

Quanticod revisited.

Neue Ansätze zur quantitativen Analyse mittelalterlicher Handschriftenbestände

Hannah Busch (Universität Trier)

Swati Chandna (Karlsruher Institut für Technologie)

Celia Krause (Technische Universität Darmstadt)

Philipp Vanscheidt (Universität Trier / Technische Universität Darmstadt)

Die Kodikologie hat sich als Hilfsdisziplin der historischen Wissenschaften seit den fünfziger Jahren etabliert. Ihr Interesse gilt in erster Linie der Beschreibung und dem Studium der äußeren Merkmale des Kodex, darunter fällt neben den materiellen Bestandteilen auch das Studium des sogenannten *Mise-en-page* – des Layouts der handgeschriebenen Seite, der Größe und des Umfangs der gebundenen Handschrift.

Ziel von kodikologischen Studien ist es, mehr über die Arbeitstechniken und den Ablauf der mittelalterlichen Buchherstellung zu erfahren sowie regionale und soziokulturelle Normierungstendenzen im Hinblick auf ihre Existenz, räumliche Verbreitung und zeitliche Überdauerung zu erforschen. Ein besonderes Augenmerk gilt dabei dem Verhältnis zwischen der zur Verfügung stehenden Oberfläche einer Seite und der Fläche, die durch Primärtext, Bilder und Marginalia eingenommen wird.

Relativ schnell kristallisiert sich heraus, dass sich solche kodikologischen Fragestellungen weniger gut an einzelnen Objekten überprüfen lassen, es vielmehr quantitativer Verfahren bedarf. Dafür wird jedoch eine kritische Masse an Kodizes benötigt, die mindestens über eine gemeinsame Eigenschaft verfügen. Diese Eigenschaft kann sich in Form der Provenienz, wie der Zugehörigkeit zu einer mittelalterlichen Bibliothek oder einem Skriptorium als zugeschriebenem Entstehungsort, der Textgattung, der Sprache oder eines bestimmten Layouts manifestieren.

Quantitative Kodikologie

Bereits seit den siebziger Jahren verfolgte eine Gruppe von italienischen und französischen Wissenschaftlern¹ der noch relativ jungen Disziplin der Kodikologie diesen Ansatz einer quantitativen Forschungsmethode an mittelalterlichen Handschriften. Durch manuell erhobene Zähl- und Messergebnisse und deren statistische Auswertung konnten so Entwicklungslinien und -tendenzen auch graphisch dargestellt werden. Auf diese Weise lassen sich beispielsweise Aussagen über die Zeichendichte auf einer ein- oder zweispaltig aufgeteilten Handschriftenseite oder über die Bedeutung von Marginalflächen treffen und

¹ Als Mitglieder der Gruppe Quanticod seien hier namentlich zu erwähnen: Carla Bozzolo, Ezio Ornato, Denis Muzerelle, Dominique Coq.

räumlich und zeitlich bedingte Layouttendenzen feststellen. Durch geometrische Berechnungen kann erforscht werden, ob sich das Seitenverhältnis an Normen, wie dem aus der Malerei bekannten Goldenen Schnitt, orientiert. Durch den Aufwand der händischen Datenerhebung und die Verteilung der Handschriften auf viele verschiedene Standorte waren die Untersuchungen aber nur auf bereits erschlossene und zum Teil vermessene Handschriftenbestände und damit sehr limitierte Korpora anwendbar.

Der dargelegte Ansatz einer quantitativen Kodikologie scheint seit den frühen neunziger Jahren nicht intensiver verfolgt worden zu sein, obwohl das Potential computergestützter Analysemöglichkeiten seitdem stetig gewachsen ist. Zwei technische Neuerungen haben nämlich inzwischen den Umgang mit historischen Buchbeständen in Universitäten, Bibliotheken und Archiven nachhaltig verändert: die Entwicklung von handlichen Geräten für die Bilddatenerfassung sowie die Einrichtung des World Wide Web. Im Internet erschließen sich Publikationswege, die den Vorteil haben, dass Inhalte für einen breiten Nutzerkreis zur Verfügung gestellt werden können. Die zunehmende Verfeinerung digitaler Reproduktionsmöglichkeiten ermöglicht die Digitalisierung alter Buchbestände mittels Scannern und Digitalkameras und deren allgemeine Freigabe im Internet. Im Zuge großer Digitalisierungsprojekte sind mittlerweile viele mittelalterliche Handschriften in Archiven und Bibliotheken ins digitale Medium überführt worden. Schon lange werden dabei nicht mehr nur berühmte und reich ausgeschmückte Handschriften der breiten Öffentlichkeit zur Verfügung gestellt. Zahlreiche Projekte erfassen inzwischen den gesamten Bestand von Bibliotheken und Skriptorien. In Folge dessen geraten nun auch Handschriften ins Blickfeld, die bisher nicht mehrfach und detailliert katalogisiert und beschrieben wurden.

eCodicology

Genau an diesem Punkt setzt das Projekt „eCodicology“ (<http://www.ecodicology.org>) an. Als Korpus dient zunächst der digital verfügbare Bestand des rekonstruierten „Virtuellen Skriptorium St. Matthias“ (<http://www.stmatthias.uni-trier.de>). Dabei handelt es sich um circa 450 Handschriften aus dem mittelalterlichen Bestand der Benediktinerabtei St. Matthias in Trier aus dem 8. – 18. Jahrhundert. Insgesamt liegen etwa 170.000 Seiten als Bilddateien vor, die ein Datenvolumen von über 5 TB umfassen.

Durch die gezielte Vorverarbeitung der Bilddateien können deutlich größere Datenmengen bewältigt werden, ohne dabei einen Qualitätsverlust zu erleiden. Dies wird ermöglicht, indem eine Farb- und Größenkalibrierung, eine Skalierung und eine Reduzierung eventuellen Bildrauschens durchgeführt werden. Erst dann findet der Vorgang der Merkmalsextraktion statt, bei dem die einzelnen Layoutmerkmale der handgeschriebenen Seite erkannt, isoliert und vermessen werden. Zu diesen Merkmalen zählen die Seitengröße, Schrifträume, Bildräume, freie Flächen und die Anzahl der Zeilen. Neben der Anzahl der jeweiligen Elemente und ihrer Ausdehnung soll auch ihre exakte Position auf der Seite festgehalten werden. Die dadurch neu gewonnen Metadaten ergänzen die Beschreibungen aus früheren Katalogen und werden in XML-Dateien nach TEI P5 gespeichert.

Fallstudien „Die Beobachtung des Banalen“

Nachdem zunächst die Vorverarbeitung der Bilddateien im Vordergrund stand, können nun die ersten vollständig automatischen Vermessungen des Seitenlayouts präsentiert und statistisch ausgewertet werden. Anhand eines Abgleiches mit manuellen Messungen, die mit Lineal und Geodreieck an einigen (analogen) Originalhandschriften durchgeführt wurden sowie vorhandenen Informationen über die Maße aus Handschriftenkatalogen, werden wir damit erste Ergebnisse zur Genauigkeit der Algorithmen für die automatische Merkmalsextraktion vorstellen können. Hierbei wird es um die Frage gehen, mit welchen Herausforderungen der quantitativ arbeitende Kodikologe bei der Vermessung konfrontiert wird. Hierzu gehört etwa die Einschätzung des Einflusses der mit der Bindung einhergehenden Wölbung der Seite auf die Messergebnisse und der daraus resultierenden Vor- und Nachteile von semi- und vollautomatischen Messungen durch den Computer. Wie wirken sich die Erkenntnisse auf die Ansprüche und die bisherigen Konventionen der Digitalisierungsverfahren aus? Der Vortrag widmet sich darüber hinaus der Frage, welche Möglichkeiten der Analyse von digitalen Sammlungen uns durch eine digitale Kodikologie eröffnet werden. Außerdem soll diskutiert werden, welche neuen Fragestellungen sich damit aufwerfen lassen und wie sich die Ergebnisse aufbereiten und visualisieren lassen.

Neben der Ergänzung und Auswertung der neu gewonnenen Bildmetadaten bietet sich die statistische Auswertung auch für die bereits in früheren Katalogen erfassten bibliografischen Metadaten an. Dies ist eine Methode, die unter dem Begriff der *Bibliometrie* in der Buchwissenschaft bereits eine längere Tradition hat. Im Zuge der Erstellung virtueller Handschriftenbibliotheken wurden die Metadaten nicht nur in klassischen Datenbankstrukturen veröffentlicht. Durch Initiativen wie der TEI (Text Encoding Initiative) ist ein XML-Schema entwickelt worden, das sich auch speziell an die Erfassung der Metadaten von Handschriften richtet. Dieses Schema findet auch im Projekt „eCodicology“ in adaptierter Form Verwendung.

Zunächst erschien es sinnvoll, die alten Daten aus den Handschriftenkatalogen und die neugewonnenen Daten zu den Maßen der Layoutmerkmale gemeinsam statistisch auszuwerten, also zu den Wurzeln der quantitativen Kodikologie zurückzukehren. Darüber hinaus wollen wir uns in einer kleinen Fallstudie die gemeinsame Art der Aufbereitung der Daten zunutze machen, um den Bestand der Abtei St. Matthias mit einem weiteren, in XML aufbereiteten Handschriftenbestand zu vergleichen: neben einem quantitativen wird somit auch ein komparatistischer Ansatz verfolgt. In einem ersten Schritt sollen dabei Katalogdaten, wie Blatt-/Seitenzahl und Größe der Kodizes anderer Bestände mit dem Skriptorium von St. Matthias verglichen werden. So ließen sich Aussagen über regionale Tendenzen treffen und beispielsweise Vergleiche zum norddeutschen Raum oder dem benachbarten romanischen Raum ziehen.

Die neuen Möglichkeiten einer digital betriebenen Handschriftenkunde können um ein weiteres kodikologisches Werkzeug bereichert werden, das es erlaubt, einzelne Exemplare aus einem Handschriftenbestand oder größere Handschriftengruppen im wahrsten Sinne des Wortes aus der Vogelperspektive zu betrachten. Hierbei handelt es sich um die Erstellung von Bildmontagen nach der Methode von Lev Manovich. Ein Kodex wird virtuell entblättert und alle Seiten der Reihe nach rasterförmig neben- bzw. untereinander angeordnet. Eine solche Bildmontage erlaubt das Aufspüren von Gestaltungsmustern innerhalb repräsentativer Handschriften.

Die Kombination aller vorgestellten Verfahren gestattet dem Kodikologen einen neuartigen Zugang zu digital erschlossenen Handschriftenbeständen. Durch die Untersuchung scheinbar banaler Merkmale lassen sich Forschungsfragen beantworten, auf die allein in Detailstudien kaum eingegangen werden könnte.