hacking 2010, 196: »Worauf es meiner Meinung nach ankommt, ist, dass aufrichtige und gewissenhafte Denker und Engagierte gegenseitigen Respekt für ihre erworbenen Fähigkeiten und natürlichen Talente aufbringen.«

⁴ Zur dieser und der folgenden Begrifflichkeit vgl. Julie Thompson Kleins taxonomische Zusammenstellung. Vgl. Klein 2010, 18: »The label *Composite ID* names another familiar practice – applying complementary skills to address complex problems or to achieve a shared goal.« »ID« steht dabei für Interdisziplinarität. Vgl. auch in anderer Begrifflichkeit Klein 2010, 19: »In *Shared ID* [...] different aspects of a complex problem are tackled by different groups. They possess complementary skills, communicate results, and monitor overall progress.«

⁵ Vgl. Klein 2010, 19: »*Cooperative ID* requires teamwork.«

⁶ Vgl. Klein 2010, 20: »The outcomes [der theoretischen Interdisziplinarität] include conceptual frameworks for analysis of particular problems, integration of propositions across disciplines, and new synthesis based on continuities between models and analogies.« Die beiden Bereiche arbeiteten bei

Methode der digitalen Bildanalyse von der Kodikologie ›entliehen‹, um die Gestaltung mittelalterlicher Handschriften zu erfassen.⁷ In einem eigens erstellten Computerprogramm ist das geisteswissenschaftlich motivierte Training der Algorithmen zur Bilderkennung ebenso Voraussetzung für die Bildanalyse wie diese für die anschließende kodikologische Annotation und quantitative Auswertung. Wenn die Adaption von Methoden in eine dauerhafte Dependenz mündet, kann von ›ergänzender‹ oder ›supplementärer‹ methodischer Interdisziplinarität gesprochen werden.⁸ Die Dauerhaftigkeit besteht in diesem Fall vor allem in dem digitalen Werkzeug, das fortan anderen, ähnlich ausgerichteten Vorhaben zur Verfügung steht. Der Beitrag von *Rißler-Pipka* in diesem Band stellt eine solche Nutzung der im Projekt *eCodicology* entwickelten Software in einem anderen Kontext vor. Ein weiteres Beispiel für die Nachnutzung der entwickelten Software ist ein Gastprojekt des Berliner Sonderforschungsbereichs *Episteme in Bewegung*, in dem automatisch erkannte Marginalien in der handschriftlichen Überlieferung des aristotelischen *De interpretatione* nach verschiedenen Kriterien klassifiziert und statistisch untersucht werden.

Auch die übrigen Beiträge dieses Bandes bewegen sich zwischen verschiedenen Disziplinen, sind aber oft ihrem Gegenstand, ihrer Frage- und Problemstellung, manchmal auch ihrer Methode nach einer traditionellen Disziplin mehr oder weniger klar zuzuordnen. Dennoch gibt es Schnittpunkte, in denen wie bei *eCodicology* Resultate einer Disziplin zum Material einer anderen werden oder Disziplinen auf andere Weise miteinander verzahnt agieren.

3 Gegenstände

Eine Möglichkeit, die Beiträge dieses Bandes zu ordnen, besteht darin, nach ihren zentralen Gegenständen zu fragen. Trotz aller Übergänge, die sich fast notwendig ergeben, lassen sich doch kleinere Gruppen spezifischen Inhalts identifizieren.

Sammlungen: Sammlungen können sowohl geschlossene historische Gebilde sein, die tatsächlich einmal bestanden haben, als auch ›künstliche‹ Zusammenstellungen von Texten nach verschiedenen Gesichtspunkten wie der Gattung oder dem Ort ihrer Herstellung. In beiden Fällen liegt der Schwerpunkt weniger auf dem einzelnen Objekt als auf dem Zusammenhang zwischen einer Vielzahl ähnlicher Objekte. Bei digitalen Sammlungen handelt es sich hauptsächlich um retrospektiv digitalisierte Quellenbe-

eCodicology aber dennoch auf der Grundlage ihrer eigenen Annahmen und Modelle.

⁷ Vgl. Klein 2010, 19: »The typical activity [in methodologischer Interdisziplinarität] is borrowing a method or concept from another discipline in order to test a hypothesis, to answer a research question, or to help develop a theory«.

⁸ Vgl. Klein 2010, 19: »If borrowing becomes more sophisticated and an enduring dependence develops, the relationship becomes *Supplementary*«. Beide Kriterien, Gewandtheit und Dauerhaftigkeit, sind relativ. Wann genau sie erfüllt sind, bedarf der Klärung.

stände kultureller Gedächtnisinstitutionen wie Bibliotheken, Archive oder Museen. Zu digitalen Sammlungen gehören aber auch Forschungsdaten, die im digitalen Medium generiert wurden. Digitale Sammlungen sind notwendige Voraussetzungen für die computergestützte Verarbeitung und Analyse. *Hartmut Beyer, Jörn Münkner, Katrin Schmidt und Timo Steyer* erschließen frühneuzeitliche Gelehrtenbibliotheken und stellen in ihrem Beitrag Möglichkeiten einer visualisierenden Auswertung vor, *Matthew Driscoll* untersucht die Überlieferung der isländischen »Geschichten der alten Männer aus den Nordländern«, der »Vorzeitsagas«.

Handschriftenbeschreibungen: Die Beschreibung von Handschriften ist ein weiteres Themenfeld, das in den versammelten Aufsätzen abgedeckt wird. Für die Erstellung von elektronischen Handschriftenkatalogen greift man oft auf die Angaben aus den gedruckten Katalogen zurück, die über einen langen Zeitraum erarbeitet worden sind. Durch eine umfassende Digitalisierung von Beständen ist es nun möglich, computerunterstützte Verfahren auf große Handschriftenbestände anzuwenden und neue Daten zu gewinnen. Handschriftenbeschreibungen können so mit zusätzlichen Informationen zu bekannten Beschreibungskategorien, aber auch mit gänzlich neuen Datenkategorien angereichert werden. *Alberto Campagnolo, Erin Connelly und Dot Porter* stellen das digitale Werkzeug VisColl vor, mit dem Lagen erfasst und dargestellt werden können. *Hannah Busch* und *Swati Chandna* beschreiben Werkzeuge, mit denen im Projekt *eCodicology* anhand von Digitalisaten mittelalterliche Handschriften automatisch vermessen und die Ergebnisse visualisiert werden können. Die Lagenbeschreibung und die Vermessung behandeln dabei zunächst jede Handschrift für sich, auch wenn im zweiten Fall herausgestellt wird, wie diese Daten anschließend für einen ganzen Bestand ausgewertet werden können. Im ersten Beitrag wird beschrieben, wie das Werkzeug den Kodikologen unmittelbar unterstützt und Außenstehenden anschließend einen Zugriff auf das Resultat erlaubt. Im zweiten Beitrag wird beschrieben, wie die Daten vom Computer gewonnen und mit bestehenden Metadaten kombiniert und verglichen werden. Auch wenn in diesen beiden Beispielen vorrangig Handschriften behandelt werden, sind die digitalen Techniken auch zur Beschreibung von Drucken geeignet.

Zeichensysteme: Ein drittes Themengebiet sind semiotische Systeme, die neben der Schrift das Buch ausmachen. *Rißler-Pipka* wendet das Werkzeug aus *eCodicology* an, um in lateinamerikanischen und spanischen Kulturzeitschriften der Moderne unterschiedliche Gestaltungsprinzipien und -praktiken zu untersuchen. Dabei betrachtet sie insbesondere das Verhältnis von Bild und Text. *Behrendt, Bain* und *Helsen* versuchen die Aspekte der Bildverarbeitung und der Kodierung im Feld der Neuenforschung zu verbinden. Die digitale Behandlung mittelalterlicher wie moderner Handschriften und Drucke hat also nicht nur mit sprachlichem Text zu tun, sondern ebenso mit Illustrationen und Noten, die oft für die Beschreibung der Handschrift oder des Drucks relevant sind.

Schriftforschung: Definitionsgemäß legt insbesondere die Paläographie besonderes Augenmerk auf das Zeichensystem der Schrift. *Svenja Gülден, Celia Krause* und *Ursula Verhoeven* befassen sich in ihrem Beitrag mit der semantischen Modellierung und Visualisierung von Daten für eine digitale Paläographie des Hieratischen. *Torsten Schaßan* weist auf die Schwierigkeiten einer gemeinsamen Sprache für die Beschreibung von Schriften hin und skizziert Lösungsmöglichkeiten mit Blick auf Semantic Web-Technologien. *Bartosz Bogacz* und *Hubert Mara* beschreiben, wie sie mit XML-basierten Vektorgrafiken Keilschrift analysieren, *Enrique Vidal* stellt ein Programm zur Unterstützung der Transkription vor. Auch innerhalb der Analyse eines einzelnen Zeichensystems gibt es – ähnlich den beiden Projekten im Bereich der Handschriftenbeschreibung – unterschiedliche Grade, die Prozesse ganz oder teilweise einem digitalen Werkzeug zu überlassen. Der Grad der Unterstützung bei der Erkennung von Schrift und der Grad der Kontrolle werden je nach Gegenstand und Erkenntnisinteresse von verschiedenen Forschenden unterschiedlich bewertet und gehandhabt.

Datierung und Stemmologie: Digitale Verfahren werden auch für die chronologische Einordnung der Forschungsobjekte eingesetzt. *Vincent Christlein, Martin Gropp* und *Andreas Maier* unterscheiden zwischen einem inhaltsbasierten und einem bildbasierten Zugang bei der Datierung von Handschriften und wenden eine automatische Methode an, indem sie Merkmale der Schrift auf Digitalisaten extrahieren und gruppieren. *Gábor Hosszú* wendet phylogenetische Methoden an, um die zeitliche Entwicklung ganzer Schriftsysteme zu untersuchen. Steht in dem einen Fall eher die Datierung des einzelnen Objektes im Vordergrund, so in dem anderen der Versuch, Ähnlichkeiten von Graphemen zu nutzen, um ihre Verwandtschaft zu beschreiben, auch wenn diese die Grenzen des einzelnen Schriftsystems überschreitet.

Materialität: Die Untersuchung der Materialität eines Buches am Digitalisat steht vor der besonderen Schwierigkeit, dass das digitale Objekt sich gerade materiell deutlich von dem eigentlich interessierenden Original unterscheidet. *Dariya Rafiyenko* beschreibt in ihrem Beitrag einen Weg, um digital mit der Materialität eines Palimpsests umzugehen und die *scriptio inferior* ohne großen technischen Aufwand sichtbar zu machen. *Campagnolo, Connelly* und *Porter* stellen in ihrem Aufsatz fest, dass bei der Präsentation von Digitalisaten oft Hinweise auf die physischen Eigenschaften von Büchern fehlen. Ihr Werkzeug VisColl soll bei der Beschreibung des Lagenschemas Abhilfe schaffen. Insbesondere Studien zum Layout von Schriftdokumenten können auch dazu beitragen, mehr über die Materialität dieser Quellen zu erfahren, wie in den Beiträgen von *Rißler-Pipka* sowie *Busch* und *Chandna* durchscheint.

4 Überschneidungen

Quer zu den oben genannten Gegenständen stehen einzelne digitale Verfahren, die in variierenden Kontexten unterschiedlichen Zwecken dienen. Die technischen Werkzeuge selbst sind durchaus vergleichbar, aber sie werden in unterschiedlichen Disziplinen zur Anwendung gebracht, um je eigenen Erkenntnisinteressen zu dienen.

Text- und Zeichenerkennung: Sowohl im *Optical Neume Recognition Project* von Behrendt, Bain und Helson als auch in der Anwendung von Handwritten Text Recognition (HTR) und Keyword Spotting (KWS) bei Vidal oder rootSIFT bei Christlein, Gropp und Maier werden Mechanismen zur automatischen Zeichenerkennung benutzt. Der phenetische Ansatz Hosszús setzt eine solche Zeichenerkennung voraus. Auch wenn diese Techniken in den vorliegenden Fällen Gegenstand der Informatik sind, zeigt die Einbindung von Trainingseinheiten, dass hier die geisteswissenschaftliche Beurteilung der Ergebnisse fest in den computerisierten Arbeitsablauf integriert ist.

Bildanalyse: Notwendige Vorstufe für die optische Erkennung von Schriftzeichen und Neumen ist die automatische Bildanalyse. Eine Bildanalyse kann sich auf 3D-Modelle oder Rastergraphiken von ganzen Schriftträgern, aber auch auf Vektorgraphiken von einzelnen Schriftzeichen oder Zeichengruppen erstrecken. Das Programm, das Busch und Chandna vorstellen und das von Riffler-Pipka auf ihr Zeitschriftenkorpus angewendet wird, konzentriert sich auf Elemente der Seitengestaltung. Auch bei der Bildanalyse dieser Art sind Geisteswissenschaftlerinnen und Geisteswissenschaftler in ein Training der Software involviert. Rafiyenkos ›Tracing‹ basiert zwar nicht auf einer vergleichbar automatischen Bildanalyse, aber auf einer definierten Routine zur Bearbeitung der Abbildung von Palimpsesten, um die verdeckte Schrift sichtbar werden zu lassen.

Kodierung: Neben der automatischen Gewinnung von Daten spielt die Kodierung als Modellierung von Information in mehreren Vorhaben eine Rolle. Die Kodierung nicht-alphabetischer handgeschriebener Schriftzeichen bildet im Aufsatz von Gùlden, Krause und Verhoeven einen Schwerpunkt. Bei automatisch gewonnenen Daten wie im Projekt *eCodicology* sind die Speicherung der Ergebnisse und die Verknüpfung mit existierenden Beschreibungen Themen. Driscoll stellt in seinem Beitrag zu den isländischen Sagas einige Kodierungsbeispiele vor, darunter solche für Fragen der Seitengestaltung oder Textdichte. Im *Optical Neume Recognition Project* werden Kodierung und Zeichenerkennung aneinander ausgerichtet.

Normierung von Daten: Normdaten kommen in verschiedenen der dargestellten Projekten zum Einsatz. Zentral sind sie für das Anliegen von Beyer, Münkner, Schmidt und Steyer. Auch Driscoll nutzt normierte Daten für die Beschreibung und Auswertung der von ihm untersuchten Textgattung. Der Zusammenhang zwischen der Erforschung von Sammlungen und dem Anreiz, auf Normdaten zurückzugreifen, ist ebenso offensichtlich wie nachvollziehbar, da hier die Nutzung von etablierten Standards, die

Verknüpfung verschiedener Objekte, das Auffinden von Querverbindungen sowie die Prozessierbarkeit und Interpretierbarkeit der Daten zentrale Anliegen sind. *Schaßan* vergleicht bestimmte kontrollierte Vokabulare für die Schriftklassifikation und stellt diese zum Beispiel Taxonomien und Ontologien gegenüber.

Mikro- und Makroanalyse: Die Normierung der Daten ist oft eine Voraussetzung für statistische Auswertungen, wie dies in *Driscolls* Aufsatz deutlich wird. Dabei wird – wohl meist von geisteswissenschaftlicher Seite – betont und angestrebt, die Auswertung größerer Datenmengen in Form von Makroanalysen mit Mikroanalysen oder Einzelfallstudien zu verbinden, wie dies *Beyer*, *Münkner*, *Schmidt* und *Steyer* nach der Erschließung der privaten Büchersammlungen vorhaben. Mit dem im Beitrag von *Busch* und *Chandna* vorgestellten Werkzeug CodiVis ist eine Verbindung von Makroanalyse zum gesamten Handschriftenbestand und Mikroanalyse zu einzelnen Kodizes oder einzelnen Seiten möglich.

Visualisierung: Ein typisches und oft gewünschtes Ergebnis statistischer Auswertungen sind statische oder dynamische graphische Aufbereitungen der Ergebnisse in verschiedensten Formen. Auffällig ist dabei, dass dem Benutzer oft auch Möglichkeiten gegeben werden, mit den Visualisierungen zu interagieren oder diese mit Metadaten zu verknüpfen. So lassen sich sowohl bei CodiVis als auch bei VisColl Werte oder Gruppen von Werten auswählen, um so durch die Datenmengen zu navigieren.

Computergestützte Suche: Das sogenannte Information Retrieval dient der Suche nach komplexen Inhalten. *Vidal* präsentiert ein Modell, das zeigt, wie Indexierung und Suche ohne Transkription des Textes unmittelbar auf den Bildern selbst durchgeführt werden können. *Bogacz* und *Mara* stellen in ihrem Aufsatz eine graphische Suchmöglichkeit von Keilschriftzeichen vor, die auf Vektorgraphiken basiert und zugleich eine automatische Annotation gleicher Zeichen innerhalb des Textes ermöglicht.

Annotation: Eine weitere Form der Interaktion mit dem digitalisierten Material ist die Annotation. Der Ansatz von *Bogacz* und *Mara* beinhaltet ein entsprechendes Verfahren. Die Werkzeuge CodiVis und VisColl erlauben, automatisch oder selbst definierte Befunde mit einer Taxonomie zu verknüpfen. Auf diese Weise können entweder eine eigene Kategorisierung eingebracht und angewendet oder auch Daten mit extern definierten Standards wie Normdaten verbunden werden.

An verschiedenen Stellen wird deutlich, dass auch zwischen diesen Überschneidungen ihrerseits wieder Überschneidungen bestehen können. Ein Beispiel hierfür ist die Kombination von Kodierung und Zeichenerkennung. Einzelne technische Komponenten lassen sich miteinander oft auch in unterschiedlicher Reihenfolge zu einem Arbeitsablauf verbinden. Eine Verknüpfung mit Normdaten etwa kann über digitale Annotationen erfolgen und die Grundlage für statistische Analysen bilden. Diese Modularisierung digitaler Komponenten ermöglicht – bestenfalls – jeweils passende Kombinationen für verschiedene Arbeitsabläufe. Bei einigen digitalen Komponenten

ist zudem die Interaktion zwischen Geisteswissenschaftler und Software entscheidend. Nicht nur werden Programme oft so konzipiert, dass der Geisteswissenschaftler mit Darstellungen seiner Daten direkt arbeiten kann, in einigen Schritten wie den genannten ›Trainingseinheiten‹ sind informatische Kompetenz und handschriftenkundliches Wissen voneinander abhängig, um den Arbeitsschritt erfolgreich abschließen zu können. Die Software muss Ergebnisse liefern, die der Geisteswissenschaftler beurteilen und verwenden kann; die Ergebnisse der Software hängen aber auch von seiner Rückmeldung beim Training der Algorithmen ab. Supplementäre methodische Interdisziplinarität wie bei *eCodicology* ist also keine Seltenheit. Zusammengesetzte Interdisziplinarität mit kooperativen Momenten scheint sogar eher die Regel als die Ausnahme zu sein.

Bibliographie

- CodiLab. <<https://github.com/JochenGraf/CodiLab/blob/master/CodiKOS.html>>.
- DigiPal: *Digital Resource and Database for Palaeography, Manuscript Studies and Diplomatic*. <<http://www.digipal.eu>>.
- Hacking, Ian. »Verteidigung der Disziplin.« In Jungert, Michael u.a. (Hrsg.). *Interdisziplinarität. Theorie, Praxis, Probleme*. Darmstadt: Wissenschaftliche Buchgesellschaft, 2010. 193–206.
- Klein, Julie Thompson. »A taxonomy of interdisciplinarity.« In Frodeman, Robert (Hrsg.). *The Oxford Handbook of Interdisciplinarity*. Oxford: University Press, 2010. 15–30.
- Mittelstraß, Jürgen. »Die Stunde der Interdisziplinarität.« In Kocka, Jürgen (Hrsg.). *Interdisziplinarität. Praxis—Herausforderung—Ideologie*. Frankfurt am Main: Suhrkamp, 1987. 152–158.
- Mittelstraß, Jürgen. *Die Häuser des Wissens. Wissenschaftstheoretische Studien*. Frankfurt am Main: Suhrkamp, 1998.
- Sonderforschungsbereich 933: *Materiale Textkulturen. Materialität und Präsenz des Geschriebenen in non-typographischen Gesellschaften*. <<http://www.materiale-textkulturen.de>>.
- Sonderforschungsbereich 950: *Manuskriptkulturen in Asien, Afrika und Europa*. <<https://www.manuscript-cultures.uni-hamburg.de>>.
- Sonderforschungsbereich 980: *Episteme in Bewegung. Wissenstransfer von der Alten Welt bis in die Frühe Neuzeit*. <<http://www.sfb-episteme.de>>.